SECURITY CLASSIFICATION OF THIS PAGE

AD-A248 681

REPORT

DTIC ELECTE APR 16 1992 S D

1a. REPORT SECURITY CLASSIFICATION
Unclassified

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
University of Notre Dame
Department of Physics

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
Office of Naval Research
Resident Representative

6c. ADDRESS (City, State, and ZIP Code)
Notre Dame, Indiana 46556

7b. ADDRESS (City, State, and ZIP Code)
536 S. Clark Street Room 286
Chicago, Illinois 60605-1588

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Office of Naval Research

8b. OFFICE SYMBOL (If applicable)
1114 SC/ONR

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-84-K-0352

8c. ADDRESS (City, State, and ZIP Code)
800 N. Quincy Street
Arlington, Virginia 22217-5000

10. SOURCE OF FUNDING NUMBERS

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

11. TITLE (Include Security Classification)
Theory of Semiconducting Superlattices and Microstructures

12. PERSONAL AUTHOR(S)
Dow, John D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 84-04-16 TO 90-09-30	92-03-01	435

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Semiconductor theory, Superlattices, Deep-level structures IV-VI compounds, PbTe, SnTe, HEMT Structures, local-density theory.

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

This final report summarizes theoretical work on semi conducting superlattices and microstructures. A focus of the work is a fundamental understanding of deep-level structures of impurities in semiconductors.

(Abstract continued on following page)

92-09782

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

21 ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code)	22c OFFICE SYMBOL
George B. Wright	(703)696-4202	114 SS/ONR

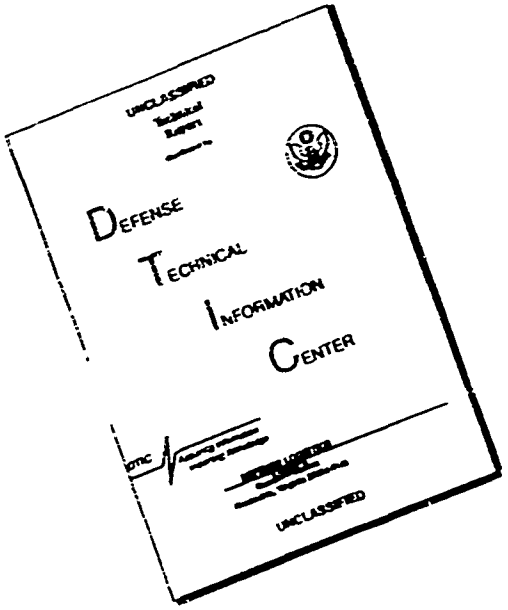
DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

*U.S. Government Printing Office: 1985-507-047

92 4 15 052

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

ABSTRACT

19. Abstract (continued)

Probably the most striking results in this area concern IV-VI semiconductors such as PbTe and SnTe, where we have originated the concept of false valence: In on a Te site in SnTe is a (triple) acceptor and has normal valence of -3 with respect to Te. But on a Te site in PbTe, In is a (triple) donor with a false valence of +3 with respect to Te because an In deep p-like one-electron level capable of trapping six electrons crosses the gap as x decreases in $Pb_{1-x}Sn_xTe$. This theory predicts that In is a donor in PbTe but an acceptor in SnTe; it argues that in IV-VI semiconductors impurities often occupy the intuitively "wrong" site or antisite; it shows that the relevant In occupies Te sites; it explains low doping efficiencies; and it shows promise for explaining solidus curves.

This work on InN illustrates how our theory can be useful for the fabrication and doping of new electronic materials. The recent development of high-mobility InN raises the possibility of making UV sensors and even lasers out of this material. We have presented global predictions on the behavior of s- and p-bonded deep levels in InN, $In_{1-x}Ga_xN$, and $In_{1-x}A_xN$ as a means of guiding experiments aimed at doping these materials for use in UV and glue-green optical devices.

This contribution to the theory of Schottky barriers by developing the theory of defects at surfaces and at interfaces. Soon we shall publish the first theory of scanning tunneling microscope images of surface antisite defects in GaAs. This theory shows that unless the microscope bias is set correctly, an antisite defect can appear to be the same as the atom it replaced! We have also investigated the role of plastic flow in the formation of scanning tunneling microscope tips by electrochemical etching.

This work has laid the foundation for treating degenerate Fermi gases in alloys, and including multi-electron relaxation effects (i.e., infrared divergence phenomena) in evaluation of their optical properties. While this approach has a way to go before being easily implemented for realistic models, it is far ahead of its main competitor, the renormalization group method.

The electronic properties of $Ge_{1-x}Sn_x$ have been worked out, as a guideline for experimental groups attempting to grow these interesting materials. The physics of order-disorder transformations in metastable alloys such as $(GaAs)_{1-x}Ge_{2x}$ has been elucidated (in collaboration with K. E. Newman) and new materials, such as $(InP)_{1-x}Ge_{2x}$, have been predicted as likely to have interesting infrared properties. The relationships of random alloys with partially ordered phases (e.g., zincblende, superlattice, chalcopyrite, etc.) have been elucidated, and a basic quasi-equilibrium

approach to understanding their electronic, vibrational, and thermodynamic properties has been developed.

This effort has produced the first comprehensive theory of deep levels in [001] and [111] III-V superlattices. This theory elucidated the various factors affecting deep levels, sets forth the conditions for obtaining shallow-deep transitions, and predicts that Si (a common donor making GaAs n-type) becomes a deep trap in GaAs/A $_x$Ga$_{1-x}$As superlattices with thin quantum wells. These results have bearing on HEMT structures.

We have developed the pseudo-function method of local-density theory, including the spin-unrestricted version. We have treated the hypothetical material MnTe with spin-unrestricted pseudo-function theory and found an antiferromagnetic ground state and conduction bands that are highly spin polarized. We have obtained the results of earlier theories by restricting the amount of spin-polarization; hence we have ascribed the differences between our results and earlier theories to inadequate basis sets of the earlier theories. We are currently exploring applications of our methods to other problems in magnetism, and have some optimism that our spin-dependent pseudo-function approach may overcome the traditional problems of the local-density theory of magnetic systems.

The pseudo-function method is computationally fast by local-density standards, and we have used it to determine equilibrium surface relaxations of semiconductors by minimizing the total energies of the surfaces. We have also looked at many possible oxygen adsorption sites on GaAs (110) to find the ones with the lowest energy, in order to understand the oxidation process. The pseudo-function method gives the same answers to the classic problems, such as the bulk modulus of Si, as the other successful implementations of local density theory.

Accesion For
NTIS CRA&I
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1 | 23



OFFICE OF NAVAL RESEARCH

FINAL REPORT

for

Contract N00014-84-K-0352

Task No. NR 616-027

Theory of semiconductor superlattices and microstructures

John D. Dow
Freimann Professor of Physics
Department of Physics
University of Notre Dame
Notre Dame, Indiana 46556

March 1, 1992

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

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

THEORY OF SEMICONDUCTING SUPERLATTICES AND MICROSTRUCTURES

John D. Dow
Department of Physics
University of Notre Dame
Notre Dame, Indiana 46556

A. Impurity levels in bulk materials

One of the long-term interests of our group has been to understand the deep level structures of impurities in semiconductors. We have continued working in this area because the problems are interesting and challenging, while often being well-suited to training graduate students.

1. IV-VI semiconductors

Probably the most striking results in this area concern IV-VI semiconductors such as PbTe and SnTe, where we have originated the concept of false valence: In on a Te site in SnTe is a (triple) acceptor and has normal valence of -3 with respect to Te. But on a Te site in PbTe, In is a (triple) donor with a false valence of +3 with respect to Te because an In deep p-like one-electron level capable of trapping six electrons crosses the gap as x decreases in $Pb_{1-x}Sn_xTe$. This theory predicts that In is a donor in PbTe but an acceptor in SnTe; it argues that in IV-VI semiconductors impurities often occupy the intuitively "wrong" site or antisite; it shows that the relevant In occupies Te sites; it explains low doping efficiencies; and it shows promise for explaining solidus curves.

2. InN-related materials

Our work on InN illustrates how our theory can be useful for the fabrication and doping of new electronic materials. The recent development of high-mobility InN raises the possibility of making UV sensors and even lasers out of this material. We have presented global predictions on the behavior of s- and p-bonded deep levels in InN, $In_{1-x}Ga_xN$, and $In_{1-x}A\ell_xN$ as a means of guiding experiments aimed at doping these materials for use in UV and blue-green optical devices.

3. Charge-state splittings and meso-bonding molecular defects

It has been our position for years that large self-consistent pseudopotential calculations of deep levels are rarely necessary for substitutional s- and p-bonded impurities in Si and III-V semiconductors. This position, although opposed by significant segments of the theoretical physics community, has been vindicated by our self-consistent calculations for S in Si.

A controversy about the character of the S_2 substitutional molecular defect in Si has been resolved in favor of a meso-bonding picture, based on our self-consistent calculations of the charge-state splittings and hyperfine tensor in this system.

4. Effects of hydrostatic pressure and uniaxial stress on deep levels

We have shown that if both the energy of a deep level and its (hydrostatic or uniaxial) pressure dependence are known, one can almost uniquely determine the substitutional defect responsible. The determination is often not unique, but reduces the candidate s- and p-bonded impurities to only a few.

B. Surface defects

We have contributed to the theory of Schottky barriers by developing the theory of defects at surfaces and at interfaces. Soon we shall publish the first theory of scanning tunneling microscope images of surface antisite defects in GaAs. This theory shows that unless the microscope bias is set correctly, an antisite defect can appear to be the same as the atom it replaced! We have also investigated the role of plastic flow in the formation of scanning tunneling microscope tips by electrochemical etching.

C. Pseudo-function local-density theory

We have developed the pseudo-function method of local-density theory, including the spin-unrestricted version. We have treated the hypothetical material MnTe with spin-unrestricted pseudofunction theory and found an antiferromagnetic ground state and conduction bands that are highly spin polarized. We have obtained the results of earlier theories by restricting the amount of spin-polarization; hence we have ascribed the differences between our results and earlier theories to inadequate basis sets of the earlier theories. We are currently exploring applications of our methods to other problems in magnetism, and have some optimism that our spin-dependent pseudo-function approach may overcome the traditional problems of the local-density theory of magnetic systems.

The pseudo-function method is computationally fast by local-density standards, and we have used it to determine equilibrium surface relaxations of semiconductors by minimizing the total energies of the surfaces. We have also looked at many possible oxygen adsorption sites on GaAs (110) to find the ones with the lowest energy, in order to understand the oxidation process. The pseudo-function method gives the same answers to the classic problems, such as the bulk modulus of Si, as the other successful implementations of local density theory.

D. Phonons in alloys

We have developed the theory of phonons in alloys and shown how to treat correlated alloys by combining Ising Monte Carlo and Recursion methods.

E. Heavily doped semiconductor alloys

We have laid the foundation for treating degenerate Fermi gases in alloys, and including multi-electron relaxation effects (i.e., infrared divergence phenomena) in evaluation of their optical properties. While our approach has a way to go before being easily implemented for realistic models, it is far ahead of its main competitor, the renormalization group method.

F. Metastable alloys

The electronic properties of $Ge_{1-x}Sn_x$ have been worked out, as a guideline for experimental groups attempting to grow these interesting materials. The physics of order-disorder transformations in metastable alloys such as $(GaAs)_{1-x}Ge_{2x}$ has been elucidated (in collaboration with K. E. Newman) and new materials, such as $(InP)_{1-x}Ge_{2x}$, have been predicted as likely to have interesting infrared properties. The relationships of random alloys with partially ordered phases (e.g., zincblende, superlattice, chalcopyrite, etc.) have been elucidated, and a basic quasi-equilibrium approach to understanding their electronic, vibrational, and thermodynamic properties has been developed.

G. Special points for superlattices

We have worked out the special points (for $\vec{k}$-space sums) for [001] and [111] superlattices, and have shown that by taking advantage of hidden symmetries for certain superlattice period-ratios the number of special points is comparable with the number needed to obtain corresponding accuracy for bulk zincblende.

H. Structure-modulated superlattices

We have originated (we believe) the concept of a superlattice whose adjacent layers are the same material but with different structures, and contrasted this type of superlattice with conventional composition-modulated superlattices. An example is a zincblende/wurtzite superlattice which is lattice-matched, yet has very interesting topology. We have calculated the electronic structures and deep levels in some such superlattices.

I. Fanipi doping superlattices

We have developed the concept of false-valence doping nipi superlattices. Such a superlattice could be formed by uniformly doping PbTe/SnTe superlattices with In. The In is a donor in PbTe because of a false valence. It is an acceptor in SnTe because of normal valence. (The relevant In occupies a Te site.)

J. Deep levels in [001] and [111] superlattices

We have produced the first comprehensive theory of deep levels in [001] and [111] III-V superlattices. This theory elucidates the various factors affecting deep levels, sets forth the conditions for obtaining shallow-deep transitions, and predicts that Si (a common donor making GaAs n-type) becomes a deep trap in $GaAs/A\ell_xGa_{1-x}As$ superlattices with thin quantum wells. Thus the common n-type dopant becomes a deep trap in ultra-small devices! These results have bearing on HEMT structures.

K. Effects of band offsets and applied pressure on deep levels in superlattices

Since band offsets are generally poorly known, except for $GaAs/A\ell_xGa_{1-x}As$ superlattices, we have explored the dependences of deep level energies on the offsets -- and found them to be small. We have also studied the pressure dependences of the electronic structures of and deep

levels in superlattices.

L. Strain-induced Type II - Type I transitions in ZnSe/ZnTe

We have shown that by suitably choosing the periods of ZnSe and ZnTe in a ZnSe/ZnTe superlattice to obtain appropriate strain fields, it is possible to drive the structure from Type II to Type I. This raises the possibility of fabricating blue-green lasers from such superlattices -- if the doping problem can be solved.

M. Si-based luminescence

We have shown that with novel combinations of superlattice growth direction (such as [111]), band-folding, superlattice periods, and uniaxial stress, it is always possible to drive a Si/Ge superlattice direct and make it capable of emitting luminescence. We have also calculated dielectric functions of such microstructures to predict the luminescent intensities.

Publications

138. D. V. Froelich, M. E. Lapeyre, J. D. Dow, and R. E. Allen. Dependence of the GaAs (110) surface electronic state dispersion curves on the surface relaxation angle.. Superlatt. Microstruct. 1, 87-89 (1985).

146. C. S. Lent, M. A. Bowen, J. D. Dow, R. S. Allgaier, O. F. Sankey, and E. S. Ho. Relativistic empirical tight-binding theory of the energy bands of GeTe, SnTe, PbTe, PbSe, PbS, and their alloys. Superlatt. Microstruct. 2, 491-499 (1986).

156. C. S. Lent, M. A. Bowen, R. S. Allgaier, J. D. Dow, O. F. Sankey, and E. S. Ho. Impurity levels in PbTe and $Pb_{1-x}Sn_xTe$. Solid State Commun. 61, 83-87 (1987).

164. S. F. Ren and J. D. Dow. A mechanism of luminescence enhancement by classical-well structures or superlattices. J. Luminescence 33, 103-107 (1985).

165. M. A. Bowen, R. E. Allen, and J. D. Dow. Surface defects and core excitons at the (2×1) asymmetric-dimer (100) surface of Si. Phys. Rev. B 30, 4617-4620 (1984).

167. J. D. Dow, R. E. Allen, and O. F. Sankey. Intrinsic and extrinsic surface electronic states of semiconductors. Chem. and Phys. of Solid Surfaces, V, 483-500 (1984), ed. by R. Vanselow and R. Howe, Springer Series in Chemical Physics 35, (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo).

169. D. V. Jenkins and J. D. Dow. Electronic properties of metastable Ge_xSn_{1-x} alloys. Phys. Rev. B 36, 7994-8000 (1987).

170. R. E. Allen, T. J. Humphreys, J. D. Dow, and O. F. Sankey. Theory of surface defect states and Schottky barrier heights: Application to

InAs. J. Vac. Sci. Technol. B2, 449-452 (1984).

171. O. F. Sankey, R. E. Allen, and J. D. Dow. Theory of Si/transition-metal silicide Schottky barriers. J. Ultramicroscopy 14, 127-130 (1984).

173. M. A. Bowen and J. D. Dow. Effects of diagonal Anderson disorder on X-ray emission and photoemission spectra. Phys. Rev. B 30, 6199-6201 (1984).

174. M. A. Bowen and J. D. Dow. Off-diagonal disorder and X-ray edges. J. Phys. Chem. Solids 46, 79-81 (1984).

175. D. V. Froelich, M. A. Bowen, and J. D. Dow. Electronic states for the (100) (2x1) reconstructed Ge surface. J. Vac. Sci. Technol. B2, 390-395 (1984).

178. J. D. Dow. Localized perturbations in semiconductors. In Highlights of Condensed-Matter Theory (Proceedings of the International School of Physics "Enrico Fermi", Course 89, Varenna, 1983), ed. by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna, Italy, and North Holland, Amsterdam, 1985), pp. 465-494.

179. M. A. Bowen and J. D. Dow. X-ray spectra of model binary alloys $A_{1-x}B_x$. Phys. Rev. B 32, 5119-5123 (1985).

180. S. Lee and J. D. Dow. Electronic structure of $Pb_{1-x}Sn_xTe$ semiconductor alloys. Phys. Rev. B 36, 5968-5973 (1987).

182. O. F. Sankey, R. E. Allen, S. F. Ren, and J. D. Dow. Dangling bonds and Schottky barriers. J. Vac. Sci. Technol. B 3, 1162-1166 (1985).

183. D. W. Jenkins, K. E. Newman, and J. D. Dow. Predicted energy band gaps of $(A^{III}B^{V})_{1-x}(X^{IV})_{2x}$ metastable, substitutional, crystalline alloys. Phys. Rev. B 32, 4034-4041 (1985).

184. O. F. Sankey, R. E. Allen, and J. D. Dow. Theory of Schottky barriers for III-V and Group-IV semiconductors. Proceedings of the 17th Internatl. Conf. Phys. Semiconductors, ed. D. J. Chadi and W. A. Harrison (Springer-Verlag, New York, 1985) pp. 189-192.

185. J. D. Dow, O. F. Sankey, and R. E. Allen. Interfacial deep levels responsible for Schottky barrier formation at semiconductor/metal contacts. Appl. Surf. Sci. 22/23, 937-947 (1985).

186. A. Kobayashi, K. E. Newman, and J. D. Dow. Densities of phonon states for $(GaSb)_{1-x}Ge_{2x}$. Phys. Rev. B 32, 5312-5327 (1985).

187. J. D. Dow, O. F. Sankey, and R. E. Allen. Chemical trends of Schottky barriers. Materials Science Forum 4, 39-50 (1985). (Trans Tech Publications Ltd., Switzerland).

188. S. Y. Ren, J. D. Dow, and J. Shen. Deep impurity levels in semiconductor superlattices. Phys. Rev. B 38, 10677-10692 (1988).

189. A. Kobayashi, J. D. Dow, and E. P. O'Reilly. Phonons in $A\ell_xGa_{1-x}As$ alloys. Superlatt. Microstruct. 1, 471-9 (1985).

190. G. Kim, J. D. Dow, and S. Lee. Theory of charge-state splittings of deep levels associated with sulfur pairs in Si. Phys. Rev. B 40, 7888-7891 (1989).

191. K. E. Newman, J. D. Dow, A. Kobayashi, and R. Beserman. Explanation of the anomalous Raman spectra of $(GaSb)_{1-x}Ge_{2x}$. Solid State Commun. 56, 553-556 (1985).

192. K. E. Newman, J. D. Dow, B. A. Bunker, L. L. Abels, P. M. Raccah, S. Ugur, D. Z. Xue, and A. Kobayashi. Effects of a zincblende-diamond order-disorder transition on the crystal, electronic, and vibrational structures of metastable $(GaAs)_{1-x}(Ge_2)_x$ alloys. Phys. Rev. B 39, 657-662 (1989).

193. M.-H. Tsai, D. W. Jenkins, J. D. Dow, and R. V. Kasowski. Pseudofunction theory of the electronic structure of InN. Phys. Rev. B 38, 1541-1543 (1988).

194. J. D. Dow. Physics of deep levels. Proc. 1985 Materials Research Soc. 46, 71-81 (1985), Microscopic Identification of Electronic Defects in Semiconductors, ed. N. M. Johnson, S. G. Bishop, and G. D. Watkins.

195. R. E. Allen, O. F. Sankey, and J. D. Dow. Theoretical interpretation of Schottky barriers and Ohmic contacts. Proc. Internatl. Conf. Formation of Semiconductor Interfaces, Marseilles, 1985, ed. G. Le Lay, Surf. Sci. 168, 376-385 (1986).

196. J. D. Dow. Supercomputers in solid state physics. In "Large Scale Computational Device Modeling" Ed. K. Hess, (Co-ordinated Science Laboratory, University of Illinois, Urbana, Illinois, 1986) p.83-89.

198. R. P. Beres, P E. Allen, and J. D. Dow. Effects of the band offset on interfacial deep levels. J. Mater. Res. 3, 164-166 (1988).

199. D. W. Jenkins, R.-D. Hong, and J. D. Dow. Band structure of InN. Superlatt. Microstruct. 3, 365-369 (1987).

203. M. A. Bowen and J. D. Dow. X-ray spectra of ordered and disordered binary compounds. Intl. J. Modern Phys. B 2, 1483-1493 (1988).

206. R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow. Defects in superlattices under pressure. Materials Research Soc. Symp. Proc. 77, 545-550 (1987), Interfaces, Superlattices, and Thin Films, ed. J. D. Dow and I. K. Schuller.

207. A. C. Redfield and J. D. Dow. Theory of Raman spectra of correlated substitutional alloys. Solid State Commun. 64, 431-434 (1987).

211. R. Nicolaides, Y. Liang, W. E. Packard, Z.-W. Fu, H. A. Blackstead, K. K. Chin, J. D. Dow, J. K. Furdyna, W. M. Hu, R. C. Jaklevic, W. J. Kaiser, A. R. Pelton, M. V. Zeller, and J. Bellina, Jr. Scanning

tunneling microscope tip structures. J. Vac. Sci. Technol., A6, 445-447 (1988).

213. D. W. Jenkins and J. D. Dow. Electronic structures and doping of InN, $In_{1-x}Ga_xN$, and $In_{1-x}A\ell_xN$. Phys. Rev. B 39, 3317-3329 (1989).

214. S. Y. Ren and J. D. Dow. Hydrogenated Si clusters: Band formation with increasing size. Phys. Rev. B in press.

216. R. D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow. Hydrostatic pressure dependencies of deep impurity levels in zinc-blende semiconductors. Phys. Rev. B 38, 12549-12555 (1988).

217. D. W. Jenkins, S. Y. Ren, and J. D. Dow. Dependence on uniaxial stress of deep levels in III-V compound and group-IV elemental semiconductors. Phys. Rev. B 39, 7881-7894 (1989).

220. S. Y. Ren and J. D. Dow. Special points for superlattices and strained bulk semiconductors. Phys. Rev. B 38, 1999-2001 (1988).

221. G. Kim, J. D. Dow, and S. Lee. Sulfur nearest-neighbor defect pairs in Si. Arabian J. Sci. Engineer. 14, 513-522 (1989).

224. K. E. Newman, J. Shen, D. Teng, B.-L. Gu, S. Y. Ren, and J. D. Dow. Electronic and structural properties of ordered III-V alloys. Alloy Phase Stability, ed. by G. M. Stocks and A. Gonis, Proceedings of the NATO Advanced Study Workshop on Band Structure Energies in Semiconductors, (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1989), pp. 621-625.

225. J. D. Dow, S. Y. Ren, and J. Shen. Deep impurity levels in semiconductors, semiconductor alloys, and superlattices. NATO Advanced Science Institutes Series B 183: Properties of Impurity States in Superlattice Semiconductors, (Plenum Press, New York, 1988), pp. 175-187, edited by C. Y. Fong, Inder P. Batra, and S. Ciraci.

228. S. Y. Ren and J. D. Dow. Electronic structure and deep impurity levels in [111] GaAs/AℓAs semiconductor superlattices. J. Appl. Phys. 65, 1987-1995 (1989).

231. A. Franciosi, A. Wall, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, R. V. Kasowski, R. Reifenberger, and F. Pool. d-states, exchange splittings and Mn electronic configuration in $Cd_{1-x}Mn_xTe$. Phys. Rev. B40, 12009-12012 (1989).

233. S. Y. Ren and J. D. Dow. Electronic structure and deep impurity levels in structure-modulated zinc-blende--wurtzite semiconductor superlattices. Phys. Rev. B 39, 7796-7802 (1989).

235. M.-H. Tsai, J. D. Dow, K. E. Newman, and R. V. Kasowski. Theory of local bond-length relaxation in $Hg_{1-x}Cd_xTe$ alloys. Phys. Rev. B 41, 7744-7748 (1990).

236. M.-H. Tsai, J. D. Dow, R. V. Kasowski, A. Wall, and A. Franciosi.

Explanation of anomalous Curie constants of $Cd_{1-x}Mn_xTe$ alloys. Solid State Commun. 69, 1131-1133 (1989).

238. A. Wall, A. Franciosi, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, and R. V. Kasowski. Inverse photoemission and resonant photoemission characterization of semimagnetic semiconductors. J. Vac. Sci. Technol. A7, 656-662 (1989).

239. S. Y. Ren, J. D. Dow, and S. Klemm. Strain-assisted p-type doping of II-VI semiconductors. J. Appl. Phys. 66, 2065-2068 (1989).

241. J. D. Dow, R.-D. Hong, S. Klemm, S. Y. Ren, M.-H. Tsai, O. F. Sankey, and R. V. Kasowski. Proposed explanation of the p-type doping proclivity of ZnTe. Phys. Rev. B in press.

242. R. V. Kasowski, M.-H. Tsai, J. D. Dow, and T. N. Rhodin. Structural and electronic properties of a CO overlayer on Ni (001). Phys. Rev. B, Submitted.

244. W.-M. Hu and J. D. Dow. Scanning tunneling microscope images of native defects on the ZnSe (110) surface. J. Vac. Sci. Technol. B7, 907-909 (1989).

246. R. V. Kasowski, M.-H. Tsai, S. T. Chui, and J. D. Dow. Electronic structure of TTF-TCNQ. Phys. Rev. B, Submitted.

248. R.-D. Hong and J. D. Dow. Doping $Zn_{1-x}Mn_xSe$ n-type. Appl. Phys. Lett. 54, 2597-2598 (1989).

249. J. Shen, J. D. Dow, and S. Y. Ren. Thin quantum-well superlattices of GaAs and $(GaAs)_{1-x}(ZnSe)_x$ with ZnSe: Possibility of band gaps in the blue-green. J. Appl. Phys. 67, 3761-3763 (1990).

250. R. V. Kasowski, M.-H. Tsai, and J. D. Dow. Self-consistent antiferromagnetic ground state for La_2CuO_4 via energy band theory. Phys. Rev. B 41, 8949-8954 (1990).

251. R. V. Kasowski, M.-H. Tsai, and J. D. Dow. Formation of Schottky barriers by metal-metal bonds on GaAs. Phys. Rev. B, Submitted.

APPENDIX

The referenced technical reports generated as a part of this report are included in the Appendix.

DEPENDENCE OF THE GaAs (110) SURFACE ELECTRONIC STATE DISPERSION CURVES ON THE SURFACE RELAXATION ANGLE

David V. Froelich, Mary E. Lapeyre, and John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556
and
Roland E. Allen
Department of Physics, Texas A&M University
College Station, Texas 77843

(Received 16 July 1984)

The surface state dispersion curves $E(\bar{k})$ of the dangling bond states near the fundamental band gap, C_3 and A_5, are computed for both the established $\theta=27^\circ$ model and the recently proposed $\theta=7^\circ$ model of the (110) surface relaxation of GaAs, where θ is the surface bond rotation angle. The two models produce surface state dispersion curves that are similar to one another and to the data.

Until recently it was thought that the geometrical structure of the (110) surface of GaAs was one of the few semiconductor surface structures that was established. The accepted model was the 27° rotation model [1,2]: To a good approximation, the anions rotate rigidly out of the surface through an angle of $\theta=27^\circ$. This model was established as a result of careful analyses of low-energy electron diffraction (LEED) data, and, in addition, provided a way out of a theoretical dilemma: calculations of GaAs surface states for unrelaxed surfaces, $\theta=0^\circ$, produced surface states in the fundamental band gap (contrary to data) that receded into the valence and conduction bands when the $\theta=27^\circ$ relaxation was accounted for [3].

Recently, however, Gibson and co-workers [4] have suggested that $\theta=7^\circ$ may be a more appropriate relaxation angle, based on analyses of Rutherford back-scattering (RBS) data. Duke and co-workers have also presented analyses of LEED data that indicate that a 7° rotation, while not preferred, is acceptable [5]. Gibson et al. have stated, however, that their data might be consistent with the $\theta=27^\circ$ model, provided one allows for anomalously large surface phonon amplitudes.

With LEED and RBS analyses producing ambiguous interpretations of the data, we thought it might be useful to determine if the measured surface state dispersion curves $E(\bar{k})$ [6], when compared with theoretical predictions, preferred either the $\theta=7^\circ$ model or the $\theta=27^\circ$ model. Previous calculations of $E(\bar{k})$, assuming the $\theta=27^\circ$ model, were in sufficiently good agreement with the data to afford explanations of the principal experimental features [7] (Fig.1). However, we now find qualitatively similar dispersion relations $E(\bar{k})$ for the relevant surface states, for $\theta=0^\circ$, $\theta=7^\circ$, 14° and 27° (Fig. 1). Since the theory is only accurate to several tenths of an eV [8] near the valence band maximum, the theoretical surface state dispersion curves do not provide a means for discriminating with confidence among the relaxation models. The theory does predict that surface states do fall in the fundamental band gap for the 7° model [1] (Fig. 2 [9] and Ref. [10]): 0.1 eV below the conduction band edge and 0.1 eV above the valence band maximum -- but these energies are too small in comparison with the several tenths of an eV theoretical uncertainty to be convincing proof of the 27° model over the 7° model.

Hence we conclude that the agreement between photoemission data and the theory does not provide strong evidence for or against either the $\theta=7^\circ$ model or the $\theta=27^\circ$ model. The established $\theta=27^\circ$ model should be retained until more conclusive experimental evidence against it is presented.

Finally, as we have been completing this manuscript, we have received a preprint from Mailhiot, Duke and Chang [11], who have independently been studying this problem using the same Hamiltonian and comparably accurate theoretical techniques. They have found similar results; however they interpret their results as providing stronger support for the $\theta=27^\circ$ model.

Acknowledgments -- We are grateful to the U.S. Army Research Office and the Office of Naval Research for their support (Contract Nos. DAAG-29-83-K-0122 and N0001+-82-K-0447).

1982 PACS Number: 68.20.+t; 73.20.-r

0749-6036/85 010087-03 $02.00/0

© 1985 Academic Press Inc (London) Limited

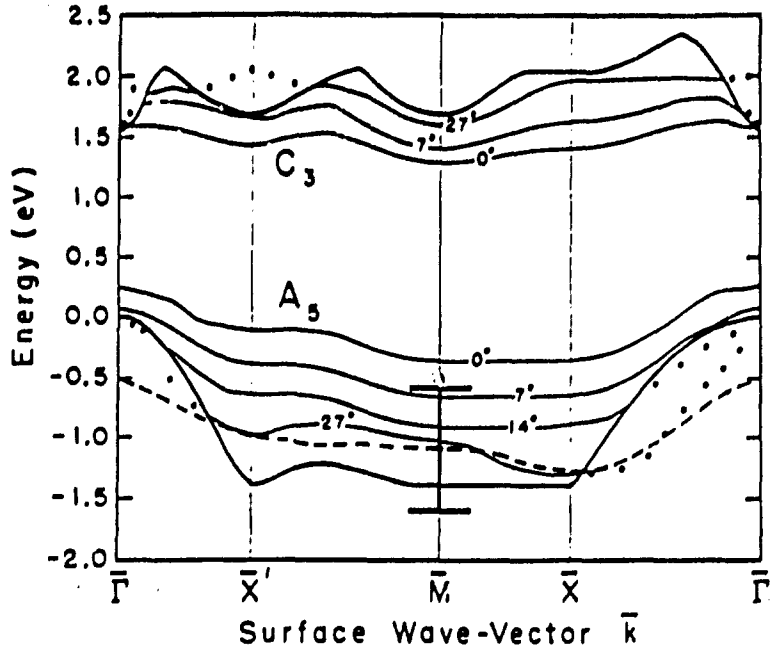


Fig. 1. Calculated surface state dispersion curves in the gap (solid lines) and resonant with the bulk bands (dotted lines), assuming a surface bond rotation angle of θ=0°, 7°, 14°, and 27° for the dangling bond (110) surface states C_3 and A_5 of GaAs. Heavy solid lines represent the bulk band edges; the dashed line denotes the data of A. Huijser, J. van Laar and T. L. van Rooy, Phys. Lett. 65A, 337 (1978) and G. P. Williams, R. J. Smith and G. J. Lapeyre, J. Vac. Sci. Technol. 15, 1249 (1978). The C_3 state is not shown for θ=14°, because this state lies too close to the θ=7° and 27° states. The 27° results are the same as those of Ref. [7]. The absolute uncertainty in the theoretical predictions is shown by the error bar at $\bar{\Gamma}$ centered on the data.

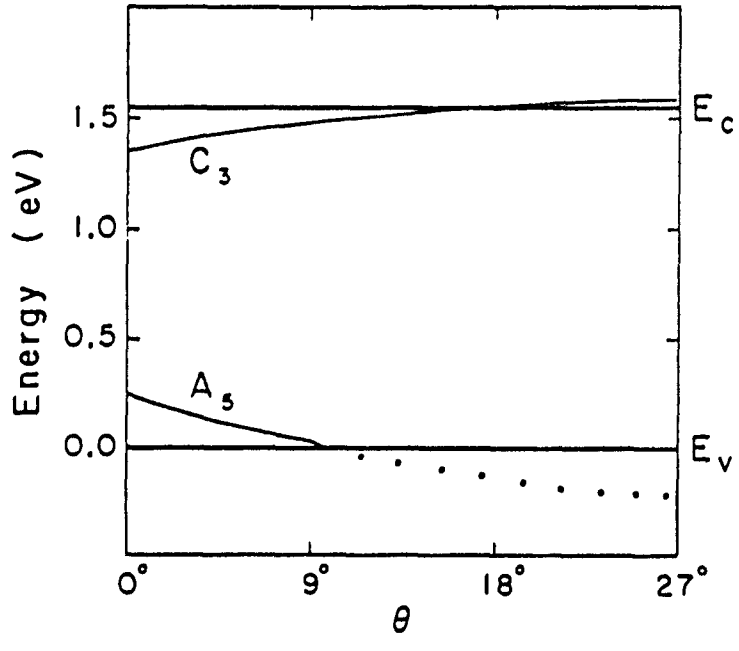


Fig. 2. Calculated energies of the dangling bond (110) surface states of GaAs (solid lines for bound states, dotted lines for resonances) at $\bar{\Gamma}$ [9], as functions of the surface bond rotation angle. The results for the C_3 state are from Ref. [10]. E_v and E_c are the valence and conduction band edges, denoted by heavy solid lines.

REFERENCES

[1] S. Y. Tong, A. R. Lubinsky, B. J. Mrstik and M. A. van Hove, Phys. Rev. B17, 3303 (1978).

[2] A. Kahn, E. So, P. Mark and C. B. Duke, J. Vac. Sci. Technol. 15, 580 (1978); C. B. Duke, R. J. Meyer and P. Mark, J. Vac. Sci. Technol. 17, 971 (1980).

[3] For a sampling of previous theories of the GaAs (110) surface states, see C. Calandra, F. Manghi and C. M. Bertoni, J. Phys. C10, 1911 (1977); J. D. Joannopoulos and M. L. Cohen, Phys. Rev. B10, 5075 (1974); E. J. Mele and J. D. Joannopoulos, Phys. Rev. B17, 1816 (1978); D. J. Chadi, J. Vac. Sci. Technol. 15, 631, 1244 (1978), Phys. Rev. B18, 1800 (1978); J. A. Knapp, D. E. Eastman, K. C. Pandey and F. Patella, J. Vac. Sci. Technol. 15, 1252 (1978); J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B20, 4150 (1979); A. Zunger, Phys. Rev. B22, 959 (1980); A. Mazur, J. Pollmann and M. Schmeits, Solid State Commun. 42, 37 (1982).

[4] W. M. Gibson and H. J. Gossman, J. Vac. Sci. Technol. B2, 343 (1984).

[5] C. B. Duke, S. L. Richardson, A. Paton and A. Kahn, Surf. Science 128, L135 (1983).

[6] J. van Laar and J. J. Scheer, Surface Sc 8, 342 (1967); J. van Laar and A. Huijser J. Vac. Sci. Technol. 13, 769 (1976); Guidat and D. E. Eastman, J. Vac. Sc Technol. 14, 831 (1976); W. E. Spicer, Landau, P. W. Gregory, C. M. Garner, Pianetta and J. Chyu, J. Vac. Sci. Techno 13, 780 (1976).

[7] R. P. Messmer, R. E. Allen and J. D. Do Solid State Commun. 45, 13 (1983).

[8] The theoretical uncertainty can be ±0.5 e

[9] To avoid singular matrices, t calculations for $\bar{\Gamma}$ are actually perform at $\bar{k}$=0.02$\bar{X}$. For $\theta \geq 21^{\circ}$, the smaller results for A_5 are extrapolated. The results are checked by considering sever different directions of $\bar{k}$, near $\bar{\Gamma}$.

[10] R. E. Allen, H. P. Hjalmarson and J. D. D Surf. Sci. 110, L625 (1981).

[11] C. Mailhiot, C. B. Duke and Y. C. Chan "The atomic geometries of ZnSe (110) a GaAs (110): determination by photoemissi spectroscopy," to be published.

RELATIVISTIC EMPIRICAL TIGHT-BINDING THEORY OF THE ENERGY BANDS OF GeTe, SnTe, PbTe, PbSe, PbS, AND THEIR ALLOYS

Craig S. Lent[a], Marshall A. Bowen[b], John D. Dow, and Robert S. Allgaier[c]
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 U.S.A.

and

Otto F. Sankey
Department of Physics, Arizona State University, Tempe, Arizona 85287 U.S.A.

and

Eliza S. Ho
Department of Physics, University of Illinois at Urbana-Champaign
Urbana, Illinois 61801 U.S.A.

(Received 29 July 1986)

The orthogonalized plane wave band structures of GeTe, SnTe, PbTe, PbSe, and PbS are fit with a nearest-neighbor, 18-orbital sp^3d^5, relativistic tight-binding model that exhibits chemical trends. The band gaps of $Pb_{1-x}Sn_xTe$, $Sn_{1-y}Ge_yTe$, and $Ge_{1-z}Pb_zTe$ alloys are predicted as functions of compositions x, y, and z. Bowing of the gap is expected to be substantial for $Ge_{1-z}Pb_zTe$, and either $Sn_{1-y}Ge_yTe$ or $Ge_{1-z}Pb_zTe$ should exhibit a Dimmock reversal.

1. Introduction

The rocksalt-structure IV-VI semiconductor compounds, such as PbTe, SnTe, GeTe, PbSe, and PbS all have small band gaps, high dielectric constants, interesting defect levels, and a variety of very unusual thermodynamic, vibrational, electronic, and infrared properties [1]. Exploitation of these properties for the fabrication of technologically important opto-electronic devices has been partially impeded by an incomplete understanding of the intrinsic and extrinsic electronic states of these materials. The IV-VI's have attracted relatively little theoretical attention, however, because their electronic band structures are complicated, having large relativistic splittings. At first glance, it would appear that the electronic states of bulk defects or surfaces of these materials can be understood only if one executes a very tedious, relativistic theory.

In this paper, we show that the apparently complicated energy bands of the IV-VI compounds can be parameterized by a simple nearest-neighbor tight-binding model Hamiltonian. The parameters of this model exhibit chemical trends and can be used to predict the electronic structures of alloys such as $Pb_{1-x}Sn_xTe$. Moreover, theories of defect energy levels and surface states in IV-VI's can be constructed using this simple Hamiltonian, as we shall demonstrate in subsequent work.

[a] Permanent address: Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556
[b] Permanent address: Department of Computer Science, Western Illinois University, Macomb, Illinois 61455.
[c] Present address: Theodore Associates, Inc., 10510 Streamview Court, Potomac, Maryland 20854.

PACS Number: 71.25.Tn

2. Tight-binding theory

The relativistic Hamiltonian that produces the energy band structures has the form [2]

$$H = (p^2/2m) + V + H_{so} + \hbar^2\nabla^2 V/8m^2c^2 - p^4/8m^3c^2 \quad (1)$$

where V is the crystal potential, the spin orbit interaction is

$$H_{so} = \hbar\vec{\sigma}\cdot(\nabla V\times\vec{p})/4m^2c^2,$$

and the remaining terms are the Darwin terms and the relativistic mass correction term [3].

Employing the ideas of Slater and Koster [4], Harrison [5], Chadi [6], and Vogl et al [7], we construct the nearest-neighbor tight-binding Hamiltonian:

$$H_0 = \Sigma_{\vec{R},\sigma,i}\{|a,i,\sigma,\vec{R}\rangle E_{i,a}\langle a,i,\sigma,\vec{R}| + |c,i,\sigma,\vec{R}+\vec{d}\rangle E_{i,c}\langle c,i,\sigma,\vec{R}+\vec{d}|\} + \Sigma_{\vec{R},\vec{R}',\sigma,i,j}\{|a,i,\sigma,\vec{R}\rangle V_{i,j}\langle c,j,\sigma,\vec{R}'+\vec{d}| + h.c.\} + H_{so}. \quad (2)$$

0749-6036/86/050491 + 09 $02.00/0

© 1986 Academic Press Inc. (London) Limited

TABLE I. Nearest-neighbor tight-binding parameters of GeTe, SnTe, PbTe, PbSe, and PbS, as fit to the band structure of Herman et al. [9], in eV. The column labelled GeTe* refers to Dimmock-reversed GeTe with the valence (conduction) band extremum at L_6^- (L_6^+). $V_{d,d\pi}$, $V_{d,s}$, and $V_{s,d}$ are taken to be zero.

	GeTe	GeTe*	SnTe	PbTe	PbSe	PbS
$E_{s,c}$	-7.847	-7.992	-6.578	-7.612	-7.010	-6.546
$E_{s,a}$	-10.974	-10.855	-12.067	-11.002	-13.742	-13.827
$E_{p,c}$	1.454	1.657	1.659	3.195	4.201	3.486
$E_{p,a}$	0.444	0.250	-0.167	-0.237	-1.478	-1.153
$E_{d,c}$	9.08	9.08	8.38	7.73	8.72	9.27
$E_{d,a}$	25.85	26.75	7.73	7.73	11.95	10.38
λ_c	0.505	0.577	0.592	1.500	1.693	1.559
λ_a	0.447	0.351	0.564	0.428	0.121	-0.211
$V_{s,s}$	-0.617	-0.631	-0.510	-0.474	-0.402	-0.364
$V_{s,p}$	0.877	0.788	0.949	0.705	0.929	0.936
$V_{p,s}$	0.790	0.876	-0.198	0.633	0.159	0.186
$V_{p,p}$	2.189	2.181	2.218	2.066	1.920	2.073
	-0.478	-0.498	-0.446	-0.430	-0.356	-0.281
$V_{p,d}$	-1.14	-1.65	-1.11	-1.29	-1.590	-1.142
$V_{p,d\pi}$	1.56	1.78	0.624	0.835	1.45	1.16
$V_{d,p}$	-1.55	-1.50	-1.67	-1.59	-1.09	-1.54
$V_{d,p\pi}$	0.976	0.742	0.766	0.531	0.0497	0.517
$V_{d,d}$	-3.79	-3.87	-1.72	-1.35	-1.90	-1.67
$V_{d,d\delta}$	0.887	0.892	0.618	0.668	0.692	0.659

Table II. Experimental values of the fundamental gap for GeTe, SnTe, PbTe, PbSe, and PbS used in fitting the tight-binding parameters of Table I (in eV).

	GeTe	SnTe	PbTe	PbSe	PbS
E_{gap}	0.2[a]	0.3[b]	0.186[c]	0.165[d]	0.286[d]

[a] L. Esaki, J. Phys. Soc. Japan, 1966, 21, 589 [Kyoto Conference Supplement], measurements at 4.2°K.
[b] Ref. [10], measurements at 4.2°K.
[c] Ref. [10], measurements at 12°K.
[d] D. L. Mitchell, E. D. Palik, and J. N. Zemel, Proc. Seventh Int. Conf. Phys. Semicond., 1964, p. 325 (1964), measurements at 4.2°K.

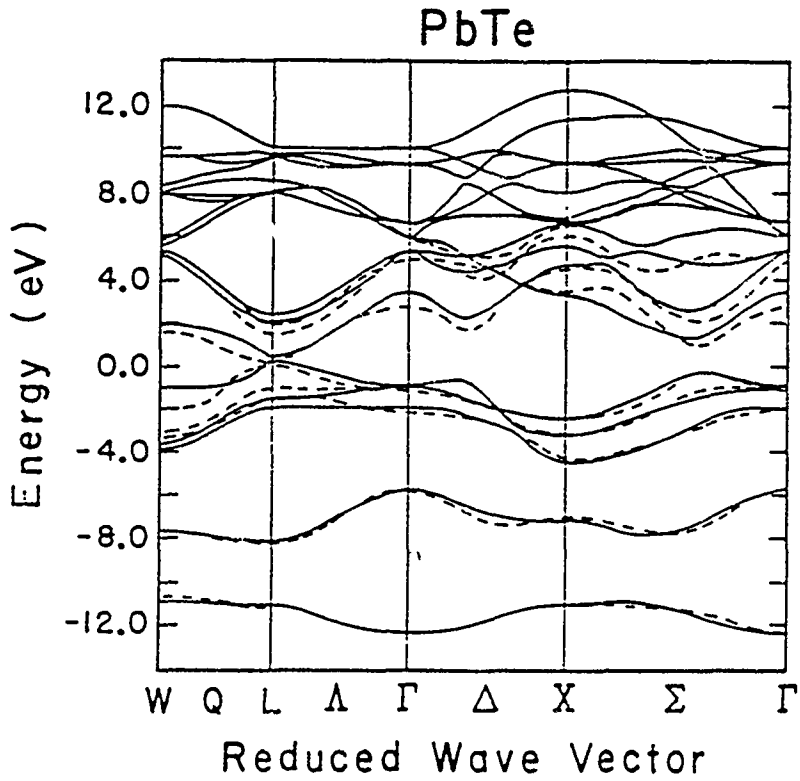


Fig. 1. The energy band structure in eV of PbTe, published by Herman et al. [9] (dashed) in comparison with the present work (solid). Note that the zero of energy is the valence band maximum and that the fundamental band gap is at L.

where h.c. means Hermitian conjugate, $\vec{R}$ are the rock-salt lattice positions of the anion, i and j are the basis orbitals for the cation and anion respectively, σ is the spin (up or down), a and c refer to the anion and cation respectively, and $\vec{d}$ is the position of the cation relative to the anion in the $\vec{R}$-th cell; $\vec{d} = (a_L/2)(1,0,0)$. The spin-orbit Hamiltonian is

$$H_{so} = \Sigma_{\vec{R},\sigma,\sigma',i} \{|c,i,\sigma,\vec{R}\rangle \lambda_c \vec{L}_c \cdot \vec{\sigma}_c \langle c,i,\sigma',\vec{R}|\} + \Sigma_{\vec{R},\sigma,\sigma',j} \{|a,j,\sigma,\vec{R}\rangle \lambda_a \vec{L}_a \cdot \vec{\sigma}_a \langle a,j,\sigma',\vec{R}|\}. \quad (3)$$

We use nine orbitals per atom in our basis, each with up and down spin: s, p_x, p_y, p_z, $d_{x^2-y^2}$, $d_{3z^2-r^2}$, d_{xy}, d_{yz}, d_{zx}. Because of the importance of the d bands near the bottom of the conduction band at the X point we found it necessary to include all five d bands in the model. This approach is to be preferred over that of Robertson [8], which included only two of the five d orbitals. We did neglect (i) the somewhat smaller couplings $V_{s,d\sigma}$ between the s states and the d states and (ii) $V_{d,d\pi}$, the π-type bonding between d states.

The resulting 36×36 Hamiltonian matrix is given in Appendix A.

3. Determination of the empirical Hamiltonian matrix elements

The parameters of this model are listed in Table I. They were obtained by fitting the eigenvalues of the matrix to the energy bands published by Herman et al. [9] (See Fig. 1). Analytic expressions for the eigenvalues at high symmetry points were used to make an initial guess for the parameters. Then a least-squares fit of the parameters to the calculated energy bands was performed. The symmetry of the states on either side of the fundamental gap was also included in the fitting procedure. This is necessary to assure the Dimmock reversal [10] in the ordering of bands that occurs in $Pb_{1-x}Sn_xTe$ between PbTe (with a conduction band minimum at L_6^- and valence band maximum at L_6^+) and SnTe (with the opposite ordering). The energy bands were fit to the values obtained by Herman et al. for wavevectors at the Γ, X, and L points of the Brillouin zone; but Herman's conduction band energies at L were all shifted by the same small amount in order to guarantee that the fundamental band gap agreed with experiment. The resulting band structures are displayed in Figs. 1-5. The fit of the band structure of GeTe

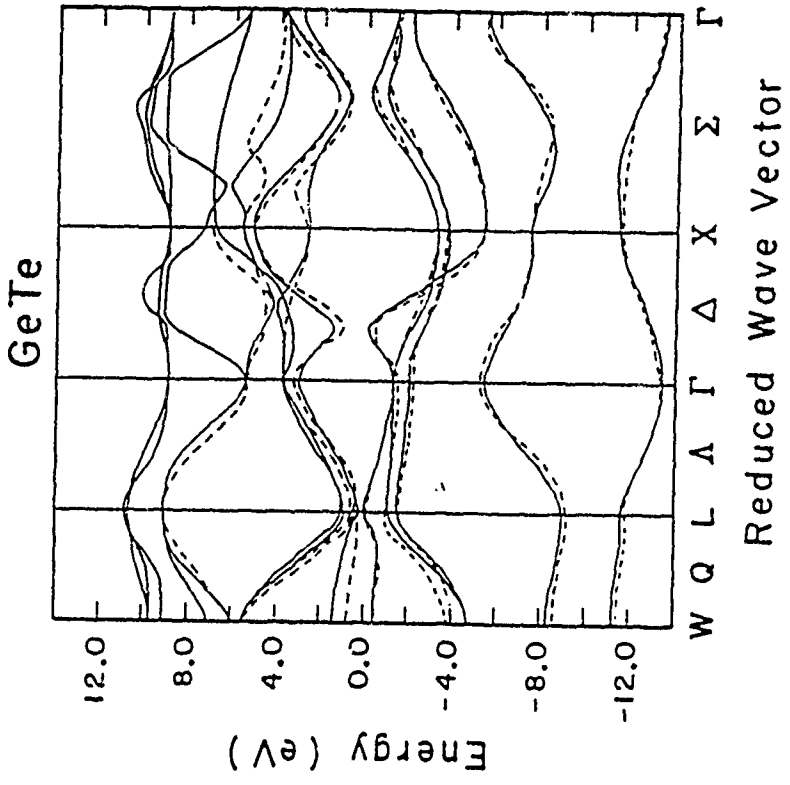


Fig. 3. The calculated energy bands of cubic (rocksalt) GeTe.

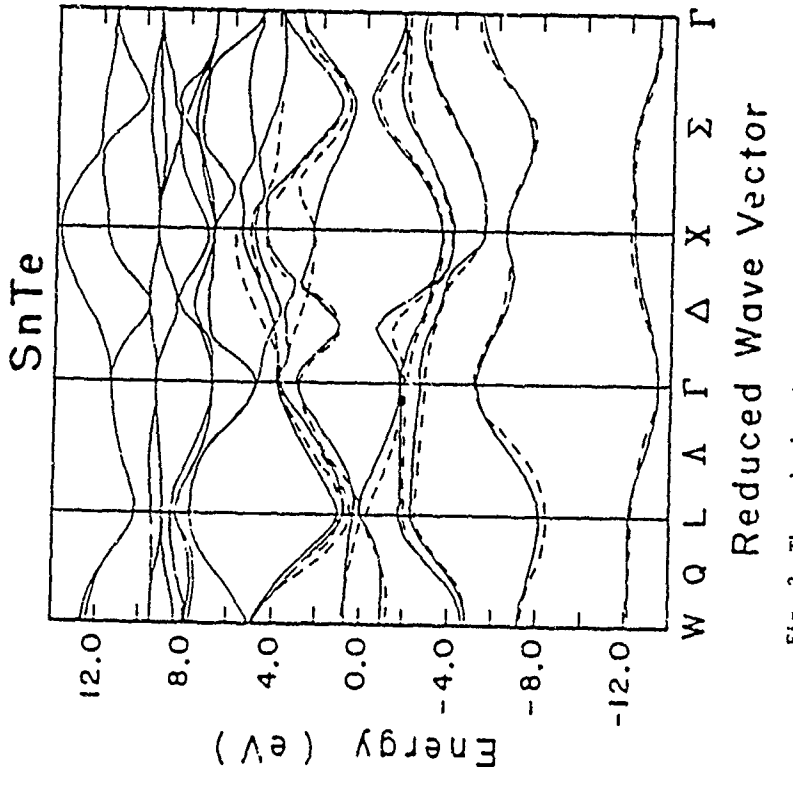


Fig. 2. The calculated energy bands of SnTe.

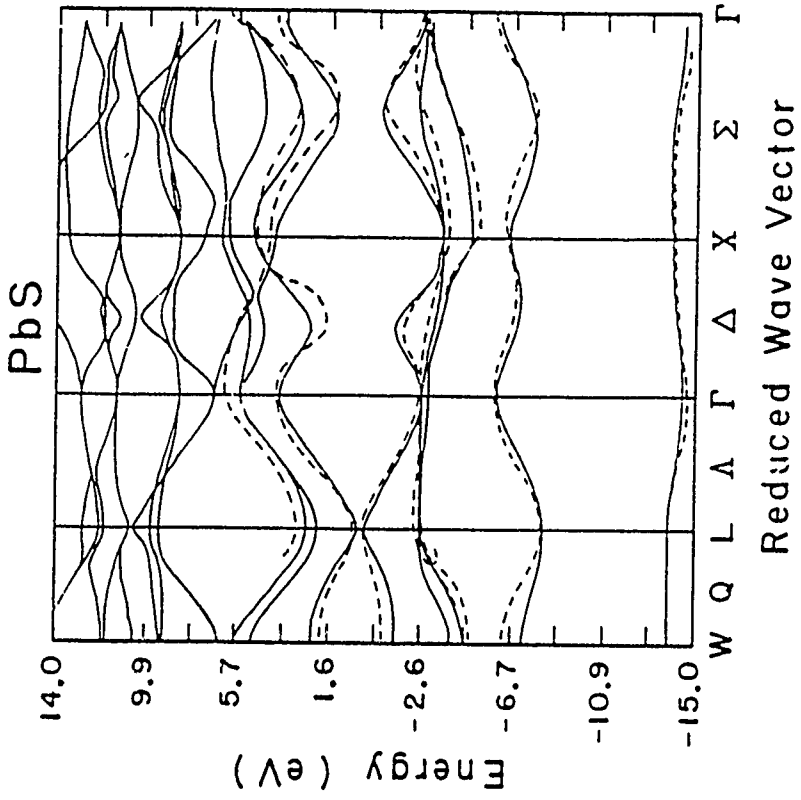


Fig. 5. The calculated energy bands of PbS.

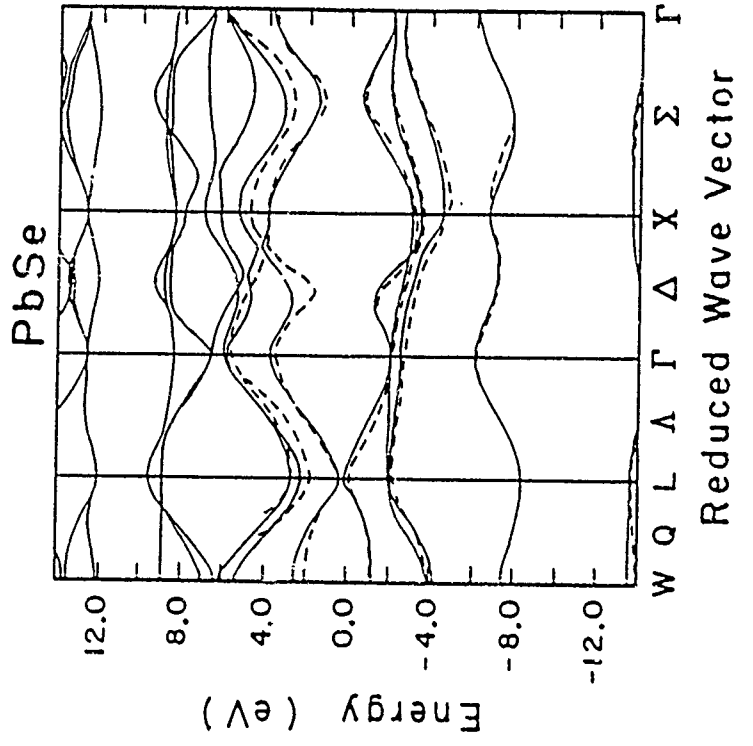


Fig. 4. The calculated energy bands of PbSe.

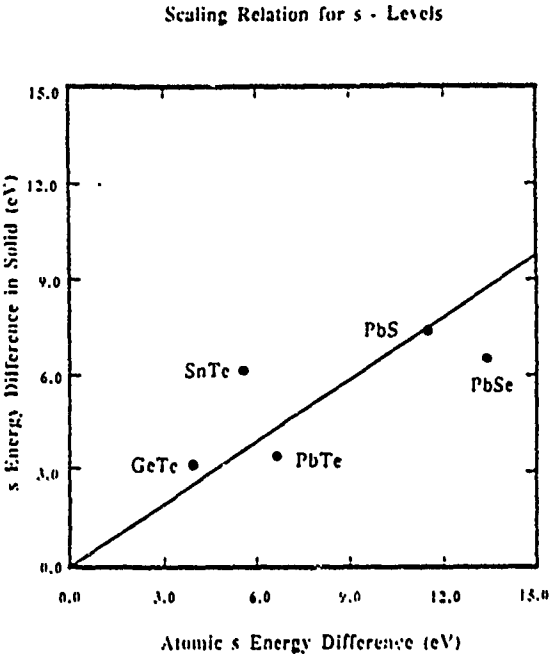


Fig. 6. The s-orbital energy differences in the solid, (Table I) versus the s-orbital energy differences in the atom [7].

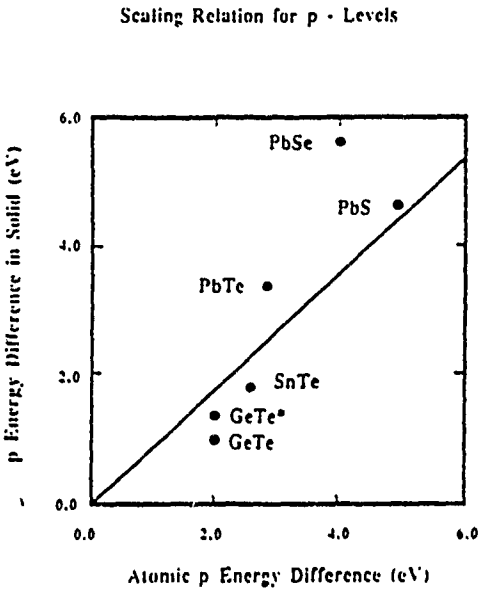


Fig. 7. The p-orbital energy differences in the solid, (Table I) versus the p-orbital energy differences in the atom [7].

assumes a PbTe-like ordering of the conduction and valence bands: L_6^- above L_6^+. The possibility exists, however, that GeTe's band structure is Dimmock-reversed, as SnTe's is, with L_6^+ being the conduction band minimum. We denote Dimmock-reversed GeTe by GeTe*, and obtain for it the slightly different matrix elements listed in Table I.

For the parameters of the model, the differences in the diagonal matrix elements $E_{s,c}-E_{s,a}$ and $E_{p,c}-E_{p,a}$ are approximately proportional to the corresponding differences of atomic energies. The Vogl constant of proportionality β [7] is about 0.65 for the s-state and 0.9 for the p-state. (See Figs. 6 and 7.) These proportionalities or scaling rules for the matrix elements of the empirical Hamiltonian allow the theory to make sensible predictions of chemical trends for intrinsic and extrinsic electronic states of different IV-VI semiconductors.

4. Application to Alloys

In this section we apply the theory to $Pb_{1-x}Sn_xTe$, $Sn_{1-y}Ge_yTe$, and $Ge_{1-z}Pb_zTe$ alloys and compute the alloy band gaps as functions of the compositions x, y, and z, using the virtual crystal approximation. These materials are substitutional alloys miscible for all compositions.

$Pb_{1-x}Sn_xTe$ is an interesting alloy because the band gap of SnTe is "inverted" in comparison with the gap of PbTe: the fundamental band gap occurs with the conduction band minimum being the L_6^+ point of the band structure, rather than at the the L_6^- point. This phenomenon was elucidated by Dimmock et al. [10] several years ago, who pointed out that a level-crossing phenomenon occurs with increasing x as the band-gap of $Pb_{1-x}Sn_xTe$ decreases and attempts to become negative. We calculate that the gap vanishes at x=0.35, in good agreement with the experimental value.

As a function of alloy composition, this Dimmock reversal in $Pb_{1-x}Sn_xTe$ must undo itself in either $Sn_{1-y}Ge_yTe$ or $Ge_{1-z}Pb_zTe$. We predict that the second Dimmock reversal must occur either near $y \approx 0.6$ in $Sn_{1-y}Ge_yTe$ or near $z \approx 0.3$ in $Ge_{1-z}Pb_zTe$. In the former case, GeTe must have the same ordering of L_6 bands as PbTe, whereas in the latter case, GeTe has the SnTe-like GeTe* electronic structure. (See Fig. 8.) The calculations also indicate that one should expect considerable bowing in the fundamental band gap versus alloy composition for $Ge_{1-z}Pb_zTe$, in contrast to the linear x-dependence of the gap for $Pb_{1-x}Sn_xTe$. This striking prediction of the calculations is in qualitative agreement with the measurements of Nikolic [11,12].

We compute the fundamental band gaps of alloys such as $Pb_{1-x}Sn_xTe$ by diagonalizing the virtual-crystal [13] empirical tight-binding Hamiltonian. The covalent radii of Pb and Sn differ by so little (≈ 4%) and all of the Hamiltonian matrix elements of PbTe and SnTe are sufficiently similar that a virtual crystal

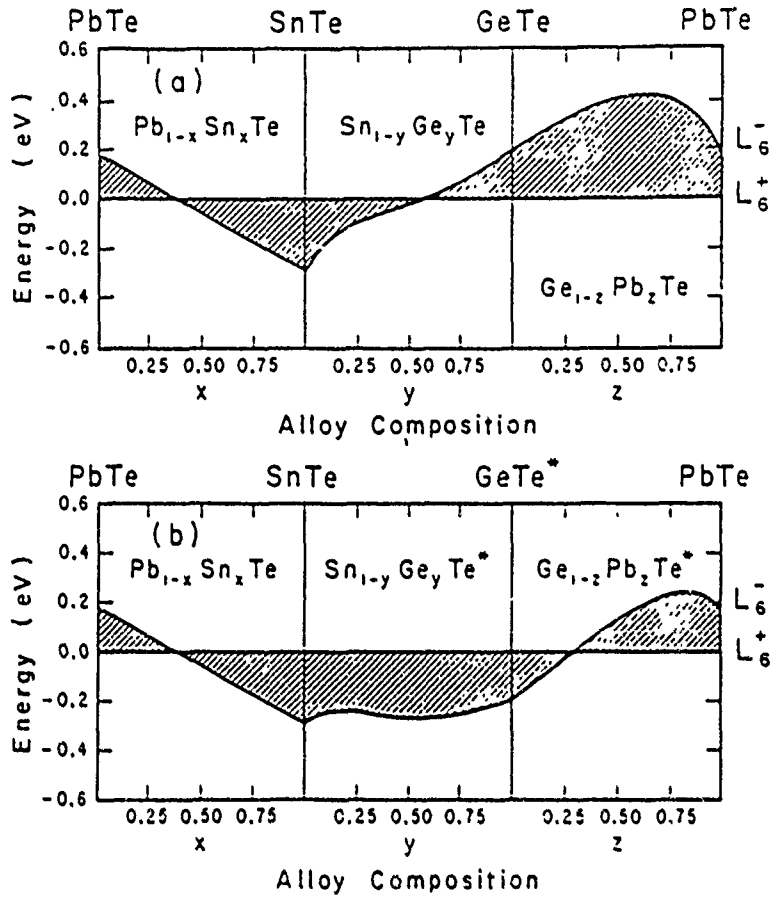


Fig. 8. The calculated band gaps and L_6^+ and L_6^- band edges of $Pb_{1-x}Sn_xTe$, $Sn_{1-y}Ge_yTe$, and $Ge_{1-z}Pb_zTe$ versus alloy compositions x, y, and z. The zero of energy is the L_6^- band extremum. The hatched area is the gap. (a) For ordinary GeTe; (b) for Dimmock-reversed GeTe* (see text).

approximation to the matrix elements of $Pb_{1-x}Sn_xTe$ is appropriate for states near the fundamental band gap [15]; the diagonal Hamiltonian matrix elements of $Pb_{1-x}Sn_xTe$ are (1-x) times the PbTe elements plus x times the SnTe matrix elements [14]. The off-diagonal matrix elements, multiplied by the square of the lattice constant [5], are similarly averaged, using Vegard's Law for the lattice constant. The band edges of Fig. 8 were eigenvalues obtained by diagonalizing this Hamiltonian for a wavevector at the L-point of the Brillouin zone.

5. Conclusion

We conclude that the present tight-binding parameterization of the IV-VI energy bands is adequate for reproducing chemical trends, including the Dimmock band reversal phenomenon in $Pb_{1-x}Sn_xTe$. Therefore, it should provide a satisfactory starting point for general theories of localized electronic states in these very small band-gap materials, such as "deep traps" [15] or surface states. Subsequent work will use this Hamiltonian to study a wide variety of problems involving localized electronic states in IV-VI semiconductors.

Acknowledgements -- The authors thank the Office of Naval Research and the Army Research Office for the support of this work (Contract Nos N00014-84-K-0352 and ARO-DAAG29-81-K-0068). One of us (E. S. H.) was supported in part by the

Department of Energy, Division of Materials Science (Contract No. DE-AC02-76ER01198) and another (R. S. A.) was supported in part by the Independent Research Fund of the Naval Surface Weapons Center. We thank R. Hong for carefully proof-reading the manuscript and checking the tight-binding parameters. We are especially grateful to G. B. Wright, whose enthusiastic support and encouragement led to the initiation of this work.

APPENDIX A: The 36x36 Hamiltonian matrix

The basis set for the 36x36 Hamiltonian is: $|s,c,\uparrow\rangle$, $|s,c,\downarrow\rangle$, $|s,a,\uparrow\rangle$, $|s,a,\downarrow\rangle$, $|p_x,c,\uparrow\rangle$, $|p_y,c,\uparrow\rangle$, $|p_z,c,\uparrow\rangle$, $|p_x,c,\downarrow\rangle$, $|p_y,c,\downarrow\rangle$, $|p_z,c,\downarrow\rangle$, $|p_x,a,\uparrow\rangle$, $|p_y,a,\uparrow\rangle$, $|p_z,a,\uparrow\rangle$, $|p_x,a,\downarrow\rangle$, $|p_y,a,\downarrow\rangle$, $|p_z,a,\downarrow\rangle$, $|d_1,c,\uparrow\rangle$, $|d_2,c,\uparrow\rangle$, $|d_3,c,\uparrow\rangle$, $|d_4,c,\uparrow\rangle$, $|d_5,c,\uparrow\rangle$, $|d_1,c,\downarrow\rangle$, $|d_2,c,\downarrow\rangle$, $|d_3,c,\downarrow\rangle$, $|d_4,c,\downarrow\rangle$, $|d_5,c,\downarrow\rangle$, $|d_1,a,\uparrow\rangle$, $|d_2,a,\uparrow\rangle$, $|d_3,a,\uparrow\rangle$, $|d_4,a,\uparrow\rangle$, $|d_5,a,\uparrow\rangle$, $|d_1,a,\downarrow\rangle$, $|d_2,a,\downarrow\rangle$, $|d_3,a,\downarrow\rangle$, $|d_4,a,\downarrow\rangle$, and $|d_5,a,\downarrow\rangle$,

where we have $d_1 = d_{x^2-y^2}$, $d_2 = d_{3z^2-r^2}$, $d_3 = d_{xy}$, $d_4 = d_{yz}$, and $d_5 = d_{zx}$.

The Hamiltonian can be written in block form as follows (only the lower triangular part of the Hamiltonian is given since it is Hermitian):

$$\begin{bmatrix} H_{s,s} & & & & \\ H_{pc,s} & H_{pc,pc} & & & \\ H_{pa,s} & H_{pa,pc} & H_{pa,pa} & & \\ 0 & 0 & H_{dc,pa} & H_{dc,dc} & \\ 0 & H_{da,pc} & 0 & H_{da,dc} & H_{da,da} \end{bmatrix}$$

$H_{s,s}$ is a Hermitian 4x4 matrix which connects s-states to s-states:

$$H_{s,s} = \begin{bmatrix} E_{s,c} & & & \\ 0 & E_{s,c} & & \\ g_0 V_{ss} & 0 & E_{s,a} & \\ 0 & g_0 V_{s,s} & 0 & E_{s,a} \end{bmatrix}$$

and $H_{pc,s}$ can be written as:

$$H_{pc,s} = \begin{bmatrix} 0 & H_1 \\ 0 & H_2 \end{bmatrix}$$

where we have

$$H_1 = \begin{bmatrix} -2g_1 V_{p,s} & 0 \\ -2g_2 V_{p,s} & 0 \\ -2g_3 V_{p,s} & 0 \end{bmatrix}$$

and

$$H_2 = \begin{bmatrix} 0 & -2g_1 V_{p,s} \\ 0 & -2g_2 V_{p,s} \\ 0 & -2g_3 V_{p,s} \end{bmatrix}$$

$H_{pa,s}$ can be written as:

$$H_{pa,s} = \begin{bmatrix} H_3 & 0 \\ H_4 & 0 \end{bmatrix}$$

where we have

$$H_3 = \begin{bmatrix} -2g_1 V_{s,p} & 0 \\ -2g_2 V_{s,p} & 0 \\ -2g_3 V_{s,p} & 0 \end{bmatrix}$$

and

$$H_4 = \begin{bmatrix} 0 & -2g_1 V_{s,p} \\ 0 & -2g_2 V_{s,p} \\ 0 & -2g_3 V_{s,p} \end{bmatrix}.$$

$H_{pc,pc}$ is a 6x6 Hermitian matrix with all the diagonal elements equal to $E_{p,c}$. Its other non-zero matrix elements in the lower triangular region are:

$\langle p_y,c,\uparrow| H |p_x,c,\uparrow\rangle = i\,\lambda_c/2$,
$\langle p_z,c,\downarrow| H |p_x,c,\uparrow\rangle = \lambda_c/2$,
$\langle p_z,c,\downarrow| H |p_y,c,\uparrow\rangle = i\,\lambda_c/2$,
$\langle p_x,c,\downarrow| H |p_z,c,\uparrow\rangle = -\lambda_c/2$,
$\langle p_y,c,\downarrow| H |p_z,c,\uparrow\rangle = -i\lambda_c/2$,
and
$\langle p_y,c,\downarrow| H |p_x,c,\downarrow\rangle = -i\lambda_c/2$.

$H_{pa,pa}$ is a 6x6 Hermitian matrix of the same form as $H_{pc,pc}$, but with the diagonal elements equal to $E_{p,a}$ and the other non-zero elements as above with λ_c replaced with λ_a.

$H_{pa,pc}$ is a diagonal 6x6 matrix with,

$\langle p_x,a,\uparrow| H |p_x,c,\uparrow\rangle = V_{x,x}$
$\langle p_y,a,\uparrow| H |p_y,c,\uparrow\rangle = V_{y,y}$
$\langle p_z,a,\uparrow| H |p_z,c,\uparrow\rangle = V_{z,z}$
$\langle p_x,a,\downarrow| H |p_x,c,\downarrow\rangle = V_{x,x}$
$\langle p_y,a,\downarrow| H |p_y,c,\downarrow\rangle = V_{y,y}$
$\langle p_z,a,\downarrow| H |p_z,c,\downarrow\rangle = V_{z,z}$

and

$V_{x,x} = 2g_4 V_{p,p} + 2(g_5+g_6)\, V_{p,p\pi}$

$V_{y,y} = 2g_5 V_{p,p} + 2(g_4+g_6)\, V_{p,p\pi}$

$V_{z,z} = 2g_6 V_{p,p} + 2(g_4+g_5)\, V_{p,p\pi}$

$H_{da,pc}$ can be written in block form as:

$$H_{da,pc} = \begin{bmatrix} H_5 & 0 \\ 0 & H_5 \end{bmatrix}$$

where H_5 is the 5x3 matrix.

$$\begin{bmatrix} -\sqrt{3}g_1V_{p,d} & \sqrt{3}g_2V_{p,d} & 0 \\ g_1V_{p,d} & g_2V_{p,d} & -2g_3V_{p,d} \\ -2g_2V_{p,d\pi} & -2g_1V_{p,d\pi} & 0 \\ 0 & -2g_3V_{p,d\pi} & -2g_2V_{p,d\pi} \\ -2g_3V_{p,d\pi} & 0 & -2g_1V_{p,d\pi} \end{bmatrix}$$

$H_{dc,pa}$ is of the same form, but with $V_{p,d}$ and $V_{p,d\pi}$ replaced by $V_{d,p}$ and $V_{d,p\pi}$

$H_{da,dc}$ is a 10x10 matrix with only four non-zero off-diagonal elements. The diagonal elements are:

$$\langle d_1,a,\uparrow| H |d_1,c,\uparrow\rangle = \langle d_1,a,\downarrow| H |d_1,c,\downarrow\rangle$$
$$= 3/2\ (g_4+g_5)V_{d,d} + (2g_6+g_4/2+g_5/2)V_{d,d\delta}$$
$$\langle d_2,a,\uparrow| H |d_2,c,\uparrow\rangle = \langle d_2,a,\downarrow| H |d_2,c,\downarrow\rangle$$
$$= 3/2\ (g_4+g_5)\ V_{d,d\delta} + (2g_6+g_4/2+g_5/2)V_{d,d}$$
$$\langle d_3,a,\uparrow| H |d_3,c,\uparrow\rangle = \langle d_3,a,\downarrow| H |d_3,c,\downarrow\rangle$$
$$= 2\ (g_4+g_5)\ V_{d,d\pi} + 2\ g_6\ V_{d,d\delta}$$
$$\langle d_4,a,\uparrow| H |d_4,c,\uparrow\rangle = \langle d_4,a,\downarrow| H |d_4,c,\downarrow\rangle$$
$$= 2\ (g_5+g_6)\ V_{d,d\pi} + 2\ g_4\ V_{d,d\delta}$$
$$\langle d_5,a,\uparrow| H |d_5,c,\uparrow\rangle = \langle d_5,a,\downarrow| H |d_5,c,\downarrow\rangle$$
$$= 2\ (g_4+g_6)\ V_{d,d\pi} + 2\ g_5\ V_{c,d\delta}\ .$$

The non-zero off-diagonal elements are all equal:

$$\langle d_1,a,\uparrow| H |d_2,c,\uparrow\rangle = \langle d_2,a,\uparrow| H |c_1,c,\uparrow\rangle$$
$$= \langle d_1,a,\downarrow| H |d_2,c,\downarrow\rangle = \langle d_2,a,\downarrow| H |d_1,c,\downarrow\rangle$$
$$= (\sqrt{3})/2\ (g_5-g_4)\ (V_{dd} - V_{d,d\delta})$$

where we have

$$g_0(\vec{k}) = 2[\cos(k_xa_L/2)+\cos(k_ya_L/2)+\cos(k_za_L/2)],$$
$$g_1(\vec{k}) = i\ \sin(k_xa_L/2),$$
$$g_2(\vec{k}) = i\ \sin(k_ya_L/2),$$
$$g_3(\vec{k}) = i\ \sin(k_za_L/2),$$
$$g_4(\vec{k}) = \cos(k_xa_L/2),$$
$$g_5(\vec{k}) = \cos(k_ya_L/2),$$

and

$$g_6(\vec{k}) = \cos(k_za_L/2).$$

The parameters $V_{s,s}$, $V_{s,p}$ and $V_{p,d}$ correspond to the integrals $(ss\sigma)_1$, $(sp\sigma)_1$, $(pd\sigma)_1$ in Ref. [4].

$H_{d,c}$ and $H_{d,a}$ are both 10x10 diagonal matrices whose elements are $E_{d,c}$ and $E_{d,a}$ respectively

REFERENCES

[1] G. Nimtz and B. Schlicht in Narrow Gap Semiconductors, (Springer Tracts in Modern Physics Vol. 98, Springer-Verlag, Berlin, 1983, ed. G. Hohler), p. 1.

[2] N. J. Parada, Phys. Rev B3, 2042 (1971).

[3] H. A. Bethe and E. E. Salpeter, Quantum Mechanics of One- and Two-electron Atoms, (Springer-Verlag, Berlin, 1957), p. 56.

[4] J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).

[5] W. A. Harrison, Phys. Rev. B8, 4487 (1973); W. A. Harrison, Electronic Structure and the Properties of Solids, (W. H. Freeman and Company, San Francisco, 1980)

[6] D. J. Chadi, Phys. Rev. B16, 790 (1977).

[7] P. Vogl, H. P. Hjalmarson and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[8] J. Robertson, Phys. Rev. B28, 4671 (1983).

[9] F. Herman, R. L. Kortum, I. B. Ortenburger and J. P. Van Dyke, J. Physique, 29, Suppl. C4, 62 (1968).

[10] J. O. Dimmock, I. Melngailis and A. J. Strauss, Phys. Rev. Letters, 16, 1193 (1966).

[11] The theory cannot confidently predict the bowing with quantitative accuracy. Rather it suggests that strong bowing should occur for $Ge_{1-x}Pb_xTe$ with x near unity.

[12] P. M. Nikolic, J. Phys. D 2, 383 (1969).

[13] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968). The range of validity of the virtual crystal approximation is a subject of current controversy. For summaries of the differing viewpoints, see L. C. Davis, Phys. Rev. B28, 6961 (1983) and references therein, and S. Lee, Ph.D. thesis, Physics Department, University of Illinois, 1986, and references therein.

[14] We fit the band gap of PbTe to experimental results at 12°K, but that of SnTe to results at 4.2K°, as given by Dimmock [10]. However, the changes in the gaps between the two temperatures amounts to only 1 or 2 meV, for both PbTe and SnTe. See the empirical relationsnip on page 42 of Ref. [1].

[15] C. S. Lent, M. A. Bowen, R. S. Allgaier, J. D. Dow, O. F. Sankey, and E. S. Ho, "Impurity levels in PbTe and $Pb_{1-x}Sn_xTe$," to be published.

Solid State Communications, Vol.61, No.2, pp.83-87, 1987. 0038-1098/87 $3.00 + .00
Printed in Great Britain. Pergamon Journals Ltd.

IMPURITY LEVELS IN PbTe AND $Pb_{1-x}Sn_xTe$

Craig S. Lent,[a] Marshall A. Bowen,[b] Robert S. Allgaier,[c] and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 U.S.A.

and

Otto F. Sankey
Department of Physics, Arizona State University, Tempe, Arizona 85287 U.S.A.

and

Eliza Y. Ho
Department of Physics, University of Illinois at Urbana-Champaign
Urbana, Illinois 61801 U.S.A.

(Received 17 September 1986 by A. A. Maradudin)

The doping characters of vacancies, the existence of impurity resonances, Fermi-level saturation and de-saturation, the sensitivity of the resonances to host chemistry, the behavior of In as both a donor and an acceptor in $Pb_{1-x}Sn_xTe$, and the occurrence of other apparently anomalous valences of impurities in IV-VI semiconductors are shown to be simple and direct consequences of a covalent defect theory, with several of the important defects necessarily assigned to "incorrect" or "anti" sites.

In this paper we present a simple chemical theory of s- and p-bonded substitutional point defects in PbTe and $Pb_{1-x}Sn_xTe$ which: (i) Corroborates experimental evidence [1] and the predictions of Pratt and Parada [2] (confirmed by Hemstreet [3]) that cation vacancies are double acceptors while anion vacancies are double donors; (ii) Naturally produces resonant defect levels near the fundamental band gap ("nearby resonances") that can limit, pin, or saturate the bulk Fermi energy so that increased doping beyond a critical value does not normally increase the Fermi energy (Experimentally this Fermi-level saturation has been inferred from "saturation anomalies" in the Hall coefficient, the Shubnikov de Haas effect, and many other effects [1,4].); (iii) Also predicts resonant levels further away from the gap, termed "distant resonances" (e.g., associated with Column-I or -VII impurities [1,4]) that can de-saturate the Fermi level and overcome the saturation anomalies (as observed in In/I- and Tl/Na-doped PbTe [1,4,5]); (iv) Explains why defects in different hosts have quite different Fermi-level saturation energies [1,4]; (v) Explains why some impurities, such as In (or other Column-III impurities), can be donors in PbTe but acceptors in SnTe [1,4]; (vi) Suggests that antisite defects and impurities on the "incorrect" site are common; and (vii) Shows why PbTe, with its large dielectric constant (that should almost fully screen Coulomb effects) nevertheless appears to support different ionic valence states of Pb: divalent Pb^{2+} and tetravalent Pb^{4+}.

The central idea of the theory is that s- and p-bonded defects each produce one s-like and three p-like "deep" levels with energies in the vicinity of the fundamental band gap. These levels are absent in the conventional effective-mass picture of impurity levels. Consider a Cd impurity on a Pb site in PbTe. One can imagine creating this impurity in three steps: (i) add two holes to PbTe, to account for the difference in the number of valence electrons between Cd and Pb; (ii) add two negative nuclear charges to the Pb nucleus on the impurity site, again to account for the valence difference, and (iii) adjust the potential in the impurity central-cell, so that the Cd potential, not the Pb potential, is present at the impurity site. (That is, turn on the "defect potential" which represents the difference between the potentials of Cd and Pb.) The first two steps are implemented in an ordinary effective-mass theory of shallow impurity levels [6], and do not produce strong impurity resonance levels. The third step, i.e., the introduction of the central-cell defect potential, has two effects. (a) a slight shift of the shallow impurity levels (the shift is small because very little of the effective-mass wave-function lies within the central-cell), and

(a) Permanent address: Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556.
(b) Permanent address Department of Computer Science, Western Illinois University, Macomb, Illinois 61455
(c) Present address: Theodore Associates, Inc., 10510 Streamview Court, Potomac, Maryland 20854.

(b) the production of s- and p-like "deep" levels (often resonant) associated with the perturbed chemical bonds between the impurity and the host. Most effective-mass theories do not include the effect (b). In IV-VI semiconductors the static dielectric constants are very large ($\epsilon \approx 10^3$ in PbTe [1]), and so all of the Coulomb effects are almost fully screened and negligible. Hence the dominant impurity effect in IV-VI's is the production of energy levels associated with the perturbed bonds. Because the fundamental band gaps of IV-VI semiconductors are so small, most of those energy levels lie outside the gap and are resonant with either the valence or the conduction band -- levels we term "deep," because they are due to the central-cell potential [7,8].

To understand the physics of these deep levels, consider first neutral defects[9] on the anion site of PbTe (Fig. 1). Begin with the simplest defect, Te, that is, no defect at all.

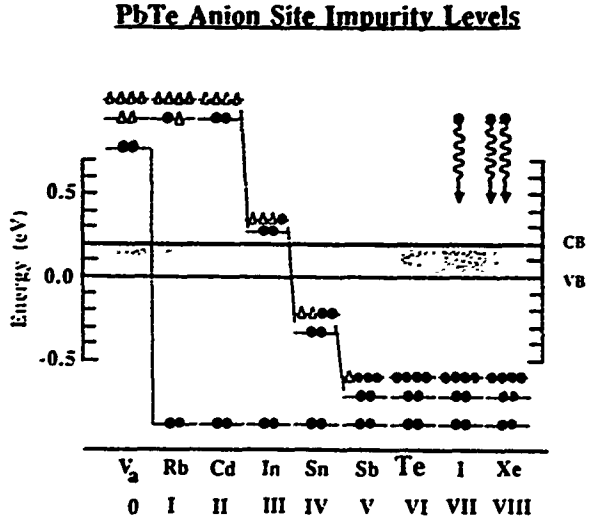


Fig. 1. Predicted energy levels and their electronic occupancies for neutral anion-site Rb-Row substitutional defects in PbTe. V_a denotes the anion vacancy, which can be thought of as originating from Column-0 of the Period Table. The Columns of the Periodic Table label the lower part of the figure. Typically the levels are ordered with increasing energy: s, $p_{1/2}$, and $p_{3/2}$. Electrons in levels above the conduction band minimum decay to the Fermi level (which, for otherwise perfect PbTe is the conduction band minimum). Holes in the valence band bubble up to the Fermi level (valence band maximum). Hence Cd_{Te} is a double donor. Electrons (holes) are denoted by closed circles (open triangles). Electrons that originate from higher levels appear at the top of the diagram for I and Xe. The dotted lines between defect levels emphasize the chemical trends across the Row of the Periodic Table. A level plotted off the scale is meant to suggest that such a level exists outside the window of the diagram.

Because Te is s- and p-bonded we can think of it as having various multiplets of s and p "levels" that are "broad resonances" -- that is, the various bands. One such multiplet of s and p levels is the valence band. So we can think of the Te "impurity" as having s and p "deep levels" that lie resonant with the valence band, which is generally acknowledged to have primarily Te p-character with some s-character [2,10]. (Of course, the six-fold degenerate p levels are split by the spin-orbit interaction into a four-fold degenerate $p_{3/2}$ level and a two-fold degenerate $p_{1/2}$ level, as shown in the figures.)

Now imagine continuously converting Te successively into the other Row-5 elements Sb, Sn, In, Cd, Rb, and a vacancy. Moving to the left in the Periodic Table corresponds roughly to all of the following equivalent operations: (i) increasing the atomic energy levels of the defect; (ii) increasing the strength of the defect potential; and (iii) decreasing the electronegativity of the defect. (A vacancy can be thought of as an atom with very large atomic energies -- so large that the "atom" is completely out of resonance with, and does not couple to, its neighbors [11].) Therefore as one moves to the left in the Periodic Table, these deep levels move up in energy. The deep resonances also become narrower, as the defect moves out of resonance with the host.

Pratt and Parada showed that the Te-vacancy, which corresponds to "Column-zero" of the Periodic Table and a strongly repulsive defect potential, has its p-levels above the conduction band edge. Thus, somewhere between the Te "defect" (Column-VI) and the vacancy (Column-0), the p-levels of the valence band cross the gap into the conduction band. Our calculations for PbTe, which reproduce the well-accepted Pratt-Parada result for the vacancy (also obtained by Hemstreet) indicate that the p-levels cross the gap near Column-III (Fig. 1). Similarly the s-levels cross into the conduction band as well (again a feature of the Pratt-Parada theory, Hemstreet's work, and the present computations for the Te-vacancy); with our calculations, the crossing occurs between the vacancy (Column-0) and Column-I. Note that the basic structure of the theory displayed in Fig. 1 is a consequence of continuity and the well-accepted Pratt-Parada theory of the vacancy. The only features of the results that depend on the Row of the Periodic Table or on any quantitative theoretical factor are the strengths of the defect potential (or, equivalently, the Columns of the Periodic Table) at which the s- and p-levels cross the fundamental band gap.

A similar analysis of the Pb-site begins with the Pb "defect" on the Pb-site which has an s-level in the valence band and p-levels in the conduction band (Fig. 2): the conduction band is largely Pb p-like and the Pb s states lies in the valence band [2,10]. Making the defect more electropositive (going from Column-IV to Column-0) drives these levels up in energy until eventually the s-level crosses into the conduction band (between Column-0 and Column-I) to form the deep levels of the vacancy. Making the defect electronegative drives the levels down, so that the p-levels ultimately drop into the valence band. For the Pb-row of the Periodic

PbTe Cation Site Impurity Levels

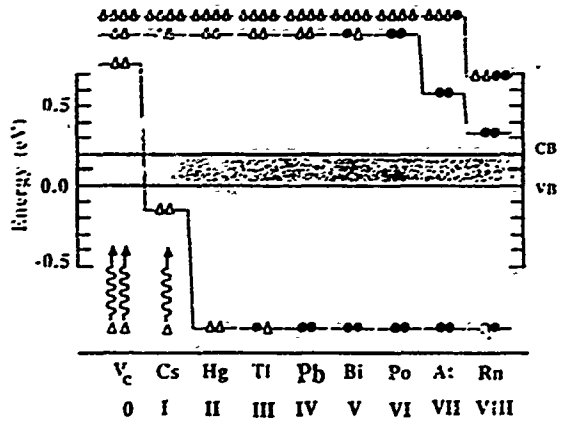


Fig. 2. Defect levels of cation-site substitutional defects in PbTe for the Cs-Row. V_c denotes a cation vacancy.

Indium Impurity Level in $Pb_{1-x}Sn_xTe$

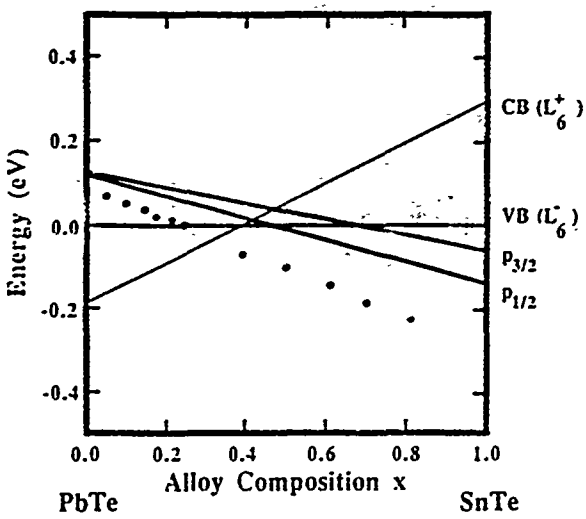


Fig. 3. Energies of the theoretical conduction band and valence band edges (hatched) and In_{Te} defect levels in $Pb_{1-x}Sn_xTe$ versus alloy composition x, compared with the data (circles) of Refs. [1] and [4].

Table, our calculations indicate that the p-levels cross the gap to the right of Column-VIII.

The consequences of this simple picture of level movement with defect potential strength are considerable. Consider first an "antisite" In atom on a Te site. It produces a donor in PbTe [12], with its "deep" p states forming a nearby conduction-band resonance in both the theory and the data [13]. Furthermore the p states are split into a doublet of $p_{3/2}$ and $p_{1/2}$ levels by relativistic effects [14]; such a doublet has been observed for Ga [15], which is predicted to have a very similar level structure to that of In (Fig. 1). The resonant In states lead to saturation of the Fermi energy: once the concentration of free electrons is so high that the Fermi energy coincides with the p-resonance, the previously unstable resonance becomes a stable electronic state capable of holding six electrons per In atom (three more than In itself provides) and the Fermi energy cannot move higher until these In resonant levels are completely filled by electrons from some other donor with a still higher resonance level, such as Cd_{Te} or iodine on either site (Figs. 1 and 2 [16]). The energies of the p resonances with respect to the valence band edge are sensitive to the host and appear to pass through the band gap into the valence band as a function of alloy composition x in $Pb_{1-x}Sn_xTe$ (Fig. 3). Once the levels are in the valence band and below the Fermi energy, In becomes an acceptor, as observed [1,4,17]. The levels resonant with the conduction band of PbTe occur naturally in the theory for the In_{Te} defect, and appear to move down through the gap as the Sn content of the host is increased -- in excellent agreement with the data [1,4] (Fig. 3). At the same time, the $Pb_{1-x}Sn_xTe$ gap undergoes a Dimmock reversal [18,19]. We cannot explain these data for the In resonance with "normal-site" In; with antisite In, the explanation is simple and natural.

The provocative suggestion that In (as well as Ga and Al) occupies a Te site when it produces a resonance in the PbTe conduction band receives support from the fact that Cd's tendency to be a donor in PbTe is also incompatible with its occupying a Pb site exclusively, as can be demonstrated without a calculation: The p deep levels of Cd on the Pb site are unquestionably resonant with the conduction band and Cd's deep s level can conceivably lie either in the valence band, in the gap, or in the conduction band. In any case, this level, which contains two electrons for the corresponding Pb atom on the cation site, can be thought of as having those two electrons removed on account of the valence difference between Cd and Pb. If the level lies in the valence band [20], the two holes bubble up to the valence band edge, making Cd a double acceptor. If the s-level lies in the gap, Cd is a deep trap for two electrons. If the s-level is above the conduction band edge, Cd is electrically inactive: neither a donor nor an acceptor. In any case, Cd on a Pb site cannot be a donor. This result is independent of any calculation.

In our model, we calculate the Cd s-level to lie in the valence band (similar to Hg in Fig. 2), making Cd a double acceptor on the Pb site. Thus the observation of donor action as a result of Cd doping indicates that Cd is not on its "normal" site and is compelling evidence that either "antisite" Cd (on the Te site) or interstitial Cd [3] commonly occurs. One provocative consequence of the theory presented here is that "antisite" impurities are

relatively common and responsible for several of the striking experimental observations: for example, the theory cannot explain the Fermi-level saturation data for In, Ga, or Al without assuming that those impurities occupy Te sites.

The predicted electronic occupancies of the various levels (assuming otherwise perfect PbTe) are also shown for neutral defects in Figs. 1 and 2. They are obtained by adding electrons or holes to account for the valence difference between the impurity and the host atom it replaces. For example, to simulate Cd on a Te site, one adds four holes to the p-levels. This leaves a doubly occupied $p_{1/2}$ level resonant with the conduction band, which is unstable if it lies above the Fermi level; hence the electrons in the Cd_{Te} $p_{1/2}$ defect level of otherwise perfect PbTe decay to the conduction band minimum, making Cd on the Te site a double donor. Thus we suspect that "antisite" impurities are common in PbTe and other IV-VI semiconductors and that significant concentrations of Cd atoms on Te sites will be observed in EXAFS and magnetic resonance studies of Cd in PbTe.

Defects on the anion site, such as Cd, that cause levels derived from the host Te atom to cross the fundamental band gap [20] exhibit "abnormal" doping behavior: they do not obey the usual valence rules of effective-mass theory. Therefore Cd is a double donor on the Te site, whereas effective-mass intuition would have predicted it to be a quadruple acceptor. This difference in predictions of six is due to the fact that six spin-orbital levels have crossed the gap [20] during the conversion of Te into Cd (Fig. 1). In the case of the Te vacancy, which has six fewer electrons than Te, both the s and p levels of Te cross the gap, and six of the eight electrons in those levels for Te are removed to form the vacancy; the two remaining electrons decay to th conduction band edge, making the Te vacancy . abnormal double donor.

Such abnormal doping behavior commonly occurs (theoretically) in IV-VI semiconductors for impurities that lie a few Columns distant in the Periodic Table from the host atoms they replace. Therefore, if the apparent valence of an impurity in a IV-VI semiconductor does not agree with the expectations of effective-mass theory, one should first examine the possibilities that the doping is abnormal (apparent valence differing from the effective-mass values by two, four [21], or six) or that the impurity is on the "incorrect" or "anti" site.

With a dielectric constant of $\epsilon \approx 10^3$, PbTe cannot support defects that are strongly ionic within the unit cell. Charged impurity states can exist, but much of the charge will necessarily lie in parts of the defect wave-function that are distant from the central cell. Therefore impurities are not properly viewed as being in ionic states and the oft-used terminology that an impurity replaces either divalent or tetravalent Pb is inappropriate Invariably such terminology is used to explain away abnormal doping behavior within an effective-mass picture: the deficient valence is introduced to account for a state that has crossed the gap.

The theory makes the following predictions for Te-site defects: Column-I and -II impurities on a Te site are abnormal single and double donors, respectively. Column-III impurities are mixed, with B a normal triple acceptor, and with Al, Ga, and In abnormal triple [12] donors. In all cases, the relevant $p_{3/2}$ level lies near the band gap, in the valence band for B and in the conduction band for Al, Ga, In, and Tl. According to the theory, Column-IV, -V, -VI, -VII, and -VIII (except He [22]) impurities are all normal on the anion site, producing double acceptor, single acceptor, isoelectronic center, single donor, and double donor behavior.

On the cation-site the Pb vacancy is an abnormal double acceptor, while alkali impurities and Column-II defects are normal triple and double acceptors. Columns III, IV, and V produce normal single acceptors, isoelectronic centers, and single donors, respectively -- except for N, which is predicted to be a quintuple acceptor, and possibly C, which is predicted to have a deep $p_{1/2}$ level slightly above the conduction band edge (but, if this level actually were to lie in the gap, it would be capable of trapping two electrons) Two possibilities are predicted for chalcogens on a Pb site in PbTe: S, Se, Te, and Po should be normal double donors; but both the $p_{3/2}$ and $p_{1/2}$ levels lie below the valence band maximum for O, making it a quadruple acceptor. F is an abnormal triple acceptor with all of its s- and p-levels in the valence band, while the theory places the $p_{1/2}$ level of Cl in the valence band, making Cl a single donor, and the $p_{1/2}$ level of Br in the band gap (a single donor and a deep hole trap) I and At are normal triple donors. The substitutional rare gas defects are generally abnormal: double acceptors (Ne, Ar), deep $p_{3/2}$ traps (Kr), or double donors (Xe). Rn is normal: a quadruple acceptor.

The theory in its present form makes no predictions concerning the solubility of specific impurities at particular sites. Nevertheless, as we have shown, the theory's predictions of levels often can be reconciled with data only if the impurity occupies a specific site. The model is based on a simple empirical sp^3d^5 tight-binding theory [10] applied to IV-VI semiconductors following ideas of Vogl [23] and Hjalmarson [7] for III-V semiconductors. Charge-state splittings (which should be negligible) and lattice relaxation (which may shift levels as much as tenths of an eV) have been omitted in order to display the chemical trends of the defect levels simply and generally. The model's predictions are remarkably insensitive to alterations of the model, and will be compared with data in subsequent work.

Acknowledgements -- We thank the Office of Naval Research for their generous support of this research (Contract No. N00014-84-K-0352). One of us (R. S. A.) was supported in part by the Independent Research Fund of the Naval Surface Weapons Center. Finally, we must thank G. B. Wright, whose enthusiasm for physics led to our collaboration.

REFERENCES

[1] H. Heinrich in Narrow-Gap Semiconductors Physics and Applications (Lecture Notes in Physics Vol. 133) Springer-Verlag, Berlin, 407 (1980); L. Palmetshofer, Appl. Phys A 34, 139 (1984).

[2] N J. Parada and G W Pratt, Phys Rev Lett. 22, 130 (1969); N. J. Parada, Phys. Rev. B 3, 2042 (1971), G. W Pratt, J. Nonmetals 1, 103 (1973)

[3] L. A. Hemstreet Phys. Rev. B 11, 2260 (1975); L. A. Hemstreet Phys Rev. B 12, 1212 (1975).

[4] V. I. Kaidanov, Yu I. Ravich, Usp. Fiz Nauk 145, 51 (1985) [Engl transl Sov Phys Usp. 28, 31 (1985)]; G Nimtz and B Schlicht in Narrow Gap Semiconductors, Springer Tracts in Modern Physics Vol. 98, Springer-Verlag, Berlin, 1983. These reviews cover most of the experimental literature.

[5] V. I. Kaidanov, S. A. Nemov, Yu I Ravich, A M. Zaitsev, Fiz. Tekh. Poluprovodn. 17, 1613 (1983) [Engl. transl.: Sov Phys Semicond. 17, 1027 (1983)].

[6] W Kohn in Solid State Physics (Academic Press, New York) vol. 5, ed. F Seitz and D. Turnbull, p. 258 (1957).

[7] H. P Hjalmarson, P. Vogl, D. J. Wolford and J D. Dow, Phys. Rev. Lett. 44, 810 (1980), and to be published. See also W. Y Hsu, J D. Dow, D. J Wolford, and B G Streetman, Phys Rev B 16, 1597 (1977).

[8] This definition supplants the previous one that a deep level is one in the fundamental band gap more than 0 1 eV from the nearest band edge.

[9] The charge-state splittings of the defect levels and the dependences of the on-electron energy levels on electronic occupancy are assumed to be negligible

[10] C. S Lent, M A. Bowen, J D Dow, E S Ho, and O F Sankey, Superlattices and Microstructures, 2, (December, 1986), in press.

[11] M Lannoo and P Lenglart, J Phys Chem Solids 30, 2409 (1969)

[12] Actually, In_{Te} is a triple donor according to the theory. Hence the filling of a Te vacancy (a double donor) by In_{Te} yields a net of one electron. And the combination of In_{Te} and In_{Pb} also yields one electron per In atom

[13] A. A Averkin, V I Kaidanov and R B Mel'nik, Fiz. Tekh. Poluprovodn. 5, 91 (1971) [Engl. transl.: Sov Phys Semicond 5, 75 (1971)].

[14] The theoretical values of the defect spin-orbit splitting may be somewhat inaccurate.

[15] A. N. Veis, V. I. Kaidanov, N A Kostyleva, R. B. Mel'nik and Yu I. Ukhanov, Fiz Tekh. Poluprovodn. 7, 928 (1973) [Engl transl . Sov. Phys. Semicond. 7, 630 (1973)]

[16] The iodine defect levels are similar to those of At in Fig. 2.

[17] Actually In_{Te} in SnTe is a triple acceptor

[18] J. O. Dimmock, I. Melngailis, and A. J. Strauss, Phys. Rev Lett , 16, 1193 (1966)

[19] By a Dimmock reversal, we mean the valence band and conduction band edges of PbTe, which are at the L_6^+ and L_6^- points of the band structure, become the conduction band and valence band edges, respectively of SnTe.

[20] We assume the Fermi level lies in the band gap, for the purposes of discussion Strictly speaking, the normal-abnormal transition occurs when impurity states cross the Fermi level, which normally lies within the conduction band or valence band for a IV-VI semiconductor in the extrinsic temperature range.

[21] Because of the spin-orbit splitting, the four $p_{3/2}$ levels can cross the Fermi level even though the $p_{1/2}$ levels do not

[22] The He s states lie in the valence band

[23] P. Vogl, H. P. Hjalmarson, and J D Dow, J Phys. Chem. Solids, 44, 365 (1983)

Journal of Luminescence 33 (1985) 103–107
North-Holland, Amsterdam

SHORT COMMUNICATION

A MECHANISM OF LUMINESCENCE ENHANCEMENT BY CLASSICAL-WELL STRUCTURES OR SUPERLATTICES

Shang-Fen REN

Department of Physics, Texas A&M University, College Station, Texas 77843 USA
and
Department of Physics, University of Science and Technology of China, Hefei, China *

John D. DOW

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 USA

Received 9 September 1983

A classical-well or superlattice structure can enhance luminescence by impeding energy transfer to killer centers.

In this paper we show that, under certain conditions, a superlattice structure or a classical-well structure can enhance the low-temperature luminescence of a material by impeding exciton transport to non-radiative killer centers [1,2,3]. (Wolford et al. have shown that alloy fluctuations have a similar effect on energy transfer in the GaP:N system [3].)

The basic idea is that the superlattice or classical-well introduces barriers to exciton migration in one of the three dimensions and therefore increases the effective mean distance from the exciton to the nearest accessible non-radiative recombination "killer" center (fig. 1). This exciton confinement effect reduces the non-radiative recombination rate of excitons and therefore increases the luminescent yield.

To see that such an effect must exist, consider a lattice with N traps (non-radiative recombination centers) dispersed in three dimensions with density D^{-3} (the mean distance between traps is of order D). Then consider a slice of this lattice of thickness $\lambda \ll D$. The mean distance between traps in this slice is of order $D(D/\lambda)^{1/2}$, larger by a factor $(D/\lambda)^{1/2}$. As λ becomes small in comparison with D, the distance from an exciton to a non-radiative trap in the

* Permanent address
PACS Numbers: 78.55.Ds; 78.55.Kz

0022-2313/85/$03.30 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

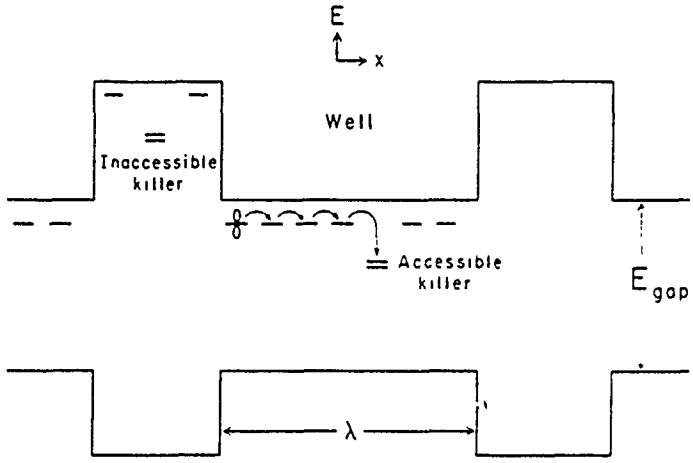


Fig. 1. A schematic one-electron diagram of a classical-well structure (energy versus position). The large band-gap layers form the confinement well in the material of band gap E_{gap}. An exciton (propeller) diffuses (arrow) to the nearest accessible killer center (double lines) through other impurity levels (solid lines). Killer centers in the barrier are inaccessible either because they are too far from the well for efficient transfer to them to be possible or because the change of their energies by the barrier renders them out of resonance [1] and impotent.

slice becomes large, so that an exciton confined to the slice can more easily radiatively decay before being trapped.

Stated more quantitatively, the probability that the nearest trap is a distance r from the exciton in a three-dimensional lattice is

$$p_\infty(r) = 4\pi r^2 xC(1-x)^{(CV)} \approx 4\pi r^2 xC \exp(-xCV),$$

where we have $V = 4\pi r^3/3$, C is the density of lattice points, xC is the density of (randomly distributed) traps ($xC = D^{-3}$), and we have assumed $x \ll 1$. The mean distance to the nearest trap is

$$R_\infty = \int_0^\infty r p_\infty(r)\,dr \Big/ \int_0^\infty p_\infty(r)\,dr$$

$$= \Gamma(4/3)(4\pi xC/3)^{-1/3},$$

where $\Gamma(z)$ is the gamma function. For a thin slice $\lambda \ll (xC)^{-1/3}$ we have the probability

$$p_\lambda(r) = 2\pi r\lambda xC(1-x)^{4\lambda C} \approx 2\pi r\lambda xC \exp(-A\lambda xC).$$

and the mean distance to the nearest trap in the slice is

$$R_\lambda = \Gamma(3/2)(\pi\lambda xC)^{-1/2}.$$

Here A is the area πr^2. Thus we have

$$R_\lambda / R_\infty = \left[\Gamma\left(\tfrac{4}{3}\right)\right]^{-1} (\pi/6)^{1/3} (xC)^{-1/6} \lambda^{-1/2} \propto (D/\lambda)^{1/2}.$$

We have executed model calculations to illustrate the effect on the luminescence of this change in effective nearest-trap-distance by the classical well. We assume that excitons walk randomly on a lattice containing traps which are distributed randomly. The jump time of the exciton between sites is a constant, τ_{jump}; the exciton concentration is small; and the excitons initially are randomly distributed on the lattice points. Excitons jump to nearest-neighbor sites only (with equal probability); and, once trapped, never reappear. The effect of the superlattice or classical well is to confine the exciton to a slab: when an exciton attempts to pass through a slab boundary, it is perfectly reflected (unless the confined slice is only one atomic layer thick, in which case the exciton is only permitted to jump to one of its four neighboring sites).

Our results, as obtained using Monte Carlo techniques [5–7] for 2000 excitons, are given in fig. 2 for a model in which we have $\tau_{jump} = 2$ ps, $x = 10^{-5}$, and a radiative lifetime of $\tau = 1$ μs [8]. (The Monte Carlo calculation has been checked by calculating the mean number of steps before trapping for various analytically solvable [7] problems involving random walkers on lattices with traps [9].) Because we have $\tau_{jump} \ll \tau$, we first compute the number of excitons remaining untrapped assuming no radiative transitions ($\tau = \infty$) and then multiply this result by $\exp(-t/\tau)$ to obtain a good approximation for the luminescence intensity $L(t)$, normalized to unity at $t = 0$. Note that confining

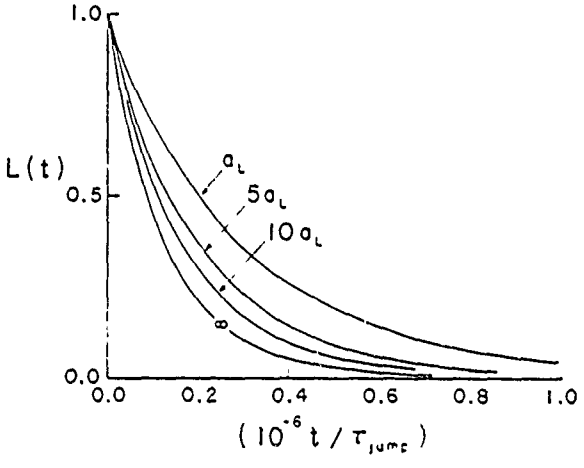


Fig. 2. Luminescence $L(t)$ versus time t in units of the jump time τ_{jump}, for an exciton confined to a well of thickness $\lambda = a_L$, $5a_L$, $10a_L$, or ∞, where a_L is the lattice constant. These curves were calculated for 2000 random walkers, with $x = 10^{-5}$, $\tau = 1$ μs, and $\tau_{jump} = 1$ ps.

wells of a_L, $5a_L$, and $10a_L$, where a_L is the lattice constant, do enhance the luminescence as expected: The total luminescent yields are enhanced by factors of 2.34, 1.59, and 1.33, respectively.

The present work is meant to demonstrate the qualitative effect on luminosity of exciton localization by superlattices or classical-well barriers, and hopefully will stimulate experiments to observe and quantitatively demonstrate the effect, which may have been present in the recent experiments of Petroff et al. [10]. The principal limitation of the present work is our assumption that the exciton is of zero radius and resides on a single lattice site at a time. This implies that the present theory is likely to be quantitatively applicable to only small-radius Frenkel excitons and to organic classical wells or superlattices. In semiconductors, excitons have large radii and, when confined in wells of less than ~ 100 Å width, exhibit quantum-well effects not included in the present model [11]. Nevertheless if the transport to the non-radiative killer center proceeds as a result of migration of one carrier through successive small-radius deep-level states that have significantly different energies in the barrier than in the well, the present model should apply semiquantitatively.

Acknowledgements

We are grateful to the Office of Naval Research (Contract No. N00014-77-C-0537) for their generous support of this work. We thank R.O. Simmons and the Department of Physics, University of Illinois, for providing us with the supportive environment in which this work was begun.

References

[1] D.L. Dexter, J. Chem. Phys. 21 (1953) 836.
[2] P.J. Wiesner, R.A. Street and H.D. Wolf, Phys. Rev. Lett. 35 (1975) 1366.
[3] D.J. Wolford and W.Y. Hsu, private communication; D.J. Wolford, Ph.D. thesis, University of Illinois, 1979.
[4] We assume that exciton transport occurs, e.g., via resonance energy transfer (refs. [1], [2] and [3]) and that the defects with levels in the gap that are responsible for the energy transfer have different energy levels in a high-band-gap barrier layer than in a free layer.
[5] Yu.A. Shreider, The Monte Carlo Method, translated from Russian by G.J. Lee and ed. by D.M. Parkyn (Pergamon Press, Oxford, New York, 1966); J.H. Killingbeck and G.H. Cole, Mathematical Techniques and Physical Applications (Academic Press, New York, 1971) pp. 622 et seq.
[6] For a discussion of some analytic techniques that can be applied to selected problems of this type, see ref. [7] and J. Klafter and R. Silbey, J. Chem. Phys. 72 (1980) 843, 849; 74 (1981) 3510; B. Movaghar, G.W. Sauer, D. Wurtz and D.L. Huber, Solid State Commun. 39 (1981) 1179; K.K. Ghosh and D.L. Huber, J. Lumin. 21 (1980) 225. D.L. Huber, Phys. Rev. B20 (1979) 2307

[7] E.W. Montroll, J. Math. Phys. 6 (1965) 167; J. Phys. Soc. Jpn. 26 (1969) 6; and J. Math. Phys. 10 (1969) 753.
[8] P.J. Wiesner, R.A. Street and H.D. Wolf, Phys. Rev. Lett. 35 (1975) 1366.
[9] These calculations were done for a one-dimensional lattice with uniform and random trap distributions, and for a two-dimensional square lattice and a three-dimensional simple-cubic lattice with uniform trap distributions.
[10] P.M. Petroff, C. Weisbuch, R. Dingle, A.C. Gossard and W. Wiegmann, Appl. Phys. Lett. 38 (1981) 965. The enhancement observed by these authors appears to be primarily due to a reduction of the density of native killer centers in the superlattice, in contrast to the present proposal, which should hold for a perfect superlattice. A.C. Gossard, Thin Films: Preparation and Properties, eds., K.N. Tu and R. Rosenberg (Academic Press), to be published.
[11] K. Hess and N. Holonyak Jr., Phys. Today 33 (1980) 40.

PHYSICAL REVIEW B VOLUME 30, NUMBER 8 15 OCTOBER 1984

Surface defects and core excitons at the (2×1) asymmetric-dimer (100) surface of Si

Marshall A. Bowen
Department of Physics, Western Illinois University, Macomb, Illinois 61455

Roland E. Allen
Department of Physics, Texas A&M University, College Station, Texas 77843

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 28 November 1983)

Predictions are given for the chemical trends in (i) deep energy levels associated with 28 sp^3-bonded substitutional defects, and (ii) energies of Hjalmarson-Frenkel core excitons at the (100) surface of Si, reconstructed according to Chadi's (2×1) asymmetric-dimer model. The predictions suggest that P at this surface should produce a deep level, and that the Hjalmarson-Frenkel surface core exciton should have a binding energy that is strongly site dependent.

In this paper we report a semiquantitative theory of defects at the Si(100)-(2×1) surface with its outer layer reconstructed according to Chadi's asymmetric-dimer model.[1] Our approach is to solve for the band-gap eigenvalues E of the secular equation,

$$\det[1-G_s(E)V]=0\ ,$$

where $G_s(E)$ is the surface Green's function and V is the defect potential of Hjalmarson *et al.*[2,3] Thus we follow the theory of bulk deep levels,[2] while evaluating G_s using the theory of Allen,[4] evanescent-wave techniques,[5] and the empirical tight-binding basis of Vogl *et al.*[3] A similar method has been applied extensively and successfully to defects at III-V(110) surfaces by Allen *et al.*,[6] and has provided a unified explanation of observed Schottky barrier heights.[7] Details of the method and calculational procedures are available elsewhere.[8]

The results of the calculation are predictions of deep

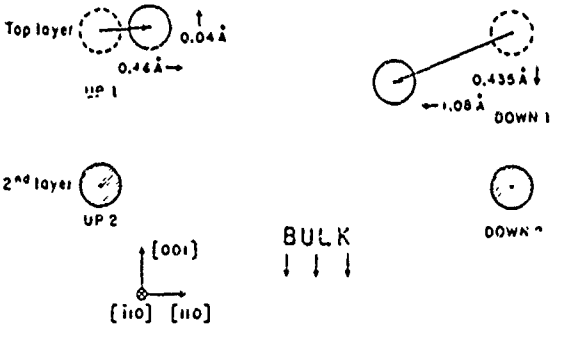


FIG. 1. Illustration of the geometrical structure of Chadi's asymmetric-dimer model for the Si(100)-(2×1) surface in the approximation that only the first plane of atoms undergoes relaxation. The surface is at the top of the figure. The circles denote rows of atoms in the (001) surface, viewed along the (110) direction (referred to the usual face-centered-cubic bulk directions). The rows of atoms in the top layer are displaced from their unreconstructed positions (dashed circles) to the open-circle positions. The displacements of the surface atoms employed for the present work are denoted by arrows. The second-layer rows of atoms are denoted by shaded circles. The plane of the up-1 and down-1 atoms in the plane of the paper lies below the corresponding up-2 and down-2 plane by a perpendicular distance $\sqrt{2}a_L/4$. The up-1 to up-2 distance is $\sqrt{3}a_L/4$, where a_L is the lattice constant.

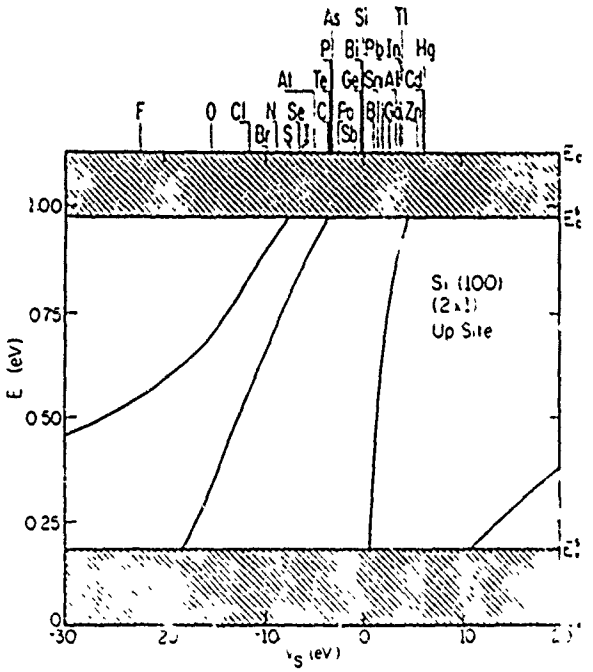


FIG. 2. Predicted deep energy levels for substitutional defects at the up site of the Si(100)-(2×1) surface, reconstructed according to Chadi's asymmetric-dimer model, versus defect potential V_s. The relevant defects appear at the top of the figure at their values of V_s. The shaded area denotes the calculated surface-state band (which are known experimentally to lie ≃0.5 eV lower than the calculation predicts[11]). The qualitative features of the calculation and chemical trends of the theory are important and meaningful results.

 ©1984 The American Physical Society

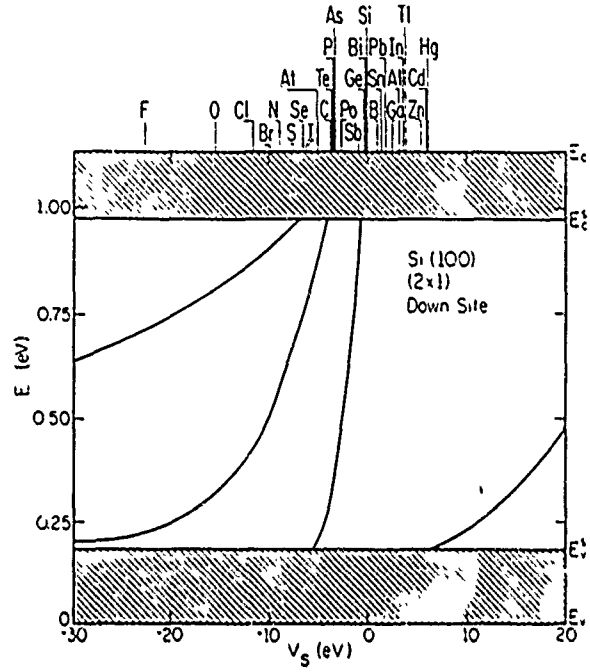


FIG. 3. Deep levels at the down site of the Si(100)-(2×1) surface.

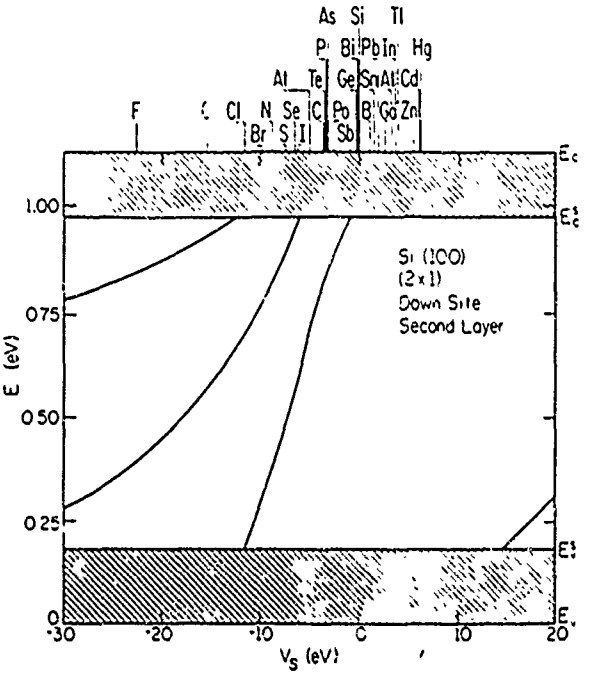


FIG. 5. Deep levels for defects at the down site of the second layer from the Si(100)-(2×1) surface.

trap energies E versus defect potential V. V is actually a diagonal matrix in the local sp^3s^* basis, centered on the defect site $V=(V_s,V_p,V_p,V_p,0)$; in order to present results as a function of a single parameter V_s, we use the rule $V_p=0.5V_s$.[9] Figure 1 shows the asymmetric-dimer model reconstruction; Figs. 2–5 show the predicted defect levels for atoms at the "up" and "down" sites of the first and second layers of the asymmetric-dimer reconstructed surface (see Fig. 1).

A major result is that the "deep" defect levels for atoms at the "up" and "down" sites and in the second and first layers[10] are different. This is shown explicitly in Figs. 6 and 7 for the P substitutional impurity and Si vacancy (which corresponds to $V_s\to\infty$), respectively. (Consider only the qualitative aspects of the predictions, do not take the precise energy levels literally; the expected uncertainty is $\simeq\pm0.5$ eV.[11,12])

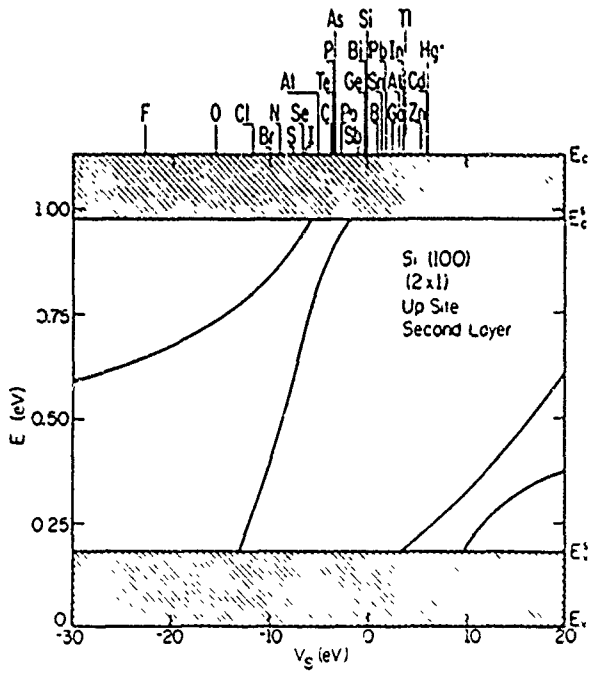


FIG. 4 Deep levels for defects at the up site of the second layer from the Si(100)-(2×1) surface.

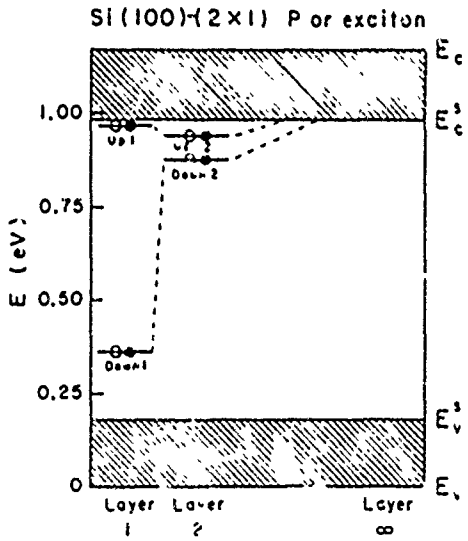


FIG. 6. The predicted energy levels within the surface band gap for substitutional P or a Hjalmarson-Frenkel core exciton (Ref. 13), as a function of layer number (∞ means bulk). The first- and second-layer deep levels all coalesce into a deep resonant level in the bulk. Closed (open) circles denote electronic (hole) occupation of the neutral defect levels. Sites are denoted, e.g., down 2 meaning the second-layer site beneath the down-1 site

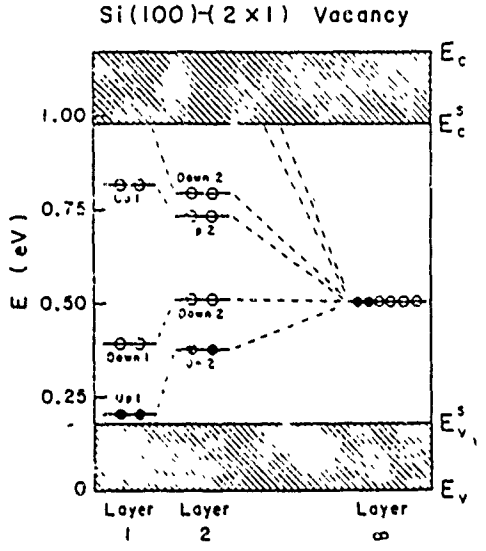


FIG. 7. The vacancy levels predicted to lie in the gap for the Si(100)-(2×1) surface as a function of layer number (∞ means bulk). Electronic (hole) occupation is denoted by closed (open) circles. The first- and second-layer levels all coalesce into the bulk T_2 vacancy deep level (Ref. 2), which, when neutral, contains two electrons and four holes.

Almost all impurities will produce at least one deep level in the gap for one of the four sites near the surface. In particular, the bulk shallow dopants P [Fig. 6 (Ref. 13)] and As, when near or at the surface, are predicted to yield deep levels in the gap. Hence, the Si(100)-(2×1) asymmetric-dimer surface of heavily doped Si should have many extrinsic surface states, with the number being roughly proportional to the surface concentration of dopants.

At the surface, P produces levels near both the middle and the top of the fundamental band gap of Si—a striking prediction that may have direct bearing on the Si $2p$ "core-exciton anomaly"—that the observed binding energy of the Si $2p$ core exciton appears to be considerably larger than predicted by shallow-impurity theory.[14] Insofar as a $2p$ core hole has a charge distribution similar to a proton, the core electron sees the effective nuclear charge Z of Si increased by unity to $Z+1$; that is, it sees a phosphorus defect potential.[13] Thus the energy level of the surface core exciton should approximately equal the energy of a P surface impurity, and the theory predicts that there should be core excitons at the Si(100) surface with apparent binding energies relative to the conduction-band edge, ranging from ≃0.2 eV on the up site of the first layer to more than half the band gap (≃0.8 eV) on the down site. The important point is not the absolute values of the predicted binding energies (which have significant theoretical uncertainties[12]) but that the variation of binding energy from site to site can be large.

It is possible that this dramatic site dependence of the Si surface core-exciton binding energy has been observed, although not recognized as such. Several authors have reported widely different Si $2p$ core-exciton binding energies,[14] from 0.1 to 0.9 eV, with some speculation that the proximity of the core exciton to the surface affects the binding energy. This variation is comparable with what we predict, 0.2–0.8 eV. Perhaps these diverse experimental results can be understood as due to experiments sensitive to core excitons at different sites.

This site-dependence occurs in part because electronic charge transfers from the down layers to the up layers at the surface. Thus the up site is a pseudo-anion site and the down site is a pseudo-cation at this surface. Chemically, the down site becomes more sp^2-bonded and the up site becomes more p backbonded.[15] The down cation site exhibits the larger exciton binding energy, as expected: In III-V semiconductors cation core excitons generally have larger binding energies than anion excitons. This is because the conduction-band states are cationlike, whereas the valence-band states are anionlike. This trend of the cation-site levels lying below the corresponding anion-site levels holds for the surface vacancies as well as for P or the core exciton.

The present results illuminate the richness of the spectra of surface impurities. At the present, however, there are few data for the energy levels of known surface impurities; we hope that the present work will stimulate more experiments in this direction and provide a guide for elucidating the chemical trends in data.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research (Contract No. N00014-77-C-0537) for their generous support, to D. V. Froelich for supplying us with a quartic root program and a proper version of Fig. 1, and to Seongbok Lee for his assistance in locating the Si bulk vacancy level predicted in Ref. 2.

[1] D. J. Chadi, Phys. Rev. Lett. 43, 43 (1979); J. Vac. Sci. Technol. 16, 1290 (1979).
[2] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).
[3] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[4] R. E. Allen, Phys. Rev. B 20, 1454 (1979).
[5] Y. C. Chang and J. N. Schulman, Phys. Rev. B 25, 3975 (1982), 26, 4400 (1982).
[6] R. E. Allen, H. P. Hjalmarson, and J. D. Dow, Surf. Sci. 110, L625 (1981); R. E. Allen and J. D. Dow, Appl. Surf. Sci. 11/12, 362 (1982); R. E. Allen, H. P. Hjalmarson, and J. D. Dow, Solid State Commun. 41, 419 (1982).
[7] R. E. Allen and J. D. Dow, Phys. Rev. B 25 1423 (1982), J. D. Dow and R. E. Allen, J. Vac. Sci. Technol. 20, 659 (1982).
[8] M. A. Bowen, an unpublished FORTRAN program for calculating deep levels at Si(100) surfaces, 1983; available from the Western Illinois Physical Sciences Reading Room, Western Illinois University, Macomb, Illinois 61455.
[9] H. P. Hjalmarson (unpublished). The rule $V_p \simeq 0.5 V_i$ permits

us to display the calculated results as a function of the single parameter V_s, and therefore to comprehend the global chemical trends in defect levels more easily. For a prescription for determining more accurately the energy levels of a specific defect from the curves of Figs. 2–5, without making the approximation $V_p \simeq 0.5 V_s$, see also H. P. Hjalmarson, Ph.D. thesis, University of Illinois, 1979; H. P. Hjalmarson, R. E. Allen, H. Büttner, and J. D. Dow, J. Vac. Sci. Technol. 17, 993 (1980).

[10]M. S. Daw and D. L. Smith, Phys. Rev. B 20, 5150 (1979); J. Vac. Sci. Technol. 17, 1028 (1980).

[11]M. A. Bowen, J. D. Dow, and R. E. Allen, Phys. Rev. B 26, 7083 (1982). The intrinsic surface states predicted by this model are $\simeq 0.5$ eV higher than observed.

[12]The theoretical uncertainty of an individual level is of order 0.5 eV, and so the precise values of the binding energies should not be overinterpreted. The relative positions of the different levels should be more accurate, however, and the qualitative predictions of the theory are reliable.

[13]H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B 24, 6010 (1981).

[14]The Si $2p$ core-exciton data have been reviewed by A. Quattropani, F. Bassani, G. Margaritondo, and G. Tinivella, Nuovo Cimento 51B, 335 (1979).

[15]J. Pollmann, in Festkörperprobleme *(Advances in Solid State Physics)*, edited by J. Treusch (Vieweg, Braunschweig, 1979), Vol. XX, p.117.

FILE: [dow.manus]milwau.mo 18 November 1983 15:45:33

Intrinsic and extrinsic surface electronic states of semiconductors

John D. Dow

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

and

Roland E. Allen

Department of Physics, Texas A&M University, College Station, Texas 77843

and

Otto F. Sankey

Department of Physics, Arizona State University, Tempe, Arizona 85281

Abstract

Recent theoretical work on the physics of surface deep impurity states, intrinsic surface states, and surface core exciton states in semiconductors is reviewed. The Schottky barrier ehights of the common semiconductors can be understood in terms of Fermi-level pinning by various surface deep levels associated with native defects or defects produced by surface treatments. The same theoretical framework, which has been successfully applied to bulk deep-level problems, also provides a good account of the physics of intrinsic surface state dispersion relations and surface core exciton states.

I. Introduction

In the bulk of a tetrahedral semiconductor, a single substitutional s-p bonded impurity or vacancy will ordinarily produce four "deep" levels with energies near the fundamental band gap: one s-like (A_1) and three p-like (T_2) [1]. These deep levels may lie within the fundamental band gap, in which case they are conventional deep levels, or they may lie within either the conduction or the valence band as "deep resonances." A sheet of N vacancies will produce 4N such deep levels — namely, the intrinsic surface state energy bands, which may or may not overlap the fundamental gap (to a good approximation, insertion of a sheet of vacancies is equivalent to creating a surface).

Intrinsic surface states have common underlying physics with deep impurities because they too result from localized perturbations of a semiconductor [2], and so their energies can be relatively easily predicted by extending to surfaces ideas developed by Hjalmarson, Vogl, Wolford, et al. [1] for the deep impurity problem. This has been done by several authors [3][4][5][6][7][8][9][10][11][12], most notably by Allen and co-workers [13][14][15][16][17].

Extrinsic and native-defect surface states also are governed by similar physics, and are especially interesting in the light of the Schottky barrier problem: Bardeen showed that modest densities of surface states on a semiconductor can "pin" the Fermi level [18], forming a Schottky barrier. The bulk Fermi energies of the semiconductor, the metal, and the semiconductor surface must align (Fig. 1). If the semiconductor is heavily doped n-type, the surface Fermi energy is the lowest empty surface state. The bands bend to

accomodate this alignment of Fermi levels, forming the Schottky barrier. Thus the Schottky barrier height is the binding energy of the lowest naturally empty surface state, relative to the conduction band edge. In 1976 Spicer et al. proposed that the Bardeen surface states responsible for pinning the Fermi energy are due to native defects [18][19][20][21][22].

Surface core excitons are similar to surface defect states, as can be seen by using the optical alchemy approximation [23] or the Z+1 rule [24]. Consider core excitation of a Ga atom at the surface of GaAs; the radius of the core hole is sufficiently small that the hole can be assumed to have zero radius (i.e., the hole is equivalent to an extra proton in the nucleus). Thus the core-excited electron feels the potential of an atom whose atomic charge Z is greater than that of Ga by unity, namely Ge. Thus the Ga core exciton spectrum is approximately the same as the spectrum of a Ge impurity on a Ga site. Hence the core exciton states in semicondutors can be either "shallow" (Wannier-Mott excitons) or "deep" (Hjalmarson-Frenkel excitons), as is the case for impurity states. The deep Hjalmarson-Frenkel excitons are similar to the surface deep levels associated with impurities.

In this paper. we show that the physics of deep impurity levels, intrinsic surface states, surface impurity states, Schottky barriers, and Hjalmarson-Frenkel core excitons are all similar.

II. Deep impurity levels at the surface: Schottky barriers and Fermi-level pinning

The basic physics of most Schottky barriers can be explained in terms of the Fermi-level pinning idea of Bardeen [18]. Stated in a slightly oversimplified form for a degenerately doped semiconductor at zero temperature, the Fermi energies of the metal, the bulk semiconductor, and the semiconductor surface all align in electronic equilibrium. For an n-type semiconductor with a distribution of electronic states at the surface, the Fermi level of the neutral surface is the energy of the lowest states that is not fully occupied by electrons. Electrons diffuse, causing band-bending near the semiconductor surface, until the surface Fermi energy aligns with the Fermi levels of the bulk semiconductor and the metal. This results in the formation of a potential barrier betweeen the semiconductor and the metal, the Schottky barrier (Fig. 1). For an n-type semiconductor, the Schottky barrier height is essentially the energy separation between the surface state that is the Fermi level and the conduction band edge. For a p-type semiconductor, the barrier height is the energy of the highest occupied electronic state of the neutral surface, relative to the valence band maximum. Thus the problem of determining Schottky barrier heights is reduced to obtaining the energy levels of the surface states responsible for the Fermi-level pinning.

In his original article, Bardeen focussed his attention on intrinsic semiconductor surface states as the most likely candidates for Fermi-level pinning. But he also pointed out that deep levels in the gap associated with impurities or native surface defects could also be responsible for the phenomenon.

Following Bardeen's work, a major advance occurred as a result of the experiments of Mead and Spitzer [25] who determined the Schottky barrier heights of many semiconductors, both n-type and p-type. Most of those old data have been confirmed by modern measurements taken under much more favorable experimental conditions.

However, after this work, the Schottky barrier problem was widely regarded as understood [26] in terms of concepts quite different from Fermi-level pinning.

In recent years Spicer and co-workers have revived the Fermi-level pinning model and have argued that the pinning is accomplished by native defects at or near the surface. Their picture is that during the deposition of the metal native defects are created at or near the semiconductor/metal interface, and that these semiconductor surface defects produce deep levels in the band gap that are responsible for Fermi-level pinning.

Spicer's viewpoint has been contested by Brillson and co-workers [27], who have emphasized the importance of chemical reactivity on barrier height. The Brillson viewpoint gains support from the observation of well-defined chemical trends in the variation of barrier height with the heat of reaction of the metal/semiconductor interface, as shown for n-InP by Williams et al. [28][29][30] (Fig. 2). (We believe that the Spicer and Brillson viewpoints can be reconciled.)

Daw, Smith, Swarts, and McGill [31] have proposed that free surface vacancies account for some of the observed Schottky barrier heights in III-V semiconductors. Allen and co-workers have argued that antisite defects

[32][33][34][35][36] "sheltered" [37] at the surface pin the Fermi energy for most Schottky barriers between III-V semiconductors and non-reactive metals, but that vacancies become the dominant pinning defect when the metal is reactive [36]. Thus the Brillson reactivity picture can be unified with the Spicer Fermi-level pinning picture: the chemical reaction merely changes the dominant pinning defect. The experimental results of Mead and Spitzer [25], Wieder [38][39][40], Williams [28][29][30], Mönch [41][42][43][44][45], their co-workers, and many others support this general viewpoint.

Moreover, the connection between the Schottky barriers formed at Si interfaces with transition metal silicides and the barriers between III-V semiconductors and metals appears to be provided by the recent work of Sankey et al. [46]: Fermi-level pinning can account for the silicide data as well. Thus a single unifying picture of Schottky barrier heights in III-V and homopolar semiconductors appears to be emerging. And although this Fermi-level pinning picture is no doubt oversimplified, it does provide a simple explanation of the first-order physics determining Schottky barrier heights, and how the physics changes when the dominant defect switches as a result of chemical reactivity.

It appears unlikely, however, that the Fermi-level pinning mechanism of Schottky barrier formation is universal. Layered semiconductors appear <u>not</u> to exhibit Fermi-level pinning, but rather seem to obey the original Schottky model [30]. This is probably because the layered semiconductors' surfaces are relatively impervious to defects and do not have defect levels in the band gap.

The Fermi-level pinning mechanism of Schottky barrier formation has the most advocates for III-V semiconductors such as GaAs and InP. However, even for these materials there are other proposed mechanisms for Schottky barrier formation, most notably those of Freeouf [47] and Ludecke [48].

Studies of Si, especially Si/transition-metal silicide interfaces, have focussed on the role of the silicide in Schottky barrier formation [49], in contrast to the studies of III-V's. Thus, prior to the recent work of Sankey et al. [46], it was widely believed that Fermi-level pinning was not responsible for the Schottky barrier at these silicide interfaces.

Thus the present state of the field is that Fermi-level pinning has its advocates for some semiconductors, but is not generally accepted as a universal mechanism of Schottky barrier formation, especially at Si/transition-metal silicide interfaces.

A central point of this paper is the Fermi-level pinning can explain an enormously wide range of phenomena relevant to Schottky barrier formation in III-V semiconductors and in Si — which no other existing model can do. In fact, the authors believe that Fermi-level pinning by native defects is responsible for the Schottky barrier formation in III-V semiconductors and in Si.

Our approach to the problem is simple: we calculate deep levels of defects at surfaces and interfaces, and we use these calculations to interpret existing data in terms of the Fermi-level pinning model. To illustrate our approach, we first consider the Si/transition-metal silicide interface and Fermi-level pinning by dangling bonds, as suggested by Sankey et al. [46].

a) Si/transition-metal silicide Schottky barriers

A successful theory of Si/transition-metal silicide Schottky barrier heights must answer the following questions: (1) How are the Schottky barrier heights at Si/transition-metal silicide interfaces related to those at interfaces of III-V semiconductors with metals and oxides? (2) Why is it that Schottky barrier heights of Si with different transition metals do not differ by ~1 eV, since changes of silicide electronic structure on this scale are known to occur [50]? (3) What is the explanation of the weak chemical trends that occur on a ~0.1 eV scale [50]? (4) Why are the Schottky barrier heights of silicides with completely different stoichiometries, such as Ni_2Si, NiSi, and $NiSi_2$ all equal to within ~0.03 eV? (5) Why are the Schottky barrier heights virtually independent of the silicide crystal structure? (6) Why is it that barriers form with less than a monolayer of silicide coverage? (7) Why do the Schottky barrier heights for n- and p-Si very nearly add up to the band gap of Si? (8) What role do the d-electrons of the transition metal play in Schottky barrier formation?

The answers to all of these questions are simple and straightforward, if one proposes (as Sankey et al. [46] have done) that the Si/transition-metal silicide Schottky barriers are a result of Fermi-level pinning by Si dangling bonds at the Si/transition-metal silicide interface. (1) The Fermi-level pinning idea unifies the Si/transition-metal silicide Schottky barriers with those found for the III-V's. (2) The Schottky barrier heights' independence of the transition-metal silicide comes from the fact that the causative agent, the Si dangling bond, is associated with the Si, and not with the silicide of transition metal. (3) The weak chemical trends in barrier heights occur

because the different transition-metal silicides repel the Si dangling bond wavefunction somewhat differently, causing it to lie slightly more or less in the Si. (4,5) The Schottky barrier heights vary very little with silicide stoichiometry and silicide crystal structure because the Si dangling-bond level is "deep-level pinned" in the sense of Hjalmarson et al. [1]: a large change in defect potential produces only a small change in the deep level responsible for Fermi-level pinning. The transition metal atoms act as inert encapsulants with the electronic properties of vacancies, because their energy levels are out of resonance with the Si. (6) Sub-monolayer barrier formation occurs because the Si dangling-bond defect responsible for the Fermi-level pinning is a localized defect that forms before a full interface is formed. (7) The Schottky barrier heights for n-Si and p-Si add up to the band gap because (in a one-electron approximation) the pinning level associated with the neutral Si dangling-bond at the interface is occupied by one electron, and so can accept either an electron or a hole: it is the surface Fermi level for both electrons and holes — both the lowest partially empty state and the highest partially filled state. (8) The d-electrons of the transition metal atoms play no essential role in the transition-metal silicide Schottky barrier formation, except to determine the occupancy of the Si dangling bond deep level; they are out of resonance with the Si at the interface.

The physics of the Si dangling-bond, Fermi-level pinning mechanism is contained in the very simple model presented by Sankey et al. [46]: to a good approximation, a Si dangling-bond at a Si/transition-metal silicide interface is the same as a vacancy in bulk Si with three of its four neighbors replaced by transition-metal atoms. To illustrate this physics, consider first a vacancy in bulk Si. This defect produces four deep levels near the band gap: a

non-degenerate A_1 or s-like level deep in the valence band (a "deep resonance") and a three-fold degenerate T_2 level in the band gap. The Si dangling bond defect at a Si/transition-metal silicide interface differs for the bulk Si vacancy in two ways: (1) some of the nearest-neighbors of the interfacial vacancy are transition-metal atoms rather than Si atoms; and (2) more distant neighbors are also different atoms at different positions -- but the experimental fact that Schottky barriers form at submonolayer coverages suggests that these differences in remote atoms are unimportant. Thus we can imagine constructing the Fermi-level pinning defect by slowly changing some of the Si atoms adjacent to a bulk Si vacancy into transition-metal atoms (Fig. 3).

To be specific we consider a $Si/NiSi_2$ interface, with a missing Si-bridge atom. Thus (Fig. 4) the Si bond dangles into the vacancy left by the removal of the Si bridge atom; this vacancy is surrounded by one Si atom and three Ni atoms.

How are the Ni atoms different from Si? First, their s and p orbital energies lie well above those of Si. Second, they each have an additional d orbital, with an energy that lies well below the Si s and p orbital energies (and is not terribly relevant here). The very positive Ni s and p energies act as a repulsive potential barrier to electrons, repelling the Si dangling bond electron from their vicinity in the silicide and forcing it to reside almost exclusively in the Si.

The effect of this positive potential barrier due to the Ni-Si difference, as it is turned on slowly in our imagination, is to drive the levels of the bulk vacancy upward in energy. In fact, for Ni, the potential is sufficiently positive to drive the T_2 bulk-Si vacancy level out of the gap into the conduction band. At the same time, the A_1 deep resonance of the Si bulk vacancy is also driven upward. For sufficiently large and positive potential, it pops into the fundamental band gap.

The A_1-derived level cannot be driven all of the way through the gap by the potential though, because an (approximate) level-crossing theorem prevents this. A simple way to see that there is an upper bound within the gap for the perturbed A_1 level is to consider a paired-defect of a vacancy V_{Si} with a neighboring atom X. If the atom X is Si, then the defect levels are the A_1 (s-like) valence band resonance and T_2 (p-like) band gap deep level of the bulk Si vacancy. A_1 and T_2 are not good irreducible representation labels of the (V_{Si},X) pair however; the A_1 level becomes σ-bonded and the T_2 level produces one σ-bonded and two π-bonded orbital, with the σ-bond oriented along the V_{Si},X axis and with the π bonds perpendicular to it. Thus the unperturbed (X=Si) σ levels of the (V_{Si},X) pair are the A_1 and T_2 bulk Si vacancy levels. The interlacing or no-crossing theorem [51] states that a perturbation cannot move a level further than the distance to the nearest unperturbed level. (It applies only approximately here.) Hence no matter how electropositive X is, the (V_{Si},X) level derived from the Si vacancy A_1 level cannot lie above the Si vacancy T_2 level. These considerations for general (V_{Si},X) pairs hold for the specific case of (V_{Si},Ni) pairs, and carry over to the dangling bond defect at the Si/transition-metal silicide interface, which is a vacancy surrounded by three Ni atoms and one Si. Thus the dangling-bond A_1 deep level is "deep-level

pinned" (as distinct from Fermi-level pinned) in the sense of Hjalmarson et al. [1], and is insensitive to even major changes in the nearby transition-metal atoms. To a good approximation, the nearby transition-metal atoms have the same effect as vacancies (which can be simulated [52] by letting the orbital energies of the transition-metal atoms approach $+\infty$, thereby decoupling the atoms from the semiconductor).

Thus the work of Sankey et al. [46] not only provides an explanation of the Si/transition-metal silicide Schottky barriers, it explains why calculations for defects at a free surface often can provide a very good description of the physics of Schottky barriers: the defects at interfaces are "sheltered" [37] or encapsulated by vacancies or by metal atoms that have orbital energies out of resonance with the semiconductor atoms; because of the deep-level pinning, the free-surface defects (which can be thought of as encapsulated by vacancies) have almost the same energies as the actual interfacial defects.

b) III-V Schottky barriers

The Fermi-level pinning story for Si/transition-metal silicides holds for Schottky barriers formed on III-V semiconductors as well. Here we summarize the main predictions of the theory.

The basic approach of the theory was to calculate the energy levels in the band gap of thirty s- and p-bonded substitutional point defects at the relaxed [53] (110) surfaces of III-V semiconductors. With these results in hand, Allen et al. examined Schottky barrier data in the context of Fermi-level pinning and eliminated from consideration all defects that

produced levels considerably farther than ~0.5 eV (the theoretical uncertainty) from the observed pinning levels. Interstitial defects were not considered; they have less of a tendency [54] to exhibit the deep-level pinning that is responsible for the experimental fact that different metals produce similar Schottky barrier heights. Moreover, extended defects were not considered initially, because it is known that paired-defect spectra are intimately related to and similar to isolated isolated-defect spectra [55]. (A more complete theory of Fermi-level pinning by paired defects, especially in GaSb where vacancy-antisite pairs are important, is in preparation.)

For clean semiconductors, the native substitutional defects potentially responsible for the commonly observed Fermi-level pinning are vacancies and antisite defects (anions on cation sites or cations on anion sites).

In GaAs, the defects proposed by Allen et al. [32] as responsible for Fermi-level pinning and Schottky barrier formation are the antisite defects. The cation-on-the-As-site defect accounts for trends with alloy composition of the Schottky barrier heights of n-type $In_{1-x}Ga_xAs$ and $Ga_{1-y}Al_yAs$ alloys (Fig. 5). The Fermi-level pinning of p-InAs [56], which shows quite different alloy dependences [57], is also explained.

This picture of Fermi-level pinning has been confirmed recently by Mönch and associates, who annealed Schottky barriers and showed that the Fermi-level pinning disappeared at the same temperature that the bulk (and presumably also the surface) antisite defect is known to anneal [58].

InP is an even more interesting material, because its Schottky barrier appears to depend on the heat of reaction of the interface [28][29][30]. This can be readily explained [36] however in terms of switching of the dominant Fermi-level pinning defect from an antisite defect for non-reactive metals to a vacancy for reactive metals (Fig. 2).

Moreover, surface treatments are known to alter the Schottky barrier height of n-InP, in a manner that can be easily understood in terms of the theory [36]: Surface treatments with Sn or S produce shallow donor levels associated with Sn_{In} or S_P at the surface, and these levels pin the surface Fermi energy for contacts between n-InP and the non-reactive noble metals. Likewise O and Cℓ treatments lead to reactions with P that leave P-vacancies, so that the surface Fermi-level of treated n-InP interfaced with non-reactive metals lies near the conduction band edge -- as though the metals were reactive.

Thus the Fermi-level pinning idea appears to provide a simple and unifying understanding of a wide variety of Schottky barrier data in the common semiconductors.

III. Intrinsic surface states

The calculations of surface defect levels for the Schottky barrier problem can be checked by simultaneously evaluating surface state energies and comparing them with the considerable body of available data. The theory underlying surface state calculations is basically the same as that for bulk point defects or surface defects. It is quite simple, and requires only (1) the well-established empirical tight-binding Hamiltonian of the semiconductor

[59] (the matrix elements of the Hamiltonian exhibit manifest chemical trends from one semiconductor to another), and (2) knowledge of the positions of the atoms at the surface. Thus a reliable treatment of the surface states of a semiconductor requires an adequate model of the geometrical structure of the surface. At present, no semiconductor surface structures are beyond controversy [60], but two seem to be rather well accepted; the (110) surface structure of III-V and II-VI semiconductors with the zincblende [53][61][62][63], and the $(i0\bar{1}0)$ surface structure of II-VI semiconductors with the wurtzite structure [62]. In particular, (110) zincblende surfaces are characterized by an outward, almost-rigid-rotation relaxation of the anion (e.g., As in GaAs), with the bond between surface anion and surface cation rotating through about 27° (III-V's) or 33° (II-VI's), and with small bond length changes and subsurface relaxations.

a) (110) surfaces of III-V and II-VI zincblende semiconductors

During the past five years, a number of groups have reported experimental and theoretical studies of intrinsic surface states at (110) zincblende surfaces [3][4][5][6][7][8][9][10] [11][12][13][14][15][16][17][18] [64][65][66][67][68][69][70][71][72][73]. In Fig. 6, we show the most recent calculation for the dispersion curves $E(\bar{k})$ at the GaAs (110) surface [14], together with the measured surface state energies fo Williams, Smith, and Lapeyre [65] and of Huijser, van Laar, and van Kooy [66]. The calculation employs the ten-band sp^3s^* empirical tight-binding model of Vogl et al. [59]. The agreement between theory and experiment is excellent. For example, along the symmetry lines $\bar{X}'\bar{\Gamma}$ and $\bar{\Gamma}\bar{X}$ (i.e., the boundary of the surface Brillouin zone), the uppermost branch of observed states appears to be explained by A_5,

the next branch by the overlapping resonances A_4 and A_2', and the three lower branches by A_1', A_3, and C_2. Here "A" and "C" refer to states localized primarily on anion and cation sites, respectively. A detailed comparison with previous theoretical studies of the GaAs (110) surface is given in Ref. [14]. The primary additional features are (i) the states A_1 through A_5 and C_1 through C_4 (in the notation of Ref. [7]) were located as bound states or resonances at all planar wavevector $\bar{k}$ along the symmetry lines of the surface Brillouin zone, and (ii) two "new" resonances, A_1' and A_2' were found. (The branch A_1' was reported in Refs. [5] and [74], but not in the other theoretical studies. The branch A_2' had not been previously reported.) The discovery of this additional resonant structure is apparently due to an improved technique for calculating bound states and resonances -- the "effective Hamiltonian" technique [14].

In Fig. 7, the theoretical dispersion curves of Beres et al. [14] are shown for the (110) surface of ZnSe, together with the measured surface state energies reported by Ebina et al. [11]. Again, the agreement between theory and experiment is quite satisfactory, being a few tenths of an eV near the band gap, and larger for more distant states. Some apparent discrepancies [11] between experiment and previous theory were found to be resolved by a more complete treatment of the resonances, using the approach described above.

Surface state dispersion relations have also been calculated for GaP, GaSb, InP, InAs, InSb, AlP, AlAs, AlSb, and ZnTe [14][15][16][17]. In none of the direct-gap materials were intrinsic surface states found within the band gap. GaP, however, was found to have a band of unoccupied surface states that overlaps the fundamental band gap and extends below the bulk conduction band

edge. This is in accord with the experimental facts: of these semiconductors only GaP has surface states in the gap [18][69][70][71][72]. Of the remaining indirect-gap materials, the theory indicates that intrinsic surface states may be observable near the top of the band gap in the indirect-gap Al-V compounds [16], although the theory is not sufficiently accurate to predict unequivocally that the states will lie within the gap.

b) Si (100) (2×1) intrinsic surface states

After many years of intensive study by numerous groups, there is still controversy over the geometrical structures of the most thoroughly studied semiconductor surfaces: Si (100) (2×1) and Si (111) (2×1). For example, four groups have recently given arguments for antiferromagnetic ordering of Si (111) surfaces [75], whereas Pandey has proposed replacing the conventional buckling model [76][77][78] of Si (111) (2×1) by a (110)-like chain model [79].

In the case of Si (100), arguments have recently been presented [80][81][82] against the (2×1) asymmetric dimer model of Chadi [83]. (In the asymmetric dimer model, adjacent rows of surface atoms dimerize, forming a pattern of paired atomic rows on the surface.) The most telling of these arguments involves the apparent disagreement between angle-resolved photoemission measurements of the surface-state dispersion curves [64][65] and theoretical calculations of these dispersion curves with conventional models of the electronic structure as applied to the asymmetric dimer geometry [83][84].

Very recently, two new calculations have been performed independently with improved models of the electronic structure [85][86]. The same conclusion was reached in both of these studies: the electronic structure calculated for the asymmetric dimer model is in agreement with the measurements. This is illustrated in Fig. 8 (taken from Ref. [48]), where both the theoretical band width of 0.65 eV and the detailed variation with the planar wavevector $\bar{k}$ are seen to be in excellent agreement with the experimental dispersion curves. In addition, there is quite satisfactory agreement between the theoretical surface band gaps and the 0.6 eV gap measured by Mönch et al. [87].

IV. Surface core exciton states

The same calculations that predict native-defect surface deep levels for the Schottky barrier problem also yield surface core exciton energies, because the optical alchemy or Z+1 rule states that the Hjalmarson-Frenkel core exciton energies are the energies of "impurities" that are immediately to the right in the Periodic Table of the core-excited atom [23][24]. Thus core-excited Ga produces a "Ge defect" and core-excited In yields "Sn."

In Figs. 9 and 10, the theoretical exciton energies for the (110) surfaces of the Ga-V and In-V compounds are compared with experiment [88]. Notice that the experimental and theoretical exciton levels for InAs and InSb lie above the conduction band edge, as resonances rather than as bound states. In the present theory this result has a simple physical interpretation: Like a deep impurity state, the Hjalmarson-Frenkel exciton energy is determined primarily by the high-density-of-states regions of the bulk band structure. There is only a small density of states near the low-lying direct conduction

band minimum (corresponding to the Γ-point of the Brillouin zone), but a large density of states near the higher, indirect X minima. Thus the conduction band minimum near Γ has relatively little influence on the position of the exciton.

The surface Hjalmarson-Frenkel core excitons have also been calculated for the (110) surface of ZnSe and ZnTe [89] and are in good agreement with the measurements [90]. We conclude that the present theoretical framework does a good job of explaining the basic physics of the "deep" Hjalmarson-Frenkel core excitons, whether bound states or resonances.

V. Unified picture

Thus one interlocking theoretical framework successfully predicts the correct physics of (1) surface deep impurity levels and Schottky barrier heights, (2) intrinsic surface states, and (3) Hjalmarson-Frenkel core exciton states.

REFERENCES

[1] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Letters 44, 810 (1980).

[2] H. P. Hjalmarson, H. Büttner, and J. D. Dow, J. Vac. Sci. Technol. 17, 993 (1980); R. E. Allen and J. D. Dow, J. Vac. Sci. Technol. 19, 383 (1981).

[3] C. Calandra, F. Mannghi, and C. M. Bertoni, J. Phys. C 10, 1911 (1977).

[4] J. D. Joannopoulos and M. L. Cohen, Phys. Rev. B10, 5075 (1974); E. J. Mele and J. D. Joannopoulos, Phys. Rev. B17, 1816 (1978).

[5] D. J. Chadi, J. Vac. Sci. Technol. 15, 631, 1244 (1978); Phys. Rev. B18, 1800 (1978).

[6] J. A. Knapp, D. E. Eastman, K. C. Pandey, and F. Patella, J. Vac. Sci. Technol. 15, 1252 (1978).

[7] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B20, 4150 (1979).

[8] A. Mazur, J. Pollmann, and M. Schmeits, Solid State Commun. 42, 37 (1982).

[9] A. McKinley, G. P. Srivastava, and R. H. Williams, J. Phys. C 13, 1581 (1980); R. H. Williams et al., to be published.

[10] M. Schmeits, A. Mazur, and J. Pollmann, Solid State Commun. 40, 1081 (1981).

[11] A. Ebina, T. Unno, Y. Suda, H. Koinuma, and T. Takahashi, J. Vac. Sci. Technol. 19, 301 (1981).

[12] F. Manghi, E. Molinari, C. M. Bertoni, and C. Calandra, J. Phys. C 15, 1099 (1982).

[13] R. E. Allen, H. P. Hjalmarson, and J. D. Dow, Surf. Sci. 110, 2625 (1981).

[14] R. P. Beres, R. E. Allen, and J. D. Dow, Solid State Commun. 45, 13 (1983).

[15] R. P. Beres, R. E. Allen and J. D. Dow, Phys. Rev. B26, 769 (1982).

[16] R. P. Beres, R. E. Allen, J. P. Buisson, M. A. Bowen, G. F. Blackwell, H. P. Hjalmarson, and J. D. Dow, J. Vac. Sci. Technol. 21, 548 (1982).

[17] R. P. Beres, R. E. Allen, and J. D. Dow, Phys. Rev. B26, 5702 (1982).

[18] J. Bardeen, Phys. Rev. 71, 717 (1947).

[19] W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. 16, 1422 (1979), and references therein.

[20] W. E. Spicer, I. Lindau, P. R. Skeath, and C. Y. Su, J. Vac. Sci. Technol. 17, 1019 (1980), and references therein.

[21] W. E. Spicer, I. Lindau, P. R. Skeath, C. Y. Su, and P. W. Chye, Phys. Rev. Letters 44, 520 (1980).

[22] P. Skeath, C. Y. Su, I. Hino, I. Lindau, and W. E. Spicer, Appl. Phys. Lett. 39, 349 (1981).

[23] J. D. Dow, D. R. Franceschetti, P. C. Gibbons, and S. E. Schnatterly, J. Phys. F 5, L211 (1975) and references therein.

[24] H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B24, 6010 (1981).

[25] C. A. Mead and W. G. Spitzer, Phys. Rev. A134, 713 (1964).

[26] S. M. Sze, Physics of Semiconductor Devices (Wiley, New York, 1981).

[27] L. J. Brillson, Phys. Rev. Letters 40, 260 (1978), and references therein.

[28] R. H. Williams, V. Montgomery, and R. R. Varma, J. Phys. C 11, L735 (1978).

[29] R. H. Williams and M. H. Patterson, Appl. Phys. Lett. 40, 484 (1982).

[30] R. H. Williams, to be published.

[31] M. S. Daw and D. L. Smith, Phys. Rev. B20, 5150 (1979); J. Vac. Sci. Technol. 17, 1028 (1980); Appl. Phys. Lett. 36, 690 (1980); Solid State Commun. 37, 205 (1981); M. S. Daw, D. L. Smith, C. A. Swarts, and T. C. McGill, J. Vac. Sci. Technol. 19, 508 (1981).

[32] R. E. Allen and J. D. Dow, Phys. Rev. B24, 911 (1981).

[33] R. E. Allen and J. D. Dow, J. Vac. Sci. Technol. 19, 383 (1981).

[34] R. E. Allen and J. D. Dow, Appl. Surf. Sci. 11/12, 362 (1982).

[35] R. E. Allen, J. D. Dow, and H. P. Hjalmarson, Solid State Commun. 41, 419 (1982).

[36] J. D. Dow and R. E. Allen, J. Vac. Sci. Technol. 20, 659 (1982).

[37] R. E. Allen, R. P. Beres, and J. D. Dow, J. Vac. Sci. Technol. B1, 401 (1983).

[38] H. H. Wieder, Inst. Phys. Conf. Ser. 50, 234 (1980).

[39] H. H. Wieder, Appl. Phys. Lett. 38, 170 (1981).

[40] H. H. Wieder, to be published.

[41] W. Mönch and H. J. Clemens, J. Vac. Sci. Technol. 16, 1238 (1979).

[42] H. Gant and W. Mönch, Appl. Surf. Sci. 11/12, 332 (1982).

[43] W. Mönch and H. Gant, Phys. Rev. Letters 48, 512 (1982).

[44] W. Mönch, R. S. Bauer, H. Gant, and R. Murschall, J. Vac. Sci. Technol. 21, 498 (1982). J. Assmann and W. Mönch, Surf. Sci. 99, 34 (1980).

[45] W. Mönch, to be published.

[46] O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1983).

[47] J. L. Freeouf and J. M. Woodall, Appl. Phys. Lett. 39, 727 (1981); J. L. Freeouf, Solid State Commun. 30, 1059 (1980); J. M. Woodall, G. D. Petit, T. N. Jackson, C. Lanza, K. L. Kavanagh, and J. W. Mayer, Phys. Rev. Letters 51, 1783 (1983).

[48] R. Ludecke and L. Esaki, Phys. Rev. Letters 33, 653 (1974); R. Ludecke and A. Kona, Phys. Rev. Letters 34, 817 (1975); Phys. Rev. Letters 39, 1042 (1977); and to be published.

[49] P. S. Ho and G. W. Rubloff, Thin Solid Films, 89, 433 (1981).

[50] J. M. Andrews and J. C. Phillips, Phys. Rev. Letters 35, 56 (1975).

[51] A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Solid State Phys. Suppl. 3, 132 (1963).

[52] M. Lannoo and P. Lenglart, J. Phys. Cs 30, 2409 (1969).

[53] S. T. Tong, A. R. Lubinsky, B. J. Mrstik, and M. A. van Hove, Phys. Rev. B17, 3303 (1978).

[54] O. F. Sankey and J. D. Dow, Phys. Rev. B27, 7641 (1983).

[55] O. F. Sankey, H. P. Hjalmarson, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. Letters 45, 1656 (1980); O. F. Sankey and J. D. Dow, Appl. Phys. Lett. 38, 685 (1981); J. Appl. Phys. 52, 5139 (1981); Phys. Rev. B26, 3243 (1982).

[56] H. Wieder, to be published.

[57] H. H. Wieder, private communication.

[58] W. Mönch, to be published.

[59] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[60] W. M. Gibson, to be published, has recently proposed that the relaxation of the (110) surface of GaAs is considerably less than once thought [53], based on his He-atom scattering data.

[61] A. Kahn, E. So, P. Mark, and C. B. Duke, J. Vac. Sci. Technol. 15, 580 (1978).

[62] C. B. Duke, R. J. Meyer, and P. Mark, J. Vac. Sci. Technol. 17, 971 (1980).

[63] C. B. Duke, A. Paton, W. K. Ford, A. Kahn, and J. Carelli, Phys. Rev. B24, 562 (1981); C. B. Duke, A. Paton, W. K. Ford, A. Kahn, and G. Scott, Phys. Rev. B24, 3310 (1981).

[64] J. A. Knapp and G. J. Lapeyre, J. Vac. Sci. Technol. 13, 757 (1976); Nuovo Cimento 39B, 693 (1977).

[65] G. P. Williams, R. J. Smith, and G. J. Lapeyre, J. Vac. Sci. Technol. 15, 1249 (1978).

[66] A. Huijser, J. van Laar, and T. L. Rooy, Phys. Lett. 65A, 337 (1978).

[67] J. van Laar and J. J. Scheer, Surf. Sci. 8, 342 (1967); J. van Laar and A.

Huijser, J. Vac. Sci. Technol. 13, 769 (1976).

[68] W. Gudat and D. E. Eastman, J. Vac. Sci. Technol. 13, 831 (1976).

[69] A. Huijser and J. van Laar, Surf. Sci. 52, 202 (1975); A. Huijser, J. van Laar, and T. L. van Rooy, Surf. Sci. 62, 472 (1977).

[70] G. M Guichar, C. A. Sebenne, and C. D. Thualt, J. Vac. Sci. Technol. 16, 1212 (1979); D. Norman, I. T. McGovern, and C. Norris, Phys. Lett. 63A, 384 (1977).

[71] P. Chiaradia, G. Chiarotti, F. Ciccacci, R. Memeo, S. Nannarone, P. Sassaroli, and S. Selci, Surf. Sci. 99, 76 (1980).

[72] V. Dose, H.-J. Gassmann, and D. Straub, Phys. Rev. Letters 47, 608 (1981).

[73] Further work is cited in the papers above and in the review by J. Pollmann, Festkorperprobleme, Vol. XX, p. 117.

[74] A. Zunger, Phys. Rev. B22, 959 (1980).

[75] See the proceedings of the Ninth International Conference on the Physics and Chemistry of Semiconductor Interfaces; J. Vac. Sci. Technol. 21 (1982).

[76] D. Haneman, Phys. Rev. 121, 1093 (1961).

[77] R. Feder, W. Mönch, and P. P. Aver. J. Phys. C 12, 2179 (1979).

[78] J. P. Buisson, J. D. Dow, and R. E. Allen, Surf. Sci. 120, L477 (1982), and references therein.

[79] K. C. Pandey, Phys. Rev. Letters 47, 1913 (1981).

[80] D. E. Eastman, J. Vac. Sci. Technol. 17, 492 (1980).

[81] D. J. Chadi, Appl. Optics 19, 3971 (1980).

[82] A. Redondo and W. A. Goddard, J. Vac. Sci. Technol. 21, 344 (1982).

[83] D. J. Chadi, Phys. Rev. Letters 43, 43 (1979).

[84] J. Ihm, M. L. Cohen, and D. J. Chadi, Phys. Rev. B21, 4952 (1980).

[85] M. A. Bowen, J. D. Dow, and R. E. Allen, Phys. Rev. B26, 7083 (1982).

[86] A. Mazur, J. Pollmann, and M. Schmeits, Phys. Rev. B26, 7086 (1982).

[87] W. Mönch, P. Koke, and S. Krueger, J. Vac. Sci. Technol. 19, 313 (1981).

[88] R. E. Allen and J. D. Dow, Phys. Rev. B24, 911 (1981). See also, R. E. Allen, H. P. Hjalmarson, H. Büttner, P. Vogl, D. J. Wolford, O. F. Sankey, and J. D. Dow, Internatl. J. Quantum Chem., Quantum Chem. Sympos. 14, 607 (1980).

[89] R. P. Beres, R. E. Allen, and J. D. Dow, Phys. Rev. B26, 769 (1982).

[90] T. Takahashi and A. Ebina, Appl. Surf. Sci. 11/12, 268 (1982).

FIGURE CAPTIONS

Fig. 1. Schematic illustration of Fermi-level pinning. Band edges of the bulk semiconductor, the semiconductor surface, and the Fermi energy of the metal, the surface of the semiconductor, and the semiconductor are all shown as functions of position. The lowest energy surface defect level that is not fully occupied (before charge is allowed to flow) is denoted by an open circle. This level and the Fermi levels of the n-type semiconductor and the metal align.

Fig. 2. Surface Fermi energy of n-type InP versus heat of reaction of InP with the metals Ni, Fe, Al, Cu, Ag, and Au, extracted from data of Ref. [28], assuming Fermi-level pinning. The theoretical Fermi-level pinning defect levels for the surface P-vacancy (V_P), the native antisite defects (In_P and P_{In}), and the extrinsic impurities S on a P-site (S_P) and Sn on a surface In site (Sn_{In}) are given at the right of the figure. The n-InP data can be interpreted as follows: non-reactive metals produce only antisite defects as the dominant defects; reactive metals and treatment of the surface with oxygen and Cl produce P-vacancies. Treatments with Sn and S produce surface Sn_{In} and S_P as dominat defects, respectively.

Fig. 3. The totally symmetric (a_1) levels for a bulk Si vacancy, surrounded by one Si atom and three X atoms, as a function of the defect potential V, normalized to the Ni defect potential, after Ref. [46]. For V=0, the X atoms are Si; for $V=V_{Ni}$, the X atoms are Ni. The parent levels of the isolated Si vacancy are shown for V=0. The experimental Fermi-level pinning position for $NiSi_2$ extracted from the data of G. Ottaviani, K. N. Tu, and J. W. Mayer, Phys. Rev. B24, 3354 (1981) are denoted by a dot with a label $NiSi_2$.

Fig. 4. One type of interfacial vacancy "sheltering" a Si dangling bond, after Ref. [46]. The geometry is that determined for the $NiSi_2/Si(111)$ interface determined by D. Cherns, G. R. Anstis, J. L. Hutchison, and J. C. H. Spence, Phil. Mag. A46, 849 (1982).

Fig. 5. Predicted dependence of Schottky barrier height on alloy compositions x and y of $In_{1-x}Ga_xAs$ and $Ga_{1-y}Al_yAs$ alloys compared with data, after Ref. [88].

Fig. 6. Predicted surface state dispersion curves $E(\bar{k})$ for surface bound states (solid lines) and surface resonances (dashed lines) at the relaxed (110) surface of GaAs, after Ref. [14]. The energy is plotted as a function of the planar wavevector $\bar{k}$ along the symmetry lines of the surface Brillouin zone, shown on the right. The labelling is the same as that of Chelikowsky and Cohen (Ref. [7]), with A_1 A_2, C_1, and C_2 mainly s-like, and A_3 A_4, A_5, C_3, and C_4 mainly p-like. A_5 and C_3 are the "dangling-bond" states. A_3, A_1', and A_2' are largely associated with in-plane p-orbitals in the first and second layers. The character of each state varies somewhat with the planar wavevector $\bar{k}$, and represents an admixture of all orbitals. The widths of the resonances are typically 0.5 to 1.0 eV, but in some cases are smaller than 0.1 eV or as large as 2.0 to 5.0 eV. The dots follow the continuous disperson curves inferred by Huijser et al. (Ref. [66]) for the "clear" and "weak" experimental features. The open squares represent the states observed by Williams et al. (Ref. [65]). The data reported in Refs. [64] and [6] are consistent with those shown here.

Fig. 7. Predicted energies of surface bound states (solid lines) and surface resonances (dashed) for the (110) surface of ZnSe, as function of the planar wavevector $\bar{k}=(k_1,k_2)$, after Ref. [15]. The surface Brillouin zone is shown on the right; $\bar{\Gamma}$ is the origin, $\bar{k}=(0,0)$. The bulk bands are shaded. E_v and E_c are the valence and conduction band edges. The experimental features identified with bound and resonant surface states in Ref. [11], along the two symmetry lines $\bar{\Gamma}\bar{X}'$ and $\bar{X}\bar{\Gamma}$, are indicated by the dotted lines.

Fig. 8. Dispersion curves for surface states and surface resonances at the (100) (2×1) surface of Si, after Ref. [85]. The energy E is shown as a function of the planar wavevector $\bar{k}$ around the symmetry lines of the surface Brillouin zone. Solid lines represent results of the present calculations; dashed lines are the measurements of R. I. G. Uhrberg, G. V. Hansson, J. M. Nicolle, and S. A. Flodstrom, Phys. Rev. B24, 4684 (1981); and the dotted line is the measurement of F. J. Himpsel and D. E. Eastman, J. Vac. Sci. Technol. 16, 1297 (1979), which were taken from $\bar{\Gamma}$ to $\bar{J}'$ along the (010) direction, rather than along the symmetry line $\bar{\Gamma}$ to $\bar{J}'$. E_v and E_c are the Si valence and conduction band edges.

Fig. 9. Predicted and observed Ga 3d core surface Frenkel excitons (double lobes) for GaAs, GaSb, and GaP, after Ref. [32]. The lower unoccupied surface states (Ref. [13][14]) are represented by closely spaced horizontal lines. E_v and E_c are, respectively, the top of the valence band and the bottom of the conduction band. The experimental results here and in Fig. 7 are those of Eastman and co-workers (D. E. Eastman and J. L. Freeouf, prl 33, 1601 (1974); 34, 1624 (1975); W. Gudat and D. E. Eastman, J. Vac. Sci. Technol. 13, 831 (1976); D. E. Eastman, T.-C. Chiang, P. Heimann, and F. J. Himpsel, prl

45, 656 (1980).).

Fig. 10. Predicted and observed In 4d core surface Frenkel excitons for InAs, InSb, and InP, after Ref. [32].

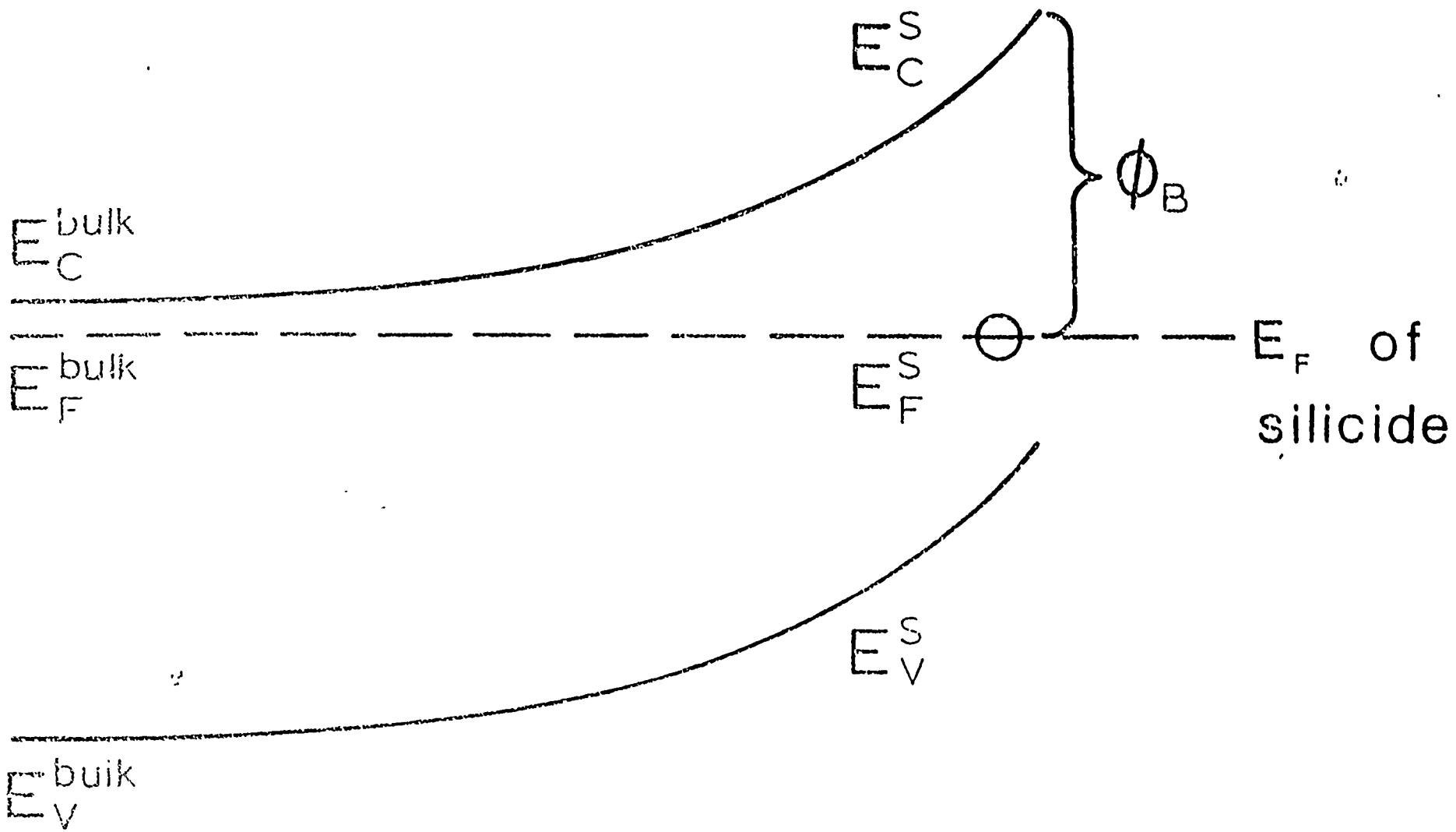

E_C^S
ϕ_B
E_C^{bulk}
E_F^{bulk}
E_F^S
E_F of silicide
E_V^S
E_V^{bulk}
Si
Interface

Fig. 1

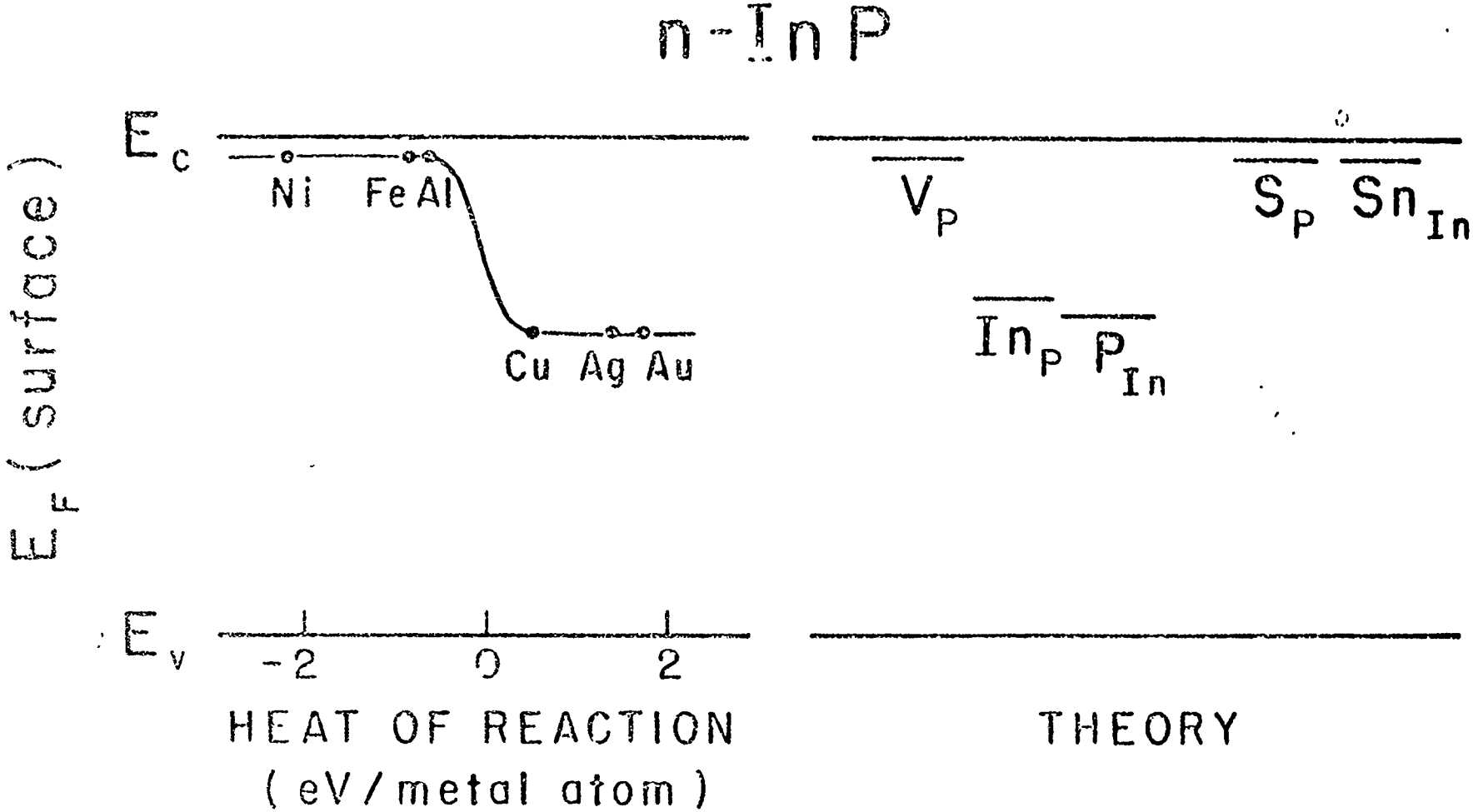

n-InP
E_c
E_v
E_F (surface)
Ni
Fe Al
Cu Ag Au
$\overline{V_P}$
$\overline{S_P}$ $\overline{Sn_{In}}$
$\overline{In_P}$ $\overline{P_{In}}$
-2
0
2
HEAT OF REACTION
(eV/metal atom)
THEORY

Fig. 2

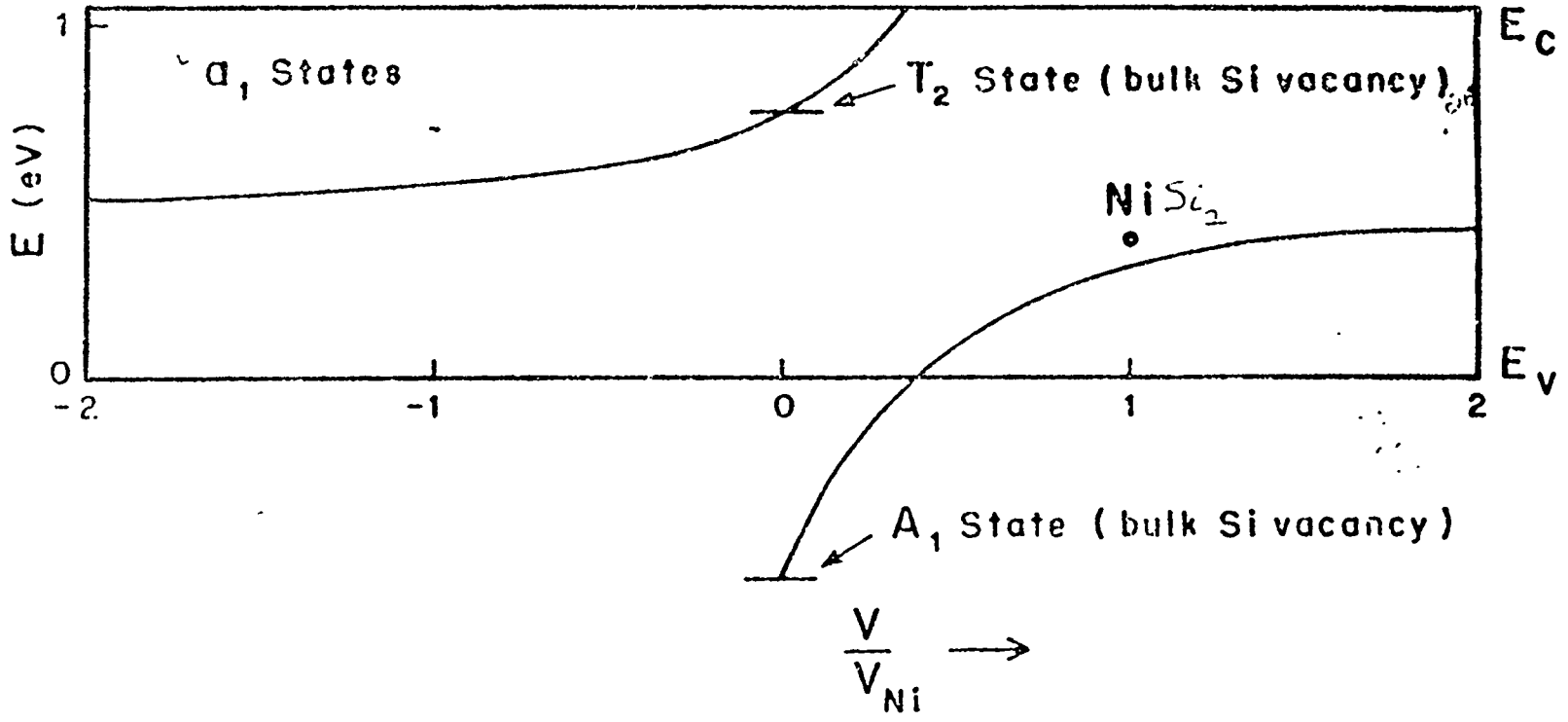

1
0
E (eV)
a_1 States
T_2 State (bulk Si vacancy)
E_C
$NiSi_2$
E_V
-2
-1
0
1
2
A_1 State (bulk Si vacancy)
$\frac{V}{V_{Ni}}$ →

Fig. 3:

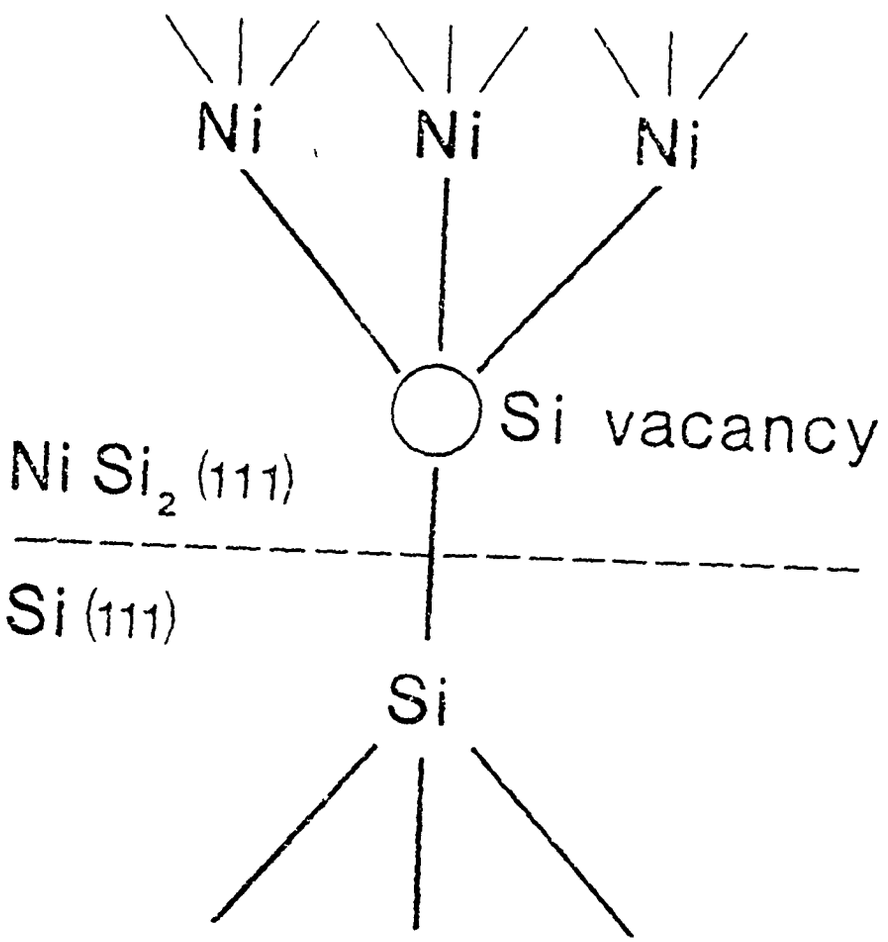


Fig. 4

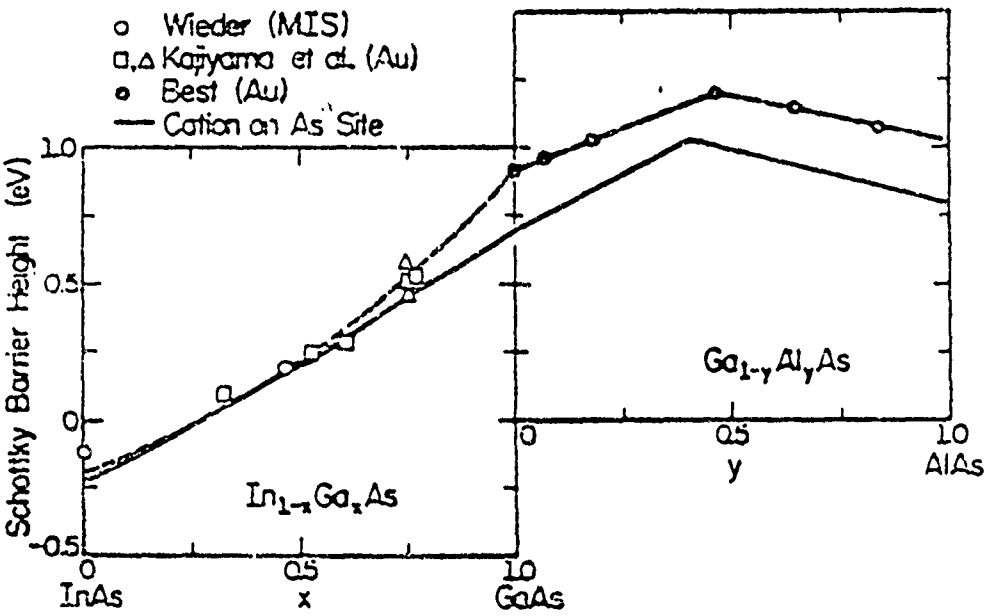
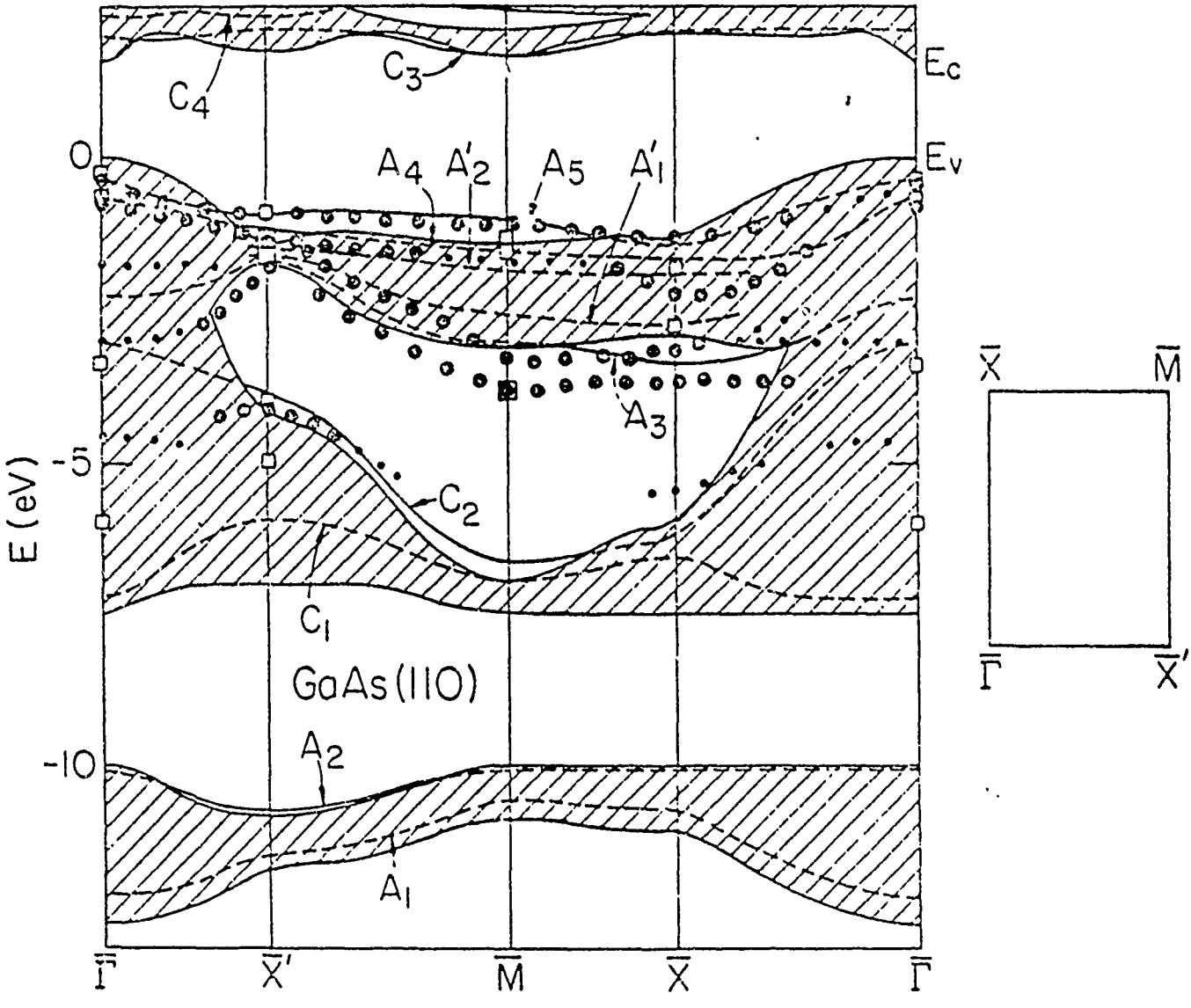

o Wieder (MIS)
□,Δ Kajiyama et al. (Au)
• Best (Au)
— Cation on As Site
Schottky Barrier Height (eV)
1.0
0.5
0
-0.5
0
InAs
0.5
x
1.0
GaAs
$In_{1-x}Ga_xAs$
$Ga_{1-y}Al_yAs$
0
0.5
y
1.0
AlAs

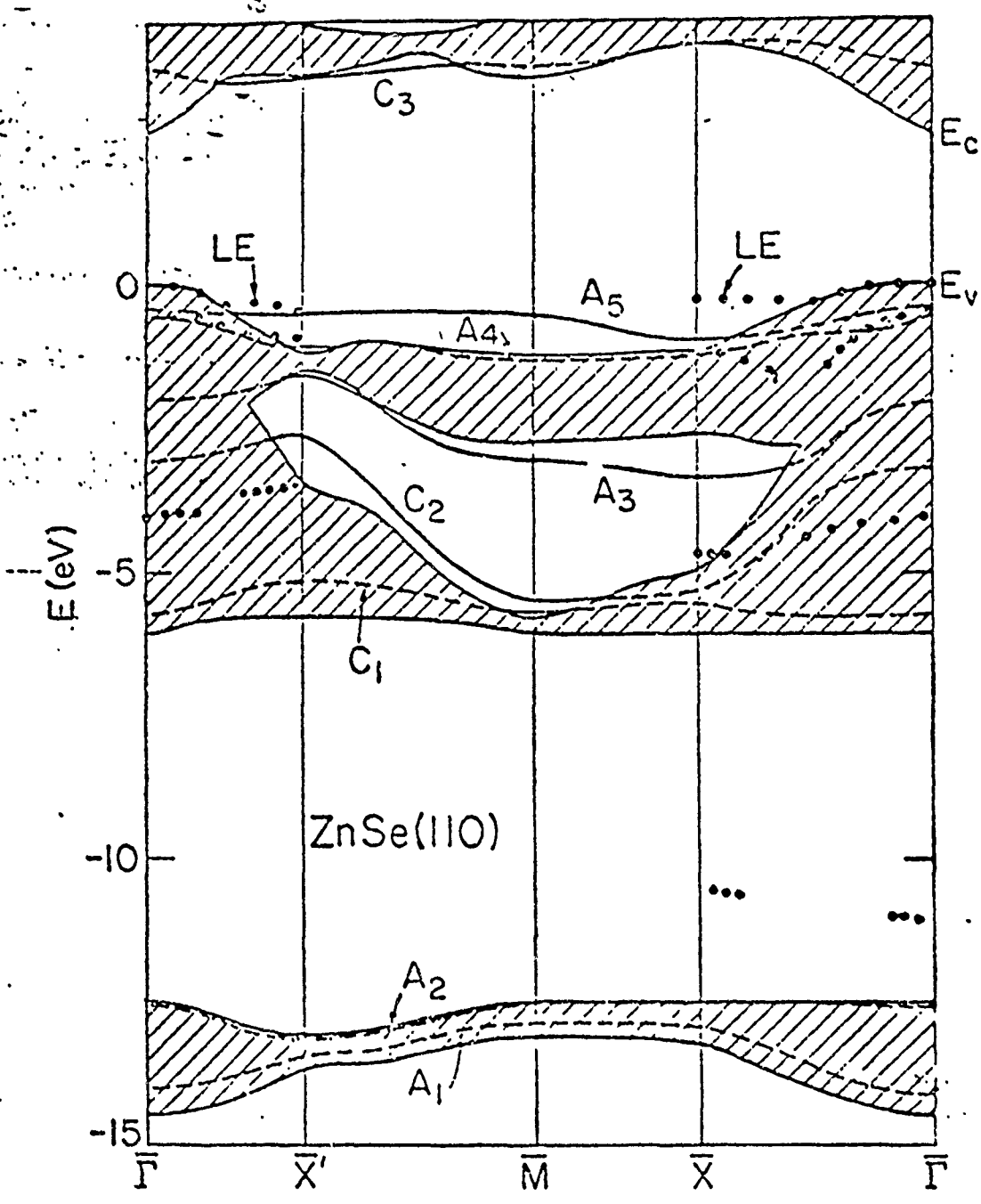
Fig. 5

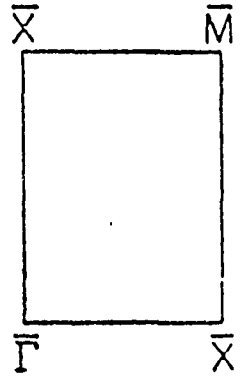
▫ Williams, Smith, and Lapeyre

o o o o clear
• • • • weak
} Huijser, van Laar, and van Rooy



F.g. 6





Experiment .....
($\bar{\Gamma}\bar{X}'$ and $\bar{X}\bar{\Gamma}$ only)

Fig 7

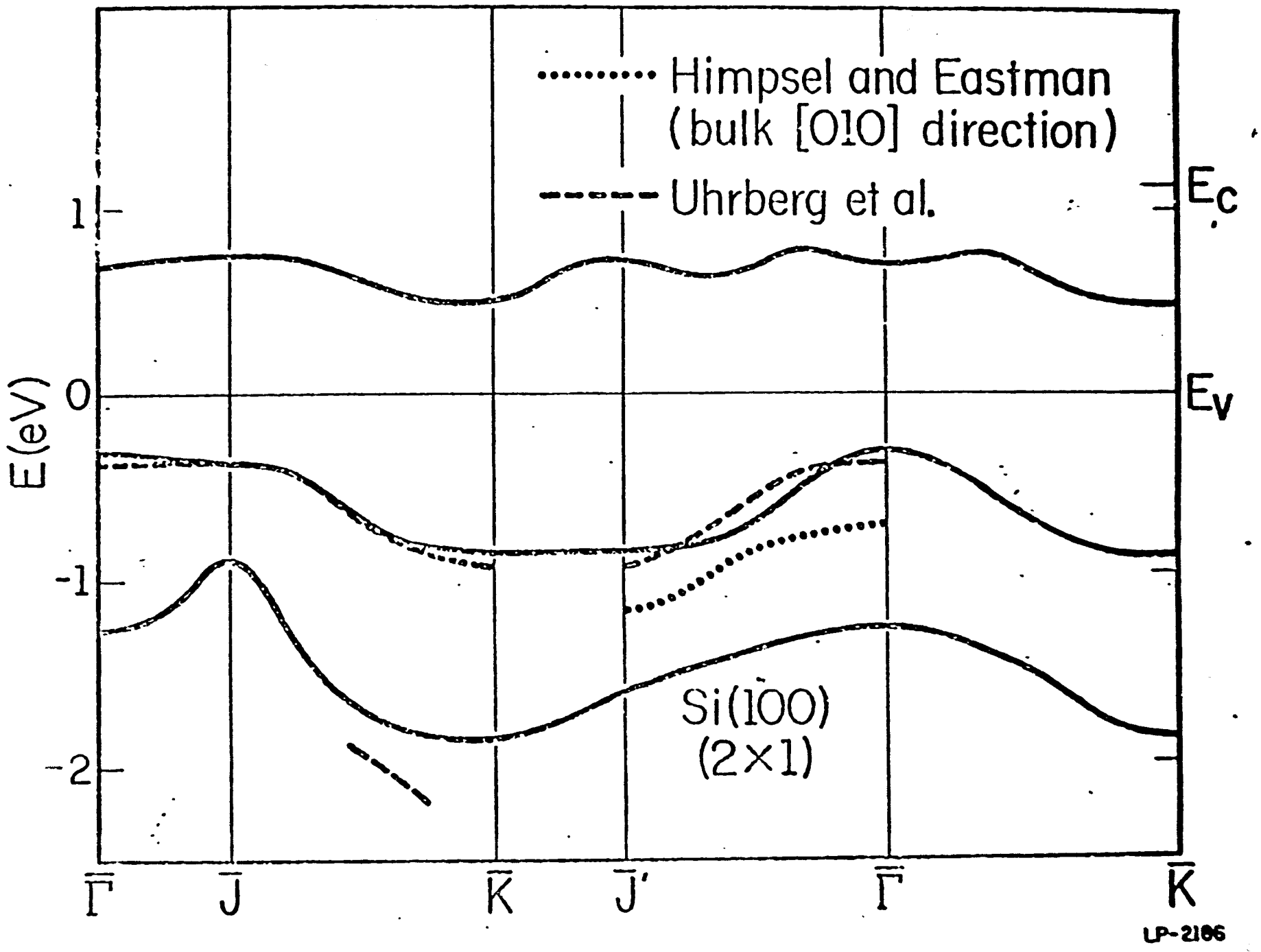

Himpsel and Eastman
(bulk [010] direction)
Uhrberg et al.
E_C
E_V
E(eV)
1
0
-1
-2
Si(100)
(2×1)
$\bar{\Gamma}$
$\bar{J}$
$\bar{K}$
$\bar{J}'$
$\bar{\Gamma}$
$\bar{K}$
LP-2186

Fig. 8

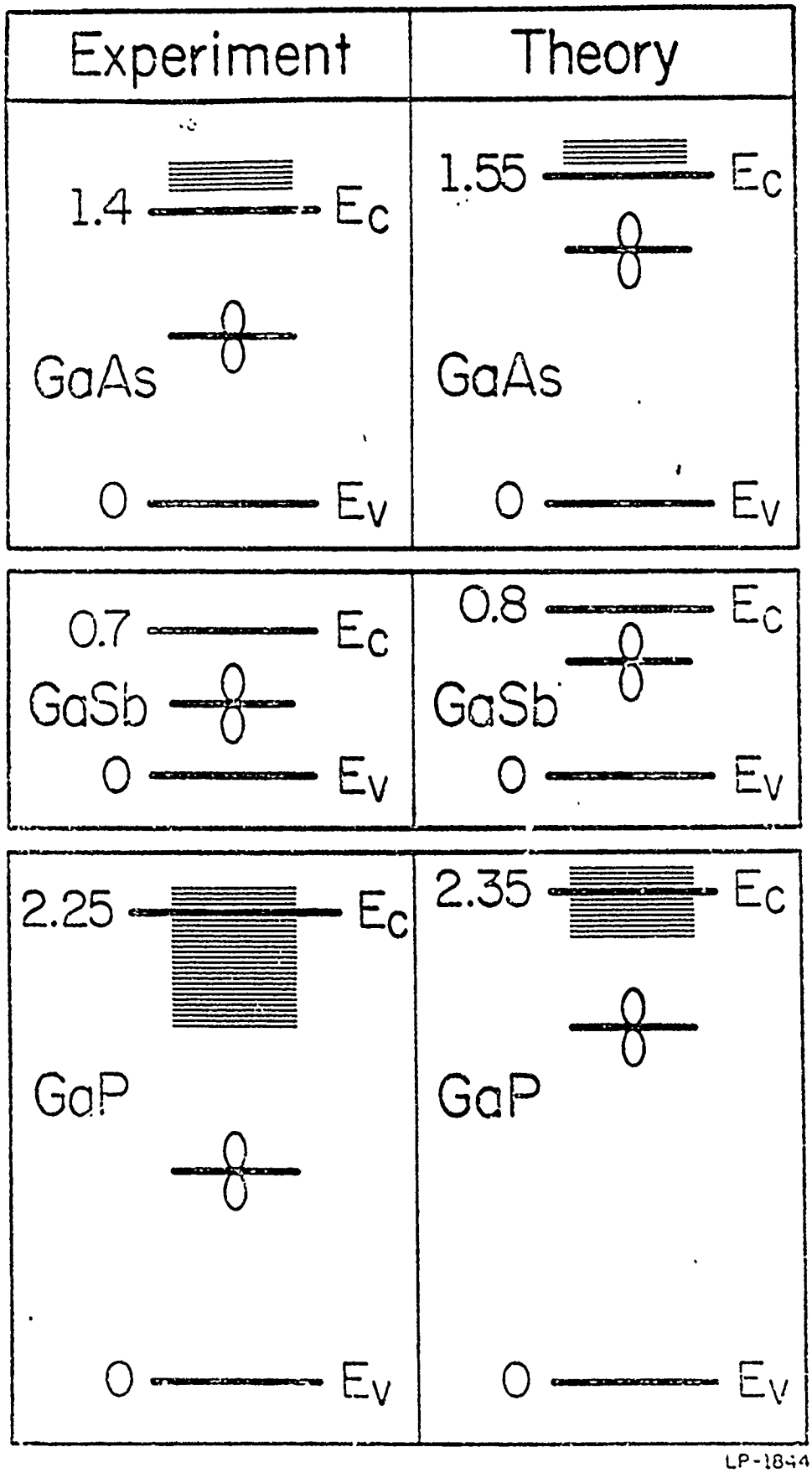
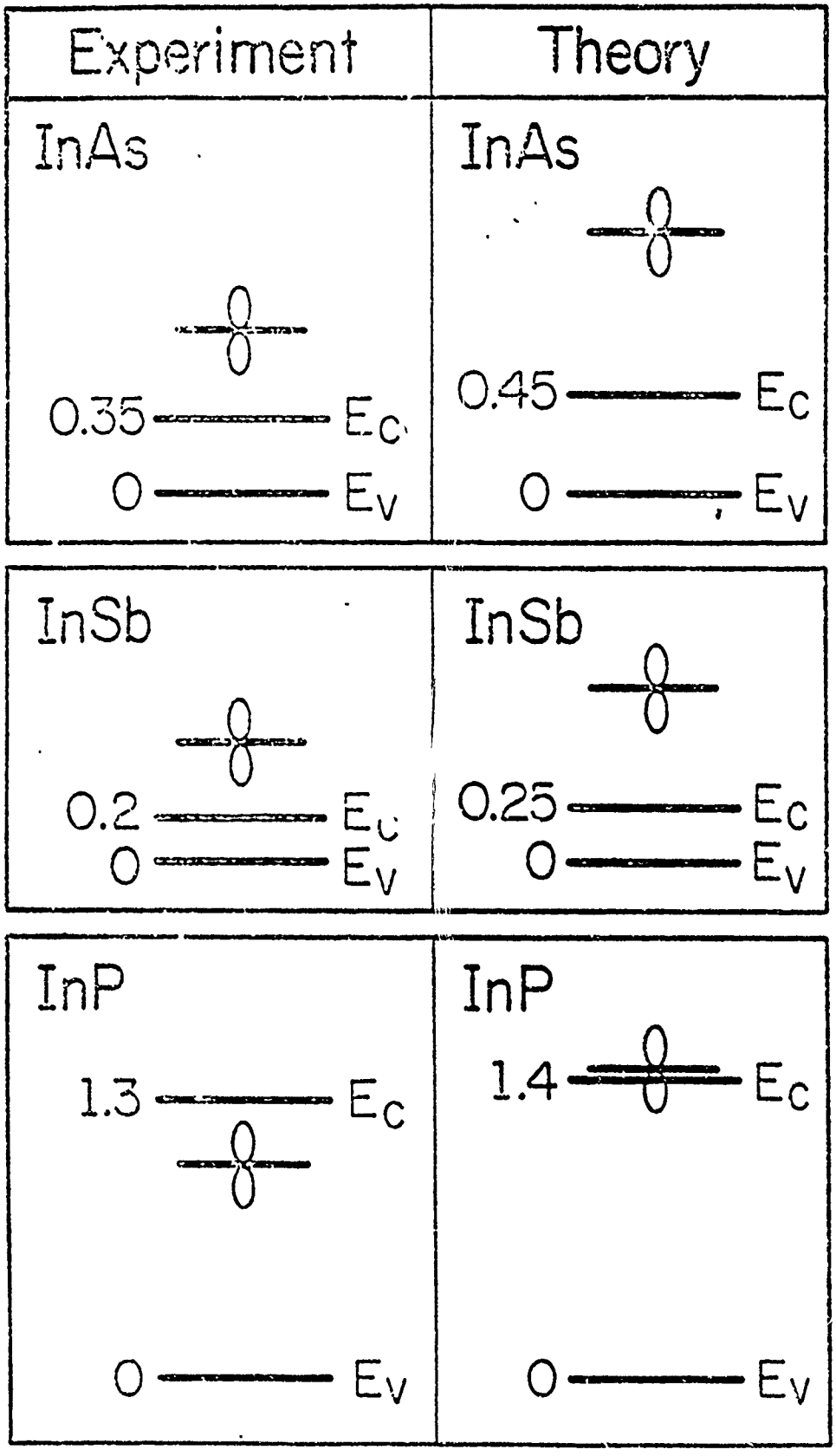

Experiment
Theory
1.4
Ec
GaAs
0
Ev
1.55
Ec
GaAs
0
Ev
0.7
Ec
GaSb
0
Ev
0.8
Ec
GaSb
0
Ev
2.25
Ec
GaP
0
Ev
2.35
Ec
GaP
0
Ev
LP-1844

Fig. 9



LP-1846

Fig. 10

Electronic properties of metastable Ge_xSn_{1-x} alloys

David W. Jenkins and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 21 May 1987)

The energy band gaps and substitutional deep impurity levels of metastable alloys Ge_xSn_{1-x} are predicted. As a function of decreasing alloy composition x the indirect band structure of semiconducting Ge first becomes direct (indicating that Ge_xSn_{1-x} may have applications as an infrared detector) and then metallic. Doping anomalies commonly occur as x decreases. Between $x \simeq 0.4$ and $x \simeq 0.8$, the Gunn effect should occur.

In this paper, we predict the band gaps and substitutional defect levels for alloys of germanium and tin: Ge_xSn_{1-x}. These materials are normally immiscible for most compositions when grown under equilibrium conditions, but have been grown in substitutional, crystalline metastable states for compositions $x > 0.78$ using nonequilibrium growth techniques.[1-3] With increasingly sophisticated growth techniques, we anticipate that metastable Ge_xSn_{1-x} alloys will soon be available for a greater range in x.[4] One purpose of this paper is to outline the electronic structure of these new alloys, and to suggest that, for a restricted range of alloy compositions, they should support Gunn-effect oscillations. Hence we hope to stimulate efforts to grow these materials.

Germanium is an indirect-gap material, the fundamental energy band gap occurring at the L point of the Brillouin zone $[\mathbf{k}=(2\pi/a_L)(\frac{1}{2},\frac{1}{2},\frac{1}{2})]$, with a magnitude of 0.76 eV at low[5] temperature. Tin is a semimetal, a material with no band gap; its valence and conduction bands overlap at the Γ point $[\mathbf{k}=(0,0,0)]$. We predict Ge_xSn_{1-x} to have a fundamental band gap that varies from zero to 0.76 eV as a function of composition x.

I. BAND STRUCTURES

The energy bands of Ge_xSn_{1-x} alloys were predicted using the virtual-crystal approximation and a second-nearest-neighbor tight-binding model of the Koster-Slater type.[6] The parameters, for a first-nearest-neighbor tight-binding Hamiltonian, are taken to be those of Vogl *et al.*,[7] which are known to reproduce valence-band structures and the principal features of the lowest conduction bands for all zinc-blende and diamond covalent semiconductors. An important and nontrivial feature of the Vogl parameters is that they incorporate chemical trends, so changes in these parameters as the semiconductor composition varies are rather well-behaved functions of changes in atomic energies and bond lengths. The Vogl model, in its published form, is lacking two ingredients essential to a proper treatment of Ge_xSn_{1-x} alloys: (i) spin-orbit splitting (which is important for the large-Z Sn atom), and (ii) second-nearest-neighbor parameters (which are needed to correctly simulate the relative conduction-band minimum near point L along Δ). The spin-orbit effect has been incorporated by a number of authors; we share common notation with Ref. 8. Similarly, the second-neighbor interactions can be incorporated as they were for $Si_{1-x}Ge_x$ alloys by Newman and Dow.[9] The resulting Hamiltonian, in a basis of tight-binding states of wave vector $\mathbf{k}$, is

$$H_0(\mathbf{k})=\begin{vmatrix} H_s & 0 & H_{ss}^{\dagger} & H_{ps}^{\dagger} \\ 0 & H_p & H_{sp}^{\dagger} & H_{pp}^{\dagger} \\ H_{ss} & H_{sp} & H_s & 0 \\ H_{ps} & H_{pp} & 0 & H_p \end{vmatrix},$$

where H_s is

	$\lvert s^*\uparrow\rangle$	$\lvert s^*\downarrow\rangle$	$\lvert s\uparrow\rangle$	$\lvert s\downarrow\rangle$
$\lvert s^*\uparrow\rangle$	E_{s^*}	0	0	0
$\lvert s^*\downarrow\rangle$	0	E_{s^*}	0	0
$\lvert s\uparrow\rangle$	0	0	E_s	0
$\lvert s\downarrow\rangle$	0	0	0	E_s

and H_p is

	$\lvert\frac{3}{2},\frac{3}{2}\rangle$	$\lvert\frac{3}{2},\frac{1}{2}\rangle$	$\lvert\frac{3}{2},-\frac{1}{2}\rangle$	$\lvert\frac{3}{2},-\frac{3}{2}\rangle$	$\lvert\frac{1}{2},\frac{1}{2}\rangle$	$\lvert\frac{1}{2},-\frac{1}{2}\rangle$
$\lvert\frac{3}{2},\frac{3}{2}\rangle$	$E_p+\lambda$	$-W_{p_1}g_b/\sqrt{3}$	$iW_{p_1}g_b/\sqrt{3}$	0	$W_{p_1}g_b/\sqrt{6}$	$-i\sqrt{2}W_{p_1}g_b/\sqrt{3}$
$\lvert\frac{3}{2},\frac{1}{2}\rangle$	$-W_{p_1}g_b^*/\sqrt{3}$	$E_p+\lambda$	0	$iW_{p_1}g_b/\sqrt{3}$	0	$-W_{p_1}g_b/\sqrt{2}$
$\lvert\frac{3}{2},-\frac{1}{2}\rangle$	$-iW_{p_1}g_b^*/\sqrt{3}$	0	$E_p+\lambda$	$W_{p_1}g_b/\sqrt{3}$	$-W_{p_1}g_b/\sqrt{2}$	0
$\lvert\frac{3}{2},-\frac{3}{2}\rangle$	0	$-iW_{p_1}g_b^*/\sqrt{3}$	$W_{p_1}g_b^*/\sqrt{3}$	$E_p+\lambda$	$-i\sqrt{2}W_{p_1}g_b/\sqrt{3}$	$W_{p_1}g_b/\sqrt{6}$
$\lvert\frac{1}{2},\frac{1}{2}\rangle$	$W_{p_1}g_b^*/\sqrt{6}$	0	$-W_{p_1}g_b^*/\sqrt{2}$	$i\sqrt{2}W_{p_1}g_b^*/\sqrt{3}$	$E_p-2\lambda$	0
$\lvert\frac{1}{2},-\frac{1}{2}\rangle$	$i\sqrt{2}W_{p_1}g_b^*/\sqrt{3}$	$-W_{p_1}g_b^*/\sqrt{2}$	0	$W_{p_1}g_b^*/\sqrt{6}$	0	$E_p-2\lambda$

 © 1987 The American Physical Society

Here H_{ss} is

$$\begin{array}{c|cccc} & |s^*\uparrow\rangle & |s^*\downarrow\rangle & |s\uparrow\rangle & |s\downarrow\rangle \\ \hline |s^*\uparrow\rangle & 0 & 0 & 0 & 0 \\ |s^*\downarrow\rangle & 0 & 0 & 0 & 0 \\ |s\uparrow\rangle & 0 & 0 & U_{ss}g_0 & 0 \\ |s\downarrow\rangle & 0 & 0 & 0 & U_{ss}g_0 \end{array},$$

H_{sp} is

$$\begin{array}{c|cccccc} & |\frac{3}{2},\frac{3}{2}\rangle & |\frac{3}{2},\frac{1}{2}\rangle & |\frac{1}{2},-\frac{1}{2}\rangle & |\frac{3}{2},-\frac{3}{2}\rangle & |\frac{1}{2},\frac{1}{2}\rangle & |\frac{1}{2},-\frac{1}{2}\rangle \\ \hline |s^*\uparrow\rangle & iU_{s^*p}g_4^*/\sqrt{2} & -\sqrt{2}U_{s^*p}g_3^*/\sqrt{3} & iU_{s^*p}g_3^*/\sqrt{6} & 0 & U_{s^*p}g_3^*/\sqrt{3} & -iU_{s^*p}g_3^*/\sqrt{3} \\ |s^*\downarrow\rangle & 0 & iU_{s^*p}g_4^*/\sqrt{6} & -\sqrt{2}U_{s^*p}g_3^*/\sqrt{3} & iU_{s^*p}g_3^*/\sqrt{2} & iU_{s^*p}g_4^*/\sqrt{3} & -U_{s^*p}g_3^*/\sqrt{3} \\ |s\uparrow\rangle & iU_{sp}g_4^*/\sqrt{2} & -\sqrt{2}U_{sp}g_3^*/\sqrt{3} & iU_{sp}g_3^*/\sqrt{6} & 0 & U_{sp}g_3^*/\sqrt{3} & -iU_{sp}g_3^*/\sqrt{3} \\ |s\downarrow\rangle & 0 & iU_{sp}g_4^*/\sqrt{6} & -\sqrt{2}U_{sp}g_3^*/\sqrt{3} & iU_{sp}g_3^*/\sqrt{2} & iU_{sp}g_4^*/\sqrt{3} & -U_{sp}g_3^*/\sqrt{3} \end{array},$$

H_{ps} is

$$\begin{array}{c|cccc} & |s^*\uparrow\rangle & |s^*\downarrow\rangle & |s\uparrow\rangle & |s\downarrow\rangle \\ \hline |\frac{3}{2},\frac{3}{2}\rangle & iU_{s^*p}g_3^*/\sqrt{2} & 0 & iU_{sp}g_3^*/\sqrt{2} & 0 \\ |\frac{3}{2},\frac{1}{2}\rangle & \sqrt{2}U_{s^*p}g_3^*/\sqrt{3} & iU_{s^*p}g_3^*/\sqrt{6} & \sqrt{2}U_{sp}g_3^*/\sqrt{3} & iU_{sp}g_3^*/\sqrt{6} \\ |\frac{3}{2},-\frac{1}{2}\rangle & iU_{s^*p}g_4^*/\sqrt{6} & \sqrt{2}U_{s^*p}g_3^*/\sqrt{3} & iU_{sp}g_4^*/\sqrt{6} & \sqrt{2}U_{sp}g_3^*/\sqrt{3} \\ |\frac{3}{2},-\frac{3}{2}\rangle & 0 & iU_{s^*p}g_4^*/\sqrt{2} & 0 & iU_{sp}g_4^*/\sqrt{2} \\ |\frac{1}{2},\frac{1}{2}\rangle & -U_{s^*p}g_3^*/\sqrt{3} & iU_{s^*p}g_3^*/\sqrt{3} & -U_{sp}g_3^*/\sqrt{3} & iU_{sp}g_3^*/\sqrt{3} \\ |\frac{1}{2},-\frac{1}{2}\rangle & -iU_{s^*p}g_4^*/\sqrt{3} & U_{s^*p}g_3^*/\sqrt{3} & -iU_{sp}g_4^*/\sqrt{3} & U_{sp}g_3^*/\sqrt{3} \end{array},$$

and H_{pp} is

$$\begin{array}{c|cccccc} & |\frac{3}{2},\frac{3}{2}\rangle & |\frac{3}{2},\frac{1}{2}\rangle & |\frac{3}{2},-\frac{1}{2}\rangle & |\frac{3}{2},-\frac{3}{2}\rangle & |\frac{1}{2},\frac{1}{2}\rangle & |\frac{1}{2},-\frac{1}{2}\rangle \\ \hline |\frac{3}{2},\frac{3}{2}\rangle & U_{xx}g_0^* & -U_{xy}g_4^*/\sqrt{3} & iU_{xy}g_3^*/\sqrt{3} & 0 & U_{xy}g_4^*/\sqrt{6} & -i\sqrt{2}U_{xy}g_3^*/\sqrt{3} \\ |\frac{3}{2},\frac{1}{2}\rangle & -U_{xy}g_3^*/\sqrt{3} & U_{xx}g_0^* & 0 & iU_{xy}g_3^*/\sqrt{3} & 0 & -U_{xy}g_4^*/\sqrt{2} \\ |\frac{3}{2},-\frac{1}{2}\rangle & -iU_{xy}g_3^*/\sqrt{3} & 0 & U_{xx}g_0^* & U_{xy}g_4^*/\sqrt{3} & -U_{xy}g_3^*/\sqrt{2} & 0 \\ |\frac{3}{2},-\frac{3}{2}\rangle & 0 & -iU_{xy}g_3^*/\sqrt{3} & U_{xy}g_3^*/\sqrt{3} & U_{xx}g_0^* & -i\sqrt{2}U_{xy}g_3^*/\sqrt{3} & U_{xy}g_3^*/\sqrt{6} \\ |\frac{1}{2},\frac{1}{2}\rangle & U_{xy}g_3^*/\sqrt{6} & 0 & -U_{xy}g_4^*/\sqrt{2} & i\sqrt{2}U_{xy}g_3^*/\sqrt{3} & U_{xx}g_0^* & 0 \\ |\frac{1}{2},-\frac{1}{2}\rangle & i\sqrt{2}U_{xy}g_3^*/\sqrt{3} & -U_{xy}g_3^*/\sqrt{2} & 0 & U_{xy}g_4^*/\sqrt{6} & 0 & U_{xx}g_0^* \end{array}.$$

We have used the notation of Kobayashi *et al.*[8] for all nearest-neighbor parameters. The second-neighbor parameters are $W_{sp}=4\langle s,\sigma,0|H_0|p,\sigma,d_2\rangle$, where d_2 is the displacement vector of a second neighbor.[10]

The first-neighbor parameters for Ge and Sn were fitted to the band structures of Chelikowsky and Cohen[5] using the method of Kobayashi *et al.*[8] The second-neighbor parameters were fit to the conduction-band edge at the L point using the same band structures. All parameters are given in Table I.

By diagonalizing this Hamiltonian, we obtain the band structures of Ge and Sn, which are in good agreement with the pseudopotential band structures of Chelikowsky and Cohen.[5] (See Figs. 1 and 2.) To obtain the virtual-crystal band structures of Ge_xSn_{1-x} alloys, we first average the parameters of Ge and Sn as follows: on-site parameters, $x[\mathrm{Ge}]+(1-x)[\mathrm{Sn}]$; off-diagonal parameters, $(x[\mathrm{Ge}]\{a(\mathrm{Ge})\}^{-2}+(1-x)[\mathrm{Sn}]\{a(\mathrm{Sn})\}^{-2})\{a(x)\}^{-2}$; where [Ge] and [Sn] are typical bulk Hamiltonian parameters of Ge and Sn, a(Ge) and a(Sn) are the lattice constants for Ge and Sn, respectively, and we assume Vegard's law:

$$a(x)=xa(\mathrm{Ge})+(1-x)a(\mathrm{Sn}).$$

This averaging procedure is a virtual-crystal approximation, and is valid because the Onodera-Toyozawa[11] ratio for Ge_xSn_{1-x} is considerably less than 0.1 for the conduction and valence bands. This ratio is the difference δ in on-site energies of the two constituent materials divid-

TABLE I. Sn and Ge tight-binding parameters (in eV, except d, which is in Å). We have added second-neighbor parameters to fit the gap at the L point. The notation is that of Kobayashi *et al.* (Ref. 8).

	Ge	Sn
E_s	−5.8800	−5.8800
E_p	1.5533	1.1733
λ	0.0967	0.2667
E_{s^*}	6.3900	5.9000
U_{ss}	−6.7800	−5.4600
U_{xx}	1.6500	1.4400
U_{xy}	4.8416	3.9042
U_{sp}	4.9520	4.0172
U_{s^*p}	4.5030	3.6459
$W_{ps'}$	0.1352	0.1229
d	2.45	2.81

Additional second-neighbor matrix elements

$$\left(\tfrac{3}{2},\tfrac{1}{2}\middle|H\middle|\tfrac{1}{2},\tfrac{3}{2}\right)=-W_{ps}g_6^*/\sqrt{3}$$
$$\left(\tfrac{3}{2},-\tfrac{1}{2}\middle|H\middle|\tfrac{3}{2},\tfrac{3}{2}\right)=-iW_{ps}g_8^*/\sqrt{2}$$
$$\left(\tfrac{1}{2},\tfrac{1}{2}\middle|H\middle|\tfrac{3}{2},\tfrac{3}{2}\right)=W_{ps}g_6^*/\sqrt{6},$$
$$\left(\tfrac{1}{2},-\tfrac{1}{2}\middle|H\middle|\tfrac{3}{2},\tfrac{3}{2}\right)=iW_{ps}g_8^*\sqrt{2/3}$$
$$\left(\tfrac{3}{2},-\tfrac{3}{2}\middle|H\middle|\tfrac{3}{2},\tfrac{1}{2}\right)=-iW_{ps}g_8^*/\sqrt{3}$$
$$\left(\tfrac{1}{2},-\tfrac{1}{2}\middle|H\middle|\tfrac{3}{2},\tfrac{1}{2}\right)=-W_{ps}g_6^*/\sqrt{2},$$
$$\left(\tfrac{3}{2},-\tfrac{3}{2}\middle|H\middle|\tfrac{1}{2},-\tfrac{1}{2}\right)=W_{ps}g_6^*/\sqrt{3}$$
$$\left(\tfrac{1}{2},\tfrac{1}{2}\middle|H\middle|\tfrac{3}{2},-\tfrac{1}{2}\right)=-W_{ps}g_7^*/\sqrt{2}$$

where

$$g_6=\sin(k_xa/2)\sin(k_ya/2)+i\sin(k_xa/2)\sin(k_za/2)$$
$$g_7=\sin(k_xa/2)\sin(k_ya/2)-i\sin(k_xa/2)\sin(k_za/2)$$
$$g_8=\sin(k_ya/2)\sin(k_za/2)$$

ed by the bandwidth W of the associated band. For Ge_xSn_{1-x}, the larger of the differences δ of s and p on-site energies is 0.38 eV. The conduction- and valence-band widths of Sn are 11.34 and 5.72 eV, respectively. (The bandwidths of Ge are comparable.) In this case, for Ge_xSn_{1-x} we have $\delta/W \lesssim 0.02$ for the valence band and

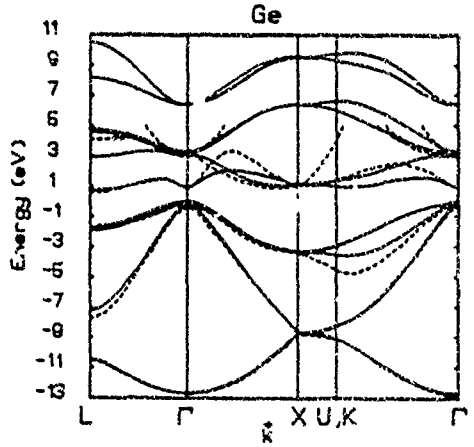


FIG. 1. Band structure of Ge using the present theory (solid lines) compared to the pseudopotential results of Chelikowsky and Cohen (Ref. 5) (dashed lines).

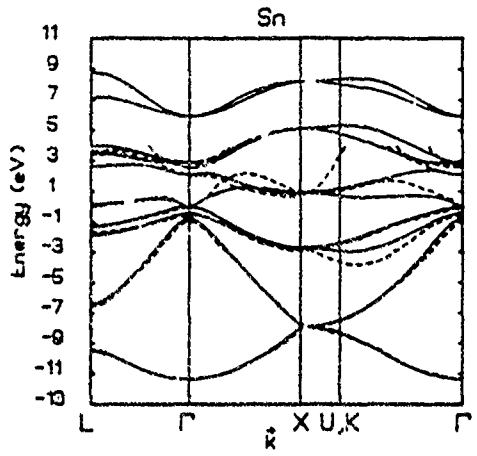


FIG. 2. Band structure of Sn (solid lines) in comparison with the results of Ref. 5 (dashed lines).

$\delta/W \lesssim 0.07$ for the conduction band. These materials satisfy the criterion as well as or better than alloys of GaAs and GaP.

The resulting virtual-crystal-approximation band structures of Ge_xSn_{1-x} are displayed in Figs. 3, 4, and 5 for $x=0.25$, 0.5, and 0.75, respectively.

Figure 6 displays, as functions of alloy composition x, the principal virtual-crystal band gaps at point Γ, point L, and point X [$k=(2\pi/a)(1,0,0)$], and Δ (Δ is the wave vector of the local minimum in the conduction band along the [100] direction; point L is a local minimum).

Interesting features of Fig. 6 are (i) that a direct (Γ)-to-indirect (L) crossover is predicted near $x\simeq 0.8$, (ii) the alloy's fundamental band gap is nonzero for $x>0.4$, and (iii) the level at Γ is lower in energy than the level at L for $x<0.8$. This means that Ge_xSn_{1-x} will be semimetallic for $x<0.4$, a semiconductor with a direct gap for $0.4<x<0.8$ (and hence a potential infrared detector or light emitter), and a potential Gunn oscillator for $0.4<x<0.8$. (See below.) For $x>0.8$, Ge_xSn_{1-x} is an indirect-gap semiconductor.

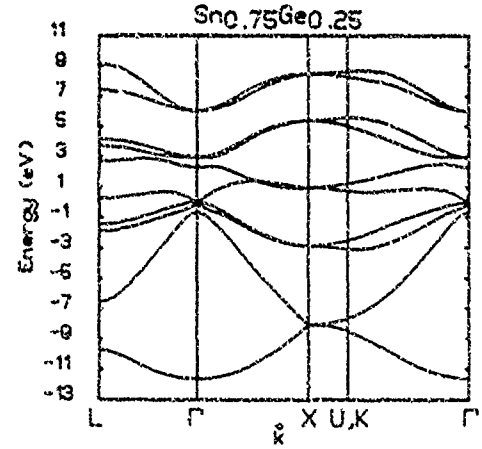


FIG. 3. Band structure of metastable $Ge_{0.25}Sn_{0.75}$

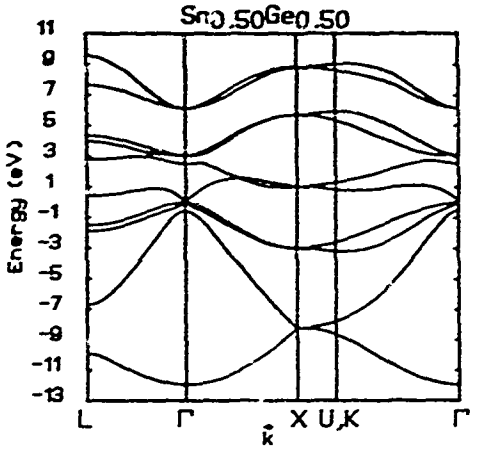


FIG. 5. Band structure of metastable $Ge_{0.5}Sn_{0.5}$.

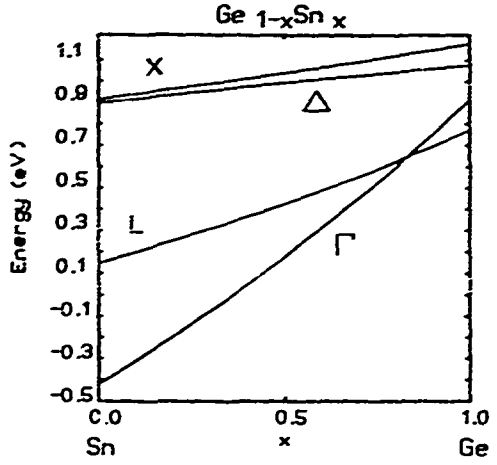


FIG. 6. Predicted lowest conduction bands at Γ, L, and X (the valence band is shaded) vs alloy composition x for Ge_xSn_{1-x}. The band gap varies from zero for $x=0.4$ to 0.76 eV for pure germanium. This covers energies corresponding to infrared light. The Gunn effect should occur for $0.4<x<0.8$ because the high-mobility low-effective-mass Γ minimum lies below the low-mobility L minimum. For $x<0.4$ the alloy is predicted to have zero gap.

II. DEEP IMPURITY LEVELS

A. General

The deep impurity levels are computed following the general approach of Hjalmarson *et al.*[7,12] Because of the chemical trends in the matrix elements, a defect potential matrix can be constructed rather easily. For substitutional defects that have the same bond length as the host atoms they replace, the matrix in a basis of localized orbitals centered at the defect site is

$$\begin{array}{c} \\ |E_6) \\ |E_8) \\ |E_7) \end{array} \begin{array}{c} \begin{array}{ccc} |E_6) & |E_8) & |E_7) \end{array} \\ \begin{pmatrix} \epsilon_i^s-\epsilon_h^s & 0 & 0 \\ 0 & \epsilon_i^{3/2}-\epsilon_h^{3/2} & 0 \\ 0 & 0 & \epsilon_i^{1/2}-\epsilon_h^{1/2} \end{pmatrix}, \end{array}$$

where ϵ is the on-site energy[13] of the host (h) or impurity (i). Note that the E_6 state is $s_{1/2}$-like, the E_7 state is $p_{1/2}$-like, and the E_8 state is $p_{3/2}$-like.

The effects of lattice relaxation around the defect and bond-length changes can be incorporated by noting that the off-diagonal matrix elements of the Hamiltonian[13] scale as the inverse square of the bond length. Here we neglect such lattice relaxation effects because (i) they are small, of order 0.1 eV, on the energy scales of relevance to the deep impurity problem—namely, the $\sim$10-eV bandwidths, and the $\sim$1–10-eV scale of the defect potential, and (ii) we are exploring the global chemical trends in the defect levels rather than attempting to predict with precision the energy levels of a specific defect in a single host—while the physics of the unrelaxed deep levels may exhibit well-defined trends, the lattice relaxation may be governed by different chemistry which might obscure the trends in the unrelaxed deep levels.

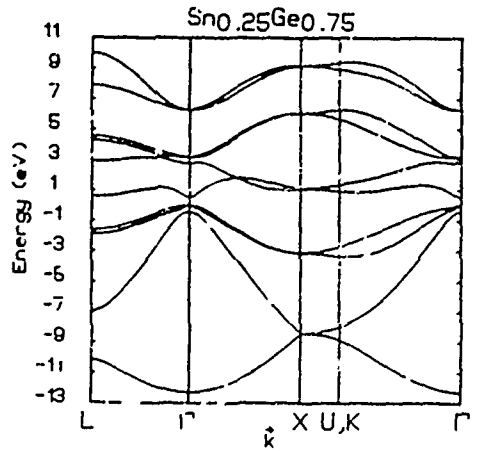


FIG. 5. Band structure of metastable $Ge_{0.75}Sn_{0.25}$

Introducing the Green's-function operator

$$G(E)=(E-H_0)^{-1},$$

where the energy E is to be interpreted as having an infinitesimal positive imaginary part when it lies in a host band, the Schrödinger equation for the deep level eigenvalues E is

$$(H_0+V)\psi=E\psi,$$

and leads to the secular equation

$$\det(1-GV)=0.$$

Here H_0 is the host-crystal Hamiltonian, $V\equiv H-H_0$ is the defect matrix, and 1 is the unit matrix. Invoking the diamond-crystal symmetry of virtual-crystal Ge_xSn_{1-x}, the secular equations reduce to the scalar equation

$$1/V_{\epsilon}=G_6(E)$$

for the doubly degenerate $s_{1/2}$-like E_6 level,

$$1/V_{1/2}=G_7(E)$$

for the doubly degenerate $p_{1/2}$-like E_7 levels, and

$$1/V_{3/2}=G_8(E)$$

for the fourfold-degenerate $p_{3/2}$-like E_8 levels. Here G_6, G_7, and G_8 are Green's functions for E_6, E_7, and E_8 symmetry, and are defined as follows:

$$G_l(E)=\sum_n \int dk^3 \frac{|\langle l',\mathbf{k},n|l',\mathbf{k},n\rangle|^2}{E-E_n(\mathbf{k})} ,$$

where $E_n(\mathbf{k})$ is the nth eigenvalue at wave vector $\mathbf{k}$ and $|l,\mathbf{k},n\rangle$ is the nth eigenvector at $\mathbf{k}$ with $l=s_{1/2}$, $p_{3/2}$, or $p_{1/2}$ for the E_6, E_8, or E_7 symmetry, respectively. The matrix elements V_l are deduced from the chemical trend:[7]

$$V_l=\beta_l(w_{l,\mathrm{imp}}-w_{l,\mathrm{host}}) ,$$

where $w_{l,\mathrm{imp}}$ and $w_{l,\mathrm{host}}$ are atomic energies for the impurity and host atoms, respectively, $l=s_{1/2}$, $p_{3/2}$, or $p_{1/2}$, and $\beta_l=0.8$ for $l=s_{1/2}$, and 0.6 for $l=p_{1/2}$ or $p_{3/2}$. We have $V_{3/2}=V_{1/2}=V_p$.

B. Deep levels

The predicted substitutional-impurity deep-level energies F in the fundamental band gap obtained by solving the secular equations for Ge_xSn_{1-x} are given in Figs. 7 and 8 for levels of E_6, E_7, and E_8 symmetry, respectively. The levels found for Ge are in generally good agreement with what is known about deep impurities in that material. The results of this model are comparable with those of other theories,[14,15] some of which are much more complicated. For example, we calculate the vacancy E_8 level to be 0.24 eV above the valence-band edge, compared to a range from 0.04 to 0.66 eV for other theories.[9] We estimate the uncertainty in our theory to be a few tenths of an eV, comparable with the claimed 0.2-eV uncertainty for self-consistent pseudopotential calculations.[15] In Table II we compare the present theory with experiment. The deep energy levels for S, Se, and Te, all from column VI of the Periodic Table, show a definite trend to higher energies for the series S to Se to Te. This trend is due to a reduction in the magnitude of the atomic orbital energies for the valence electrons of these impurities, hence the defect potential weakens. The trend is present both in theory and in experiment. While the predicted level for S precisely matches experiment (accidentally good agreement for a theory with an uncertainty of a few tenths of an eV), the theory also agrees with the data for Se, and places the Te deep level just above the conduction-band minimum, while the data reveal a level of 0.1 eV below the band edge (within the uncertainty).

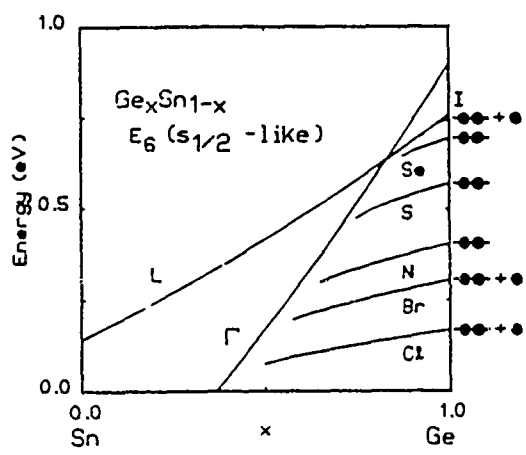


FIG. 7. Predicted substitutional deep impurity levels in Ge_xSn_{1-x} of E_6 (s-like) symmetry as a function of composition x. The zero of energy is the valence-band edge. The conduction-band edges at Γ and L are shown. Impurity levels in the gap for pure Ge are driven into the conduction band as x decreases. Occupancies of the neutral impurity states are shown on the right; electrons are solid circles and holes are open circles. An extra electron (denoted by $-\bullet$) would occupy a state near the conduction-band edge.

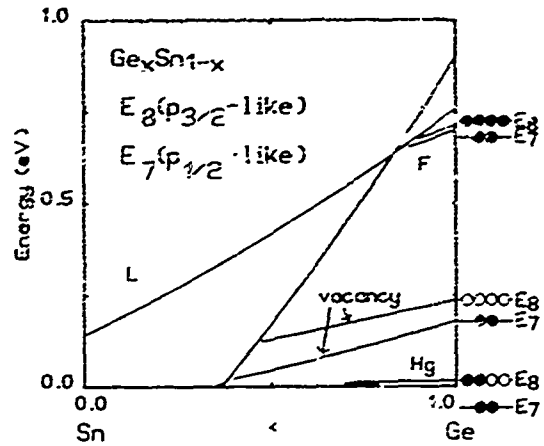


FIG. 8. Predicted substitutional impurity deep levels in Ge_xSn_{1-x} of E_7 ($p_{1/2}$-like) and E_8 ($p_{3/2}$-like) symmetry as functions of composition x. The levels are plotted relative to the valence-band edge. Occupancies of the states are shown; electrons are solid circles and holes are open circles. The conduction-band edges at Γ and L are shown. Impurity levels in the gap for pure germanium are driven into the conduction band as x decreases.

C. Doping anomalies

As the band gap decreases with increasing Sn composition, the deep levels lying in the fundamental band gap of Ge pass into either the conduction band or the valence band of the alloy. When this happens, a doping

TABLE II. Comparison of our calculated deep levels (in eV) in Ge with experimental values taken from W. W. Tyler, J. Phys. Chem. Solids 8, 59 (1959). The Te deep level in our theory is resonant with the conduction band so the ground state of the Te impurity has its two extra electrons in the effective-mass shallow levels.

	Deep levels	
Impurity	Present theory	Experiment
S	0.58	0.58
Se	0.69	0.62
Te	resonant	0.65

anomaly generally occurs. There are two types of common doping anomalies: (i) deep-shallow transitions, which occur when a deep level crosses a band edge, and (ii) false valences that result from a deep level crossing the fundamental band gap.

For clarity of discussion, we shall assume that the predicted deep-level energies are precisely correct, while cautioning the reader to make allowances for a few tenths of an eV uncertainty in the theory due to neglect of lattice relaxation and charge-state splitting of the levels:[16] for example, the Hg E_8 level, according to Fig. 8, is both an electron and a hole trap, but might actually lie below the valence-band maximum, donating its two holes to the valence band and becoming a double acceptor. (The holes are then trapped by the long-ranged Coulomb potential in shallow acceptor levels.)[1]

1. Deep-shallow transitions

All impurities with deep levels in the gap for Ge undergo a deep-shallow transition as the Sn composition increases. For example, the $p_{3/2}$-like E_8 Hg level is driven into the valence band (Fig. 8), while the other deep levels are driven into the conduction band (Fig. 7.)

When the Cl, Br, and I deep E_7 levels pass into the conduction band with decreasing x, the electrons that occupy the deep levels are autoionized, fall to the conduction-band minimum, and then are trapped in shallow levels. These impurities cease being deep hole traps (plus single donors) and instead become triple donors—the status they would hold in a naive effective-mass theory which contained no deep levels.

Similarly S and Se are deep (double-hole) traps in Ge but become double donors for smaller x (see Fig. 7). N is a deep (electron and hole) trap in Ge, but becomes a shallow donor for $x < 0.6$. The Hg E_8 level traps two electrons and two holes (if the theory is taken literally) in Ge, but Hg becomes a double acceptor with increasing Sn content. Finally the vacancy, which is a deep trap in Ge capable of capturing four electrons or two holes, becomes a double donor when both of its levels enter the conduction band (but is only a hole trap when the E_8 level is in the conduction band and the E_7 level is in the gap).

2. False valences

Substitutional oxygen displays a false valence of zero with respect to Sn or Ge, instead of -2. To see how this happens, consider Fig. 9, which displays the predictions for substitutional impurities from row 2 of the Periodic Table in Ge. The s-like and p-like levels in the conduction band of Ge for a column-IV defect (C) move down in energy as one moves to the right in the Periodic Table. The s-like level lies in the gap for N, but crosses the gap into the valence band for oxygen and F. Similarly the p-like E_7 and E_1 levels descend into the gap for F. Because its s-like E_6 deep level has crossed the gap into the valence band and contains two electrons, neutral oxygen produces neither a double donor (effective-mass intuition) nor a deep trap. Instead neutral oxygen is inert, neither trapping, nor donating, nor accepting electrons. It has a false valence of zero with respect to Ge.

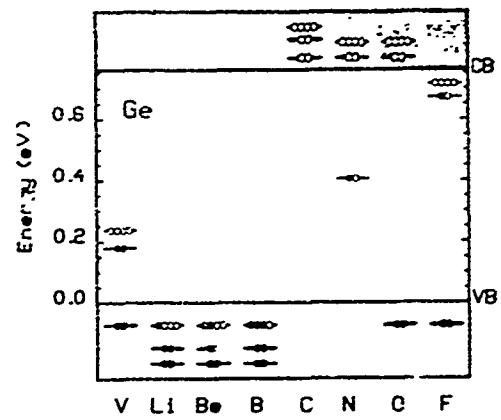


FIG. 9. Predicted deep levels for substitutional impurities from row 2 of the Periodic Table in Ge. Impurities to the right of C, namely N, O, and F, are not donors (counter to intuition). N and F are traps, while O is inert. B and Be are acceptors, C is inert, and Li and the vacancy trap both electrons and holes. Levels in the bands are not to scale.

Similarly F has a false valence of -1 instead of -3, and also has a deep level in the gap of Ge. There are no false valences for impurities on the left side of the Periodic Table, because the filled s- and p-like states in the valence band move up in energy, and cross into the gap for the vacancy levels (Fig. 9).

III. GUNN EFFECT

Gunn oscillations[17,18] result when electrons can transfer from a high-mobility region of the Brillouin zone to a low-mobility region. The mobility is

$$\mu = |e|\tau/m^* ,$$

where e is the electron's charge, and m^* is the electronic

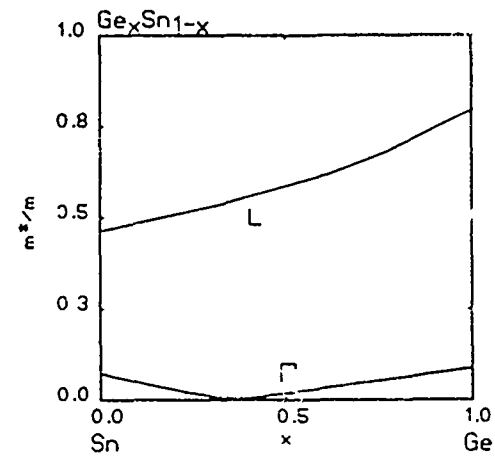


FIG. 10. Predicted effective electron masses in the Γ and L valleys vs composition x. The mass in the Γ valley is smaller than the mass of the L valley, likely resulting in a larger mobility for electrons in the Γ valley. The Gunn effect may be observed for $0.4 < x < 0.8$. The minimum in the mass of the Γ valley occurs at the composition where the energy band gap vanishes.

effective mass, and τ is the scattering time (due to phonon, impurity, and alloy scattering). In most semiconductors the mobility of electrons in the Γ valley of the conduction band is considerably higher than that in the L or X valleys, owing to the very light effective mass. We find this to be the case for Ge_xSn_{1-x} (see Fig. 7).[19] The effective masses produced by the current model may be in error by as much as a factor of 2; nevertheless, the model does give a good qualitative idea of how the masses vary with composition: the mass of the Γ minimum becomes very light near $x \simeq 0.4$, as the alloy becomes metallic (see Fig. 10).

Gunn devices are also known to produce coherent radiation.[20,21] As the potential across the device increases, it eventually causes transitions to the low-mobility state, and then the electrons slow down and form a high-resistivity domain that propagates along the device. Most of the potential drop is over the small domain. The resulting electric fields are large and can cause impact ionization, generating electron-hole pairs. As the domain passes through the material, the electron-hole pairs are left behind. The electrons (holes) fall to the conduction (valence) -band edge through phonon emission. The pairs undergo radiative recombination. Such radiation stimulates further recombination and light is coherently produced. It is an unanswered experimental question whether such effects occur in Ge_xSn_{1-x}.

IV. SUMMARY

In summary, we have predicted the electronic structure of Ge_xSn_{1-x} alloys, and find that these materials should exhibit interesting properties for some ranges of composition x, including direct band gaps in the infrared and band structures compatible with the Gunn effect. We hope that this work will stimulate further attempts to produce electronic-grade Ge_xSn_{1-x} materials.

ACKNOWLEDGMENTS

We are grateful to L. Romano, K. Newman, and D. Vasquez for their many useful discussions, and we gratefully acknowledge the generous support of the U.S. Office of Naval Research (under Contract No. N00014-84-K-0352). One of us (J.D.D.) thanks the Department of Chemical Engineering and Materials Science of the University of Minnesota for their hospitality.

[1] *Handbook of Binary Metallic Systems (Structures and Properties) 2* (Israel program for Scientific Translations, Jerusalem, 1967).

[2] S. Oguz, W. Paul, T. F Deutch, B.-Y. Tsaur, and D V. Murphy, Appl. Phys. Lett. 42, 848 (1983).

[3] L Romano, J. E. Sungren, S. A. Barnett, and J. E Greene, J. Cryst Growth (to be published).

[4] Compositions with $x \simeq 0$ have not, to our knowledge, been attempted.

[5] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B 14, 556 (1976).

[6] J. C. Slater and G F Koster, Phys. Rev 94, 1498 1954).

[7] P Vogl, H. P Hjalmarson, and J D Dow, J. Phys. Chem. Solids 44, 353 (1983).

[8] A. Kobayashi, O F. Sankey, and J. D Dow, Phys Rev B 25, 6367 (1982). Note that in this paper the definitions for $g_4(k)$ and $g_5(k)$ are different from ours Our definitions are $g_4(k)=g_2(k)-g_1(k)$ and $g_5(k)=g_2(k)-g_1(k)$

[9] K. E. Newman and J D Dow, Phys. Rev B 30, 1929 (1984).

[10] Additional second-neighbor parameters are listed in Table I. There are 12 second-neighbor vectors related by symmetry. One vector is $d_2=a_L(\frac{1}{2},\frac{1}{2},0)$, where a_L is the lattice constant.

[11] Y Onodera and Y Toyozawa, J Phys. Soc. Jpn. 24, 341 (1968).

[12] H. P. Hjalmarson, P. Vogl, D J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980). See also W Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[13] W A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), p. 48ff.

[14] For a recent review of deep-level theories, see S. Pantelides, Rev. Mod. Phys. 50, 797 (1977).

[15] See Table II of Ref. 9.

[16] S. Lee, J. D. Dow, and O. F. Sankey, Phys Rev. B 31, 3910 (1985).

[17] B. E. Streetman, *Solid State Electronic Devices*, (Prentice-Hall, Englewood Cliffs, New Jersey, 1980), p 47.

[18] J. B. Gunn, Suppl. J. Phys Soc. Jpn. 21, 505 (1966).

[19] The calculated effective masses differ from experimental values (0.02 for Sn at point Γ and 0.08 for Ge at point L [see J. I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971), p. 410]), being 0 08 and 0 79 for Ge and Sn, respectively.

[20] J. I. Pankove, *Optical Processes in Semiconductors* (Dover, New York, 1971), p 247.

[21] See, for example, A. C. Chynoweth, W. L. Feldmann, and D. E McCumber, Suppl. J. Phys. Soc. Jpn. 21, 514 (1966)

Theory of surface-defect states and Schottky barrier heights: Application to InAs

Roland E. Allen and Terry J. Humphreys

Department of Physics, Texas A&M University, College Station, Texas 77843

John D. Dow

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

Otto F. Sankey

Department of Physics, Arizona State University, Tempe, Arizona 85287

(Received 31 January 1984; accepted 1 May 1984)

Theoretical predictions of electronic energy levels associated with *s*- and *p*-bonded substitutional point defects at (110) surfaces of InAs and other III–V semiconductors are presented and discussed. The specific defects considered for InAs are: anion and cation vacancies, the (native) antisite defects In_{As} and As_{In}, and 26 impurities. The predicted surface-defect deep levels are used to interpret Schottky barrier height data for (a) *n*- and *p*-(InAs) and (b) the alloys $Al_xGa_{1-x}As$, $GaAs_{1-x}P_x$, $In_{1-x}Ga_xP$, and $In_{1-x}Ga_xAs$. The rather complicated dependence of the Schottky barrier height ϕ_B on alloy composition x provides a nontrivial test of the theory (and competing theories). The following unified microscopic picture emerges from these and previous calculations: (1) For most III–V and group IV semiconductors, Fermi-level pinning by native defects can explain the observed Schottky barrier heights. (2) For GaAs, InP, and other III–V semiconductors interfaced with nonreactive metals, the Fermi-level pinning is normally due to antisite defects. (3) When InP is interfaced with a reactive metal, surface P vacancies are created which pin the Fermi level. (4) Impurities and defect complexes are sometimes implicated. (5) At Si/transition-metal-silicide interfaces, Si dangling bonds pin the Fermi level. (6) These defects at the semiconductor/metal interfaces are often "sheltered" or "encapsulated." That is, the states responsible for Fermi-level pinning are frequently "dangling-bond" states that dangle into a neighboring vacancy, void, or disordered region. The defects are partially surrounded by atoms that are out of resonance with the semiconductor host, causing the defect levels to be deep-level pinned and to have energies that are almost independent of the metal.

PACS numbers: 73.20.Hb, 73.30.+y

I. INTRODUCTION

In this paper we report calculations of the deep levels associated with surface *s*- and *p*-bonded substitutional point defects in InAs, and we show how these results fit into an emerging unified microscopic picture of surface deep levels and Schottky barrier heights in III–V and group IV semiconductors.

II. DEFECTS AT THE InAs (110) SURFACE

The calculations we report employ existing and well-established techniques for treating defects in semiconductors and at semiconductor surfaces.[1–16] Briefly stated, the Green's function matrix $G(E) = (E - H)^{-1}$ of the host semiconductor, with a (110) surface relaxed according to the 27° rigid-rotation model,[17–19] is constructed in an empirical nearest-neighbor sp^3s^* tight-binding basis.[8,9] Because the matrix elements of the empirical tight-binding theory exhibit simple dependences on the orbital energies of the atoms and the bond lengths between neighbors, the changes in these matrix elements due to a defect can be estimated from the known atomic orbital energies of the defect atom and the (presumed known) lattice distortion around the defect. Thus, a defect potential matrix V_d can be constructed which is localized to the impurity site and a small number of its neighbors. (In practice lattice distortion around the defect has only a small, $\simeq 0.2$ eV, effect on the deep levels of interest, and so we neglect it—making V_d a diagonal matrix in a localized basis.) The resulting eigenvalue equation for the "deep" energy level E is

$$\det[1 - G(E)V_d] = 0,$$

whose solutions $E(V_s)$ are given in Figs. 1 and 2 for defects at the (110) surface of InAs. [To plot E as a function of a single variable V_s, we have made the usual approximations for the on-site matrix elements of the 5×5 matrix V_d: (1) the s^* diagonal element and all off-diagonal elements vanish, and (2) the three p diagonal elements are equal to one another and half of the s-diagonal element V_s.[9,20]] Details of the calculations, which are identical to those for other III–V semiconductors, are available in the literature.[1–5,12] It should be emphasized that this theory is best suited for predicting chemical *trends* in the energy levels of different impurities in different hosts, rather than predicting with precision the ab-

 © 1984 American Vacuum Society

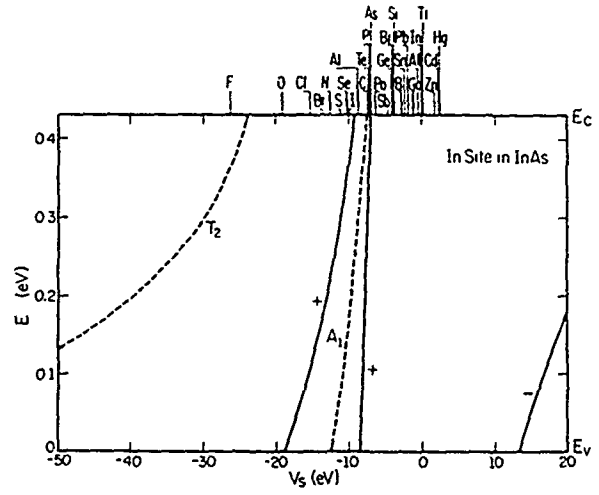


FIG. 1. Deep levels for impurities on the In site within an InAs(110) surface layer (solid lines). V_s is the impurity potential for s electrons. E_c and E_v are the conduction and valence band edges. The curves are labeled according to the parity of the defect state with respect to the $(1\bar{1}0)$ plane, which is perpendicular to the (110) surface plane. The dashed lines give the A_1 (s-like) and T_2 (p-like) levels for the In site in bulk InAs.

solute energy levels themselves. For this reason, it is especially well suited for treating small band-gap semiconductors, because a knowledge of the trends can compensate for a significant theoretical uncertainty on the scale of the band gap.

The predicted surface (solid line) and bulk (dashed line) deep levels for defects at the As site in InAs are given in Fig. 1. The surface or bulk deep levels of a specific impurity are obtained (roughly[21]) by dropping a vertical line from that impurity (at the top of the graph) and determining its intersections with the solid or dashed theoretical curves. For example, Zn on the As site in InAs is predicted to produce a bulk level roughly 0.2 eV above the valence band maximum and a surface + parity level at roughly 0.4 eV (just below the conduction band minimum). The vacancy levels are the limits as $V_s \to \infty$ of these curves.[15]

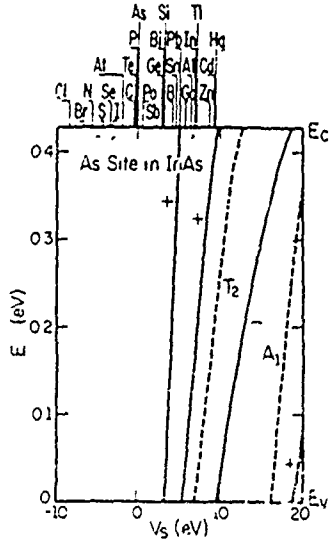


FIG. 2. Deep levels for the As site at the InAs(110) surface (solid lines) and in bulk InAs (dashed lines).

The electronic occupancy of a level is determined by noting its origin for $V_s = 0$ and the valence difference between the impurity and the host atom. For example, neutral Zn has three fewer electrons than As, or three more holes. For surface Zn, two of these holes go into the uppermost + level, which (with V_s increasing from zero) emerges from the valence band, crosses the gap, and passes into the conduction band. The third hole goes into the second + level, which (for V_s equal to the value appropriate for Zn) lies roughly 0.4 eV above the valence band maximum in the fundamental gap. Thus, the only deep level in the gap of neutral surface-Zn_{As} is occupied by one electron and one hole. Negatively charged Zn^- at the surface would have two electrons in this level. A similar argument for the bulk Zn level reveals that the orbitally threefold-degenerate T_2 state lies in the gap and is occupied by three electrons and three holes when the defect is neutral.

Similarly, the As vacancy corresponds to $V_s \to \infty$, with five holes added; i.e., the As site atom is decoupled from the host[15] and its electrons are removed. Thus, three electrons and five holes "occupy" all the levels of the vacancy, including those that have been pushed into and are resonant with the conduction band. Electrons are unstable in resonant levels, so they spill out and fall to the conduction band edge. The defect then becomes a shallow donor, with levels bound by the long-range Coulomb tail of the potential (omitted from the present model). Our calculations thus predict that the As vacancy in InAs yields only shallow donor levels in the gap.

We can estimate the energies and electronic occupancies of levels in the gap produced by the various defects of Figs. 1 and 2 in a similar way. We found above that the surface As vacancy produces no deep levels in the gap (deep being defined as bound as a result of the central-cell potential) and instead gives shallow donor states near the conduction band edge. Similarly the neutral surface In vacancy produces only a deep level occupied by one electron and one hole near the conduction band edge, at approximately[21] 0.3 eV. The surface antisite defect In_{As} produces a deep level in the gap at approximately[21] 0.05 eV, near the valence band maximum, which, for neutral In, is occupied by two electrons. Surface and bulk As_{In} produce doubly occupied donor states near the conduction band edge, that, within the theoretical uncertainty,[5] could be either shallow or deep (at $\simeq$0.65 eV, above the conduction band edge, for the surface, producing shallow donor levels in the gap).

III. SCHOTTKY BARRIER HEIGHTS

The energy levels of the native defects are relevant to the Schottky barrier problem. According to Bardeen's theory of Fermi-level pinning (in simplified form) the bulk bands bend so that the Fermi energies of the surface and the bulk align. The band bending forms a Schottky barrier. For n-type (p-type) InAs, the surface Fermi level is believed to be determined by native defects at the surface, and lies near the level to be occupied by an extra electron (hole).

For n-InAs the Schottky barrier height is approximately the energy of the conduction band edge relative to the lowest empty surface defect level—the surface Fermi level. Assum-

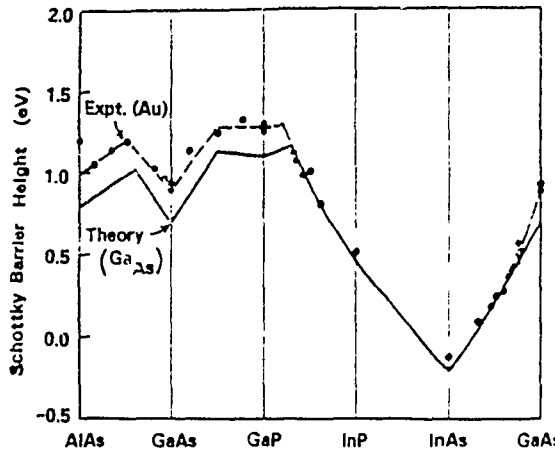


FIG 3 Schottky barrier height ϕ_B as function of alloy composition x for Au contacts to n-type $Al_xGa_{1-x}As$, $GaAs_{1-x}P_x$, $In_{1-x}Ga_xP$, and $In_{1-x}Ga_xAs$. The experimental data are those of J. S. Best, Appl. Phys. Lett. 34, 552 (1979) for $Al_xGa_{1-x}As$, W. G. Spitzer and C. A. Mead, Phys. Rev. A 133, 872 (1964), and D. A. Neamen and W. W. Grannemann, Solid State Electron. 14, 1319 (1971) for $GaAs_{1-x}P_x$, T. F. Kuech and J. O. McCaldin, J. Vac. Sci. Technol. 17, 891 (1980) for $In_{1-x}Ga_xP$, and K. Kajiyama, Y. Mizushima, and S. Sakata, Appl. Phys. Lett. 23, 458 (1973) for $In_{1-x}Ga_xAs$. For this alloy, the data of H. H. Wieder, Appl. Phys. Lett. 38, 170 (1981), for metal–insulator–semiconductor structures, are also shown.

ing only native surface defects (i) In_{As}, (ii) As_{In}, (iii) V_{As} (As vacancy), or (iv) V_{In}, the Fermi level is (i) near the conduction band edge, (ii) near the conduction band edge, (iii) near the conduction band edge, or (iv) at $\simeq 0.3$ eV. If the concentration of V_{In} at the surface is less than the concentration of As vacancies plus half the concentration of As_{In}, the As_{In} and V_{As} levels compensate the V_{In} level—so that the surface Fermi level lies at the conduction band edge. Thus, normally, the surface Fermi level of n-InAs is nearly at the conduction band edge, the Schottky barrier height is approximately zero, and the semiconductor/metal contact is, by definition, Ohmic.

For p-InAs, the Schottky barrier height is approximately the energy of the surface Fermi level for holes relative to the valence band maximum. The surface Fermi level for holes is (i) near the valence band maximum ($\simeq 0.05$ eV), (ii) near the conduction band edge, (iii) near the conduction band edge, or (iv) near $\simeq 0.3$ eV (slightly less than the 0.42 eV band gap) for In_{As}, As_{In}, V_{As}, or V_{In}, respectively. Under normal circumstances we expect the As_{In} and V_{As} defects to compensate V_{In}, causing the surface Fermi level to lie nearly at the conduction band edge. Hence, p-InAs should have a Schottky barrier height of approximately the band gap.

Mead and Spitzer[22] have reported a Schottky barrier height of zero (Ohmic) for n-InAs with Au contacts and a barrier height for p-InAs of approximately the band gap—as the theory predicts!

IV. RELATIONSHIP TO SCHOTTKY BARRIER HEIGHTS IN OTHER III–V MATERIALS

We believe the defect primarily responsible for Schottky barrier formation in InAs, is In_{As}, but that As_{In} and V_{As} play secondary roles, as discussed above. Evidence supporting the importance of the cation-on-anion site antisite defect is available from the predicted alloy dependence of Schottky barrier heights for Au on n-type semiconductors, shown in Fig. 3. (See also Refs. 1 and 2.) The theory rather dramatically mimics the data and their major and complicated chemical trends for a wide range of III–V semiconducting alloys. (For p-type materials, the predicted alloy dependences are less dramatic and quite different—and appear to be fully consistent with existing data.)

Further evidence supporting the role of surface antisite defects includes (i) experiments showing that the barrier can be "annealed" at a temperature characteristic of antisite defects,[23] but not at the vacancy annealing temperature,[24] and (ii) experiments demonstrating that the concentrations of defects responsible for pinning the surface Fermi levels of n-GaAs and p-GaAs are nearly equal[25] (see Ref. 30).

This is not to say that any one defect is responsible for Fermi-level pinning in all materials. For example, in n-InP[26–29] it appears likely[3] that reactive metal contacts produce P vacancies and make the Schottky barrier height nearly zero, whereas nonreactive metals produce primarily antisites and Schottky barrier heights of $\simeq 0.5$ eV. Extrinsic impurities, notably S and Sn, are also thought to determine the Schottky barrier height of InP under certain conditions.[3]

We speculate that the "cleavage-related defect" in GaAs[24,30] involves Ga_{As}, and that the "chemisorption-related defect" involves As_{Ga}.

V. RELATIONSHIP TO Si

A similar picture, namely, Fermi-level pinning by Si dangling bonds at Si/transition-metal–silicide interfaces, seems to explain those Schottky barriers as well.[6,31] This work also helps to explain why calculations for *free surfaces can be used to interpret data for semiconductor/metal contacts*. In reality, the defects responsible for Fermi-level pinning are not isolated point defects at an ideal free surface, but instead defects adjacent to vacancies, voids, or disordered regions of the semiconductor/metal interface, which are effectively "encapsulated" by electropositive metal atoms or other atoms at the contact. These atoms are out of resonance with some of the semiconductor atoms, and function primarily as a repulsive potential that pushes the defect state (frequently a dangling-bond state) back toward the semiconductor. Since they are out of resonance, they act electronically as pseudovacancies, mechanically "sheltering" the Fermi-level pinning defect without greatly affecting it electronically.[30] The defect level is "deep-level pinned"[9] by these encapsulants—explaining why different contacts produce similar Schottky barrier heights. The paradigm for such "sheltering" is an interfacial vacancy sheltering a Si dangling bond at a $Si(111)/NiSi_2(111)$ interface, illustrated in Refs. 6 and 31.

Here it is important to realize that *bulk* point defects, specifically bulk antisite defects, cannot explain the Fermi-level pinning data for GaAs, suggestions to the contrary[32] not withstanding. While, for example, bulk As_{Ga} appears to have a level close to the level required to explain the n-GaAs Schottky barrier, this level cannot pin the Fermi energy because (in contrast to the pinning level of surface As_{Ga}) it is a *donor*—being fully occupied when the defect is neutral. A deep *acceptor* is required to pin the Fermi level of an n-type semiconductor.[33] Moreover, the fact that As_{Ga} in the bulk has a certain energy level does *not* imply that it will have the same energy level at the surface or at a GaAs/metal inter-

face. Surface and interfacial defects generally have more deep levels in the fundamental band gap than the same defects in the bulk, and the energy levels of surface defects typically differ from those of bulk defects by ~1 eV.[14] This is the case because the p-like T_2 level of a bulk substitutional defect often lies outside of the band gap as a deep resonance in the conduction band; the surface or interface destroys the tetrahedral point-group symmetry of the defect and splits the T_2 level, often driving one or more of the split sublevels into the band gap.

We therefore conclude that the following zeroth-order model provides a very satisfactory description of Fermi-level pinning and Schottky barrier formation for both III–V and group IV semiconductors: Various native surface defects (antisites, vacancies, and dangling bonds), or their complexes, pin the Fermi energy at the semiconductor surface.[37] The basic picture initiated by Bardeen,[33] and further developed and championed by Spicer,[34–36] is thus found to be fundamentally correct. The task of future theories is to work out the corrections to this simple picture due to, for example, work-function effects and the details of specific metal/semiconductor interfaces.

ACKNOWLEDGMENTS

We gratefully thank the U. S. Office of Naval Research (Contract Nos. N00014-82-K-0447 and N00014-77-C-0537) for their generous support of this work. O. F. S was supported by the Arizona Faculty Grant-in-Aid program.

[1]R E Allen and J D. Dow, Phys. Rev. B 25, 1423 (1982).
[2]R. E. Allen and J D. Dow, J. Vac Sci Technol. 19, 383 (1981).
[3]J D Dow and R. E. Allen, J Vac. Sci. Technol 20, 659 (1982)
[4]R. E. Allen, J. D. Dow, and H. P. Hjalmarson, Solid State Commun. 41, 419 (1982)
[5]R. E Allen and J. D. Dow, Appl. Surf. Sci. 11/12, 362 (1982).
[6]O. F. Sankey, R. E. Allen, and J. D Dow, Solid State Commun. 49, 1 (1984).
[7]R. E. Allen, Phys. Rev B 20, 1454 (1979).
[8]P Vogl, H P Hjalmarson, and J D. Dow, J Phys. Chem Solids 44, 365 (1983).
[9]H P Hjalmarson, P Vogl, D J Wolford, and J. D Dow, Phys Rev Lett 44, 810 (1980)
[10]O F Sankey, H. P Hjalmarson, J D Dow, D J Wolford, and B G Streetman, Phys Rev Lett 45, 1656 (1980)
[11]H P Hjalmarson, Ph D thesis, University of Illinois, 1979
[12]R P Beres, Ph D thesis, Texas A&M University, 1982 The computer program used for the present calculations is reproduced as an appendix in this thesis, since it was also used in the intrinsic surface state calculations of Beres *et al.* [See, e g., R. P Beres, R E Allen, and J D Dow, Solid State Commun 45, 13 (1983)] The comment cards in this program provide a complete description of the method of calculation
[13]J. Pollmann, Adv Solid State Phys 20, 117 (1980)
[14]M S Daw and D L. Smith, Phys Rev B 20, 5150 (1979)
[15]M. Lannoo and P Lenglart, J Phys Chem Solids 30, 2409 (1969)
[16]The self-consistent Green's function technique for bulk defects was developed by G. A. Baraff and M. Schluter, Phys Rev Lett 41, 892 (1978) and by J Bernholc, N Lipari, and S T Pantelides, *ibid.* 41, 895 (1978).
[17]A Kahn, E So, P Mark, and C B Duke, J Vac. Sci Technol 15, 580 (1978).
[18]S Y Tong, A. R Lubinsky, B J Mrstik, and M A Van Hove, Phys. Rev B 17, 3303 (1978)
[19]D J Chadi, J Vac. Sci. Technol 15, 631 (1978)
[20]H P Hjalmarson, R E. Allen, H Buttner, and J D Dow, J Vac Sci. Technol 17, 993 (1980)
[21]See Ref 4 or 5 for the precise prescription which we have used to determine the native defect levels quoted in this paper
[22]C. A Mead and W. G Spitzer, Phys Rev A 134, 713 (1964)
[23]W Monch (to be published)
[24]W Monch, Surf. Sci 132, 92 (1983). This paper, and Refs 27 and 28, also appear in *Surfaces and Interfaces: Physics and Electronics*, edited by R. S Bauer (North-Holland, Amsterdam, 1983)
[25]W Monch and H Gant, Phys Rev Lett 48, 512 (1982)
[26]R H Williams, V Montgomery, and R R Varma, J Phys C 11, 2735 (1978).
[27]R H. Williams, Surf Sci 132, 122 (1983)
[28]H H. Wieder, Surf Sci. 132, 390 (1983)
[29]L J Brillson, Surf Sci Rep 2, 123 (1982
[30]R E Allen, R. P Beres, and J D Dow, J Vac Sci Technol B 1, 401 (1983)
[31]O F Sankey, R E Allen, and D Dow, J Vac Sci Technol B (these proceedings)
[32]J Schneider, 2nd Conference on Semi-insulating III–V Materials, Evian, 1982
[33]J. Bardeen, Phys. Rev 71, 717 (1947)
[34]W E Spicer, P W Chye, P R Skeath, C. Y Su, and I Lindau, J Vac Sci. Technol 16, 1422 (1979)
[35]W E. Spicer, I Lindau, P R Skeath, and C Y Su, J Vac Sci Technol 17, 1019 (1980)
[36]W E Spicer, I Lindau, P R Skeath, C Y Su, and P W Chye, Phys Rev Lett. 44, 420 (1980)
[37]As discussed in the text and in Ref. 4, impurities and subsurface defects appear to be involved sometimes

Ultramicroscopy 14 (1984) 127–130
North-Holland, Amsterdam

THEORY OF Si/TRANSITION-METAL SILICIDE SCHOTTKY BARRIERS

Otto F. SANKEY

Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

Roland E. ALLEN

Department of Physics, Texas A&M University, College Station, Texas 77843, USA

and

John D. DOW

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA

Received 20 January 1984; presented at Workshop January 1984

The height of the Schottky barrier formed at transition-metal/Si interfaces varies over a very small range (~ 0.4 eV) considering the wide range of electronic structures possible from one end of the transition-metal series to the other. Furthermore, the barriers are observed to form within a few monolayers of metal coverage, suggesting that the barrier is a property of the local bonding and that the true metallic states play only a minor role. A model has been developed to explain these facts in terms of the Fermi-level pinning mechanism of Schottky barrier formation. The physics contained in the model is that of a Si dangling bond sheltered from the transition-metal-silicide by an interfacial vacancy. Since (i) the dangling-bond is sheltered from the metallic-silicide and (ii) the atomic energy levels of the transition metal are out of resonance with Si, the dangling bond (which forms a level in the Si band gap) will be only weakly perturbed by the silicide. Thus this interfacial dangling bond can pin the Fermi level at nearly the same energy for all the transition-metal-silicides. A tight-binding calculation of the electronic structure of this defect at the $NiSi_2/Si(111)$ interface has been performed for an infinite interface using the transfer-matrix technique. The results of this calculation are described in terms of a very simple molecular model.

It is a remarkable fact that the Schottky barrier heights for the whole range of Si/silicide interfaces varies over a relatively narrow range of about 0.55–0.87 eV in n-Si [1]. For a Si bandgap of 1.1 eV, this places the Fermi-level in the *lower* part of the bandgap between 0.23–0.55 eV above the valence band edge. Here we argue that such barriers can be understood in terms of Fermi-level pinning [2] by a small concentration of Si dangling bonds that are "sheltered" from the transition metal by vacancies at the Si/silicide interface. This explanation, which differs substantially from previous theories of Si Schottky barrier formation [1,3,4], unifies the understanding of Si/transition-metal Schottky barriers with the generally accepted model of Fermi-level pinning by native defects [5,6] at (110) interfaces between III–V semiconductors and metals [7] or other overlayers. A more detailed account of our work will be given elsewhere [8].

The following observations place severe constraints on any theory of Schottky barrier formation at Si/silicide interfaces: (1) The barrier heights for the silicides all lie within 0.4 eV of one another for all the different transition metals, stoichiometries, and crystal structures. (2) The barriers are observed to form at low coverages before a complete metallic silicide is formed, indicating that the local atomic bonding at the interface, rather than any collective interface property, determines the barrier [1]. (3) There are only slight variations of the barrier height for different compounds of a

0304-3991/84/$03.00 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

given transition metal [9,10] indicating that the transition metal itself, rather than stoichiometry, crystal structure, etc., determines the barrier. (4) The barrier heights for n- and p-type Si very nearly add up to the Si bandgap. This means that the pinning level must only be partially occupied, so that it may act as both an acceptor and a donor.

We propose that these observations can be understood quite naturally in terms of an interfacial vacancy which shelters a Si dangling bond from the effects of the transition metal. This dangling bond has only a weak link with the silicide and is only slightly perturbed by the transition metal s-, p-, and d-orbitals, and hence is insensitive to the large variations (on a 1eV scale) one might expect to occur when the transition metal is varied or when the stoichiometry or the crystal structure of the silicide is changed.

To make these ideas specific, we consider a particular example of such a defect - the example illustrated in fig. 1 for the case of the abrupt $Si/NiSi_2(111)$ interface. If the vacancy in fig. 1 were replaced by a Si atom, one would have the bonding configuration determined by Cherns et al. [11]. For the reactive systems under consideration here (transition metals "eating" their way into Si), a reasonable concentration of vacancies ($\sim 10^{13}$ cm^{-2}) appears quite likely.

Here we consider a very simple model of the electronic structure of the defect shown in fig. 1. This model is justified only by the results of the more complete calculation described elsewhere [8], but it reveals the essential physics of the problem.

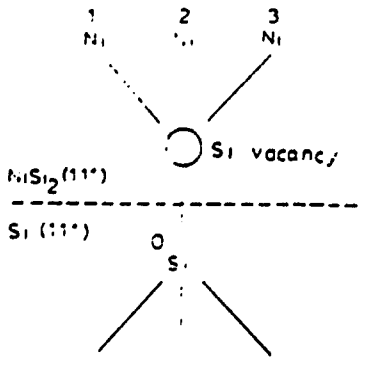


Fig. 1. An example of an interfacial vacancy sheltering a Si dangling bond. Replacing the vacancy by a Si atom gives the geometry of the $NiSi_2/Si$ (111) interface determined by Cherns et al. [11].

In this simplified model, only four atoms are explicitly considered - those surrounding the vacancy - and only one sp^3 hybrid orbital per atom - which is directed toward the vacancy. We first take all four atoms to be Si (tetrahedral symmetry) and later change three of these atoms into Ni (C_{3v} symmetry) to simulate the $Si/NiSi_2$ interface.

Taking all four atoms to be Si, we construct the A_1 and T_2 states of the bulk Si vacancy:

$$|A_1(a_1)\rangle = \tfrac{1}{2}(|\phi_0\rangle + |\phi_1\rangle + |\phi_2\rangle + |\phi_3\rangle). \quad (1a)$$

$$|T_2(a_1)\rangle = (1/\sqrt{12})(3|\phi_0\rangle - |\phi_1\rangle - |\phi_2\rangle - |\phi_3\rangle). \quad (1b)$$

$$|T_2(e),1\rangle = (1/\sqrt{2})(|\phi_1\rangle - |\phi_3\rangle). \quad (1c)$$

$$|T_2(e),2\rangle = (1/\sqrt{6})(|\phi_1\rangle + |\phi_3\rangle - 2|\phi_2\rangle). \quad (1d)$$

where the orbital $|\phi_i\rangle$ is the hybrid orbital of atom i. The energies of the A_1 and T_2 levels can be described by two parameters ϵ_h and t: here we have $\epsilon_h = \langle\phi_i|H|\phi_i\rangle$ is the orbital energy of an sp^3 hybrid, and $-t = \langle\phi_i|H|\phi_j\rangle$ for $i \neq j$ represents the interaction between two different hybrid orbitals. These two parameters represent effective interactions and are obtained by fitting to the bulk Si vacancy deep levels. The A_1 level is resonant with the valence band at $E_{A_1} = \epsilon_h - 3t$, while the triply degenerate T_2 level lies in the Si bandgap and has an energy $E_{T_2} = \epsilon_h + t$. We list in table 1

Table 1
The calculated A_1- and T_2-symmetric energy levels for the unrelaxed Si vacancy by several workers; the calculations are either pseudopotential (P) or tight-binding (TB); all energies are in eV, and the top of the valence band is defined to be the zero of energy; the Si bandgap is 1.1eV; the two parameters ϵ_h and t are simply obtained from the A_1 and T_2 energy levels (see text); the important parameter ϵ_h is the energy of a single dangling bond and is found to lie in the lower part of the Si bandgap in all cases

Type of calculation		A_1 level	T_2 level	t	ϵ_h
P	TB				
Ref. [12]		-1.10	0.70	0.45	0.25
Ref. [13]		-0.60	0.80	0.35	0.45
Ref. [14]		-1.10	0.60	0.42	0.16
	Ref. [15]	-0.55	0.75	0.33	0.43
	Ref. [16]	-0.96	0.51	0.37	0.14

the A_1 and T_2 levels for three different pseudopotential calculations [12–14] and two tight-binding [15,16] calculations for the unrelaxed vacancy in Si. From these levels the parameters ϵ_h and t can be extracted using $\epsilon_h = (E_{A_1} + 3E_{T_2})/4$ and $t = (E_{T_2} - E_{A_1})/4$. Note that ϵ_h, which is the energy of a *single* Si dangling bond, lies in the lower part of the Si bandgap in all cases.

We next change three of the atoms surrounding the vacancy into Ni atoms. We do this by raising the hybrid orbital energies of atoms 1, 2, and 3 (see fig. 1) from ϵ_h to $\epsilon_h + V$, where $V \approx 5$ eV [8] represents the (large) positive difference between a Ni and a Si sp^3 hybrid orbital. The symmetry is now reduced from T_d to C_{3v}, and the possible levels are of a_1 (σ-like) and e (π-like) symmetry. The states of e-symmetry evolve from two of the T_2 levels of the bulk Si vacancy (the T_2(e) levels in eqs. (1c) and (1d)), but are raised out of the gap roughly linearly with the potential V to become resonant with the conduction bands. Since the e-symmetric levels are not in the gap and are metal-atom derived, they play no role in pinning the interfacial Fermi level, and we will no longer consider them.

The interesting levels are those of a_1-symmetry which are admixtures of the $|A_1(a_1)\rangle$ (eq. (1a)) and the $|T_2(a_1)\rangle$ (eq. (1b)) levels of the bulk Si vacancy. However, since the Si and Ni hybrid orbitals are no longer degenerate, perturbation theory shows that the effective interaction between Si and Ni hybrid orbitals is reduced from t (~ 0.4 eV) for the bulk Si vacancy to t^2/V (~ 0.03 eV) for the interfacial vacancy. A schematic energy level diagram for a_1-symmetric states of the bulk and interfacial vacancies is shown in fig. 2. Note that because Ni (or any transition-metal element) and Si are "out of resonance", a level is formed in the lower part of the Si bandgap which is tied to the Si dangling bond energy ϵ_h and is relatively insensitive to the transition metal as long as we have $V \gg t$. This simple model leads to the important conclusion that for various transition metals, interfacial Fermi-level pinning positions are nearly equal to, but slightly below, the "defect pinning energy" ϵ_h of a single Si dangling bond.

We briefly mention the more rigorous calculations on which the simple model is based. A

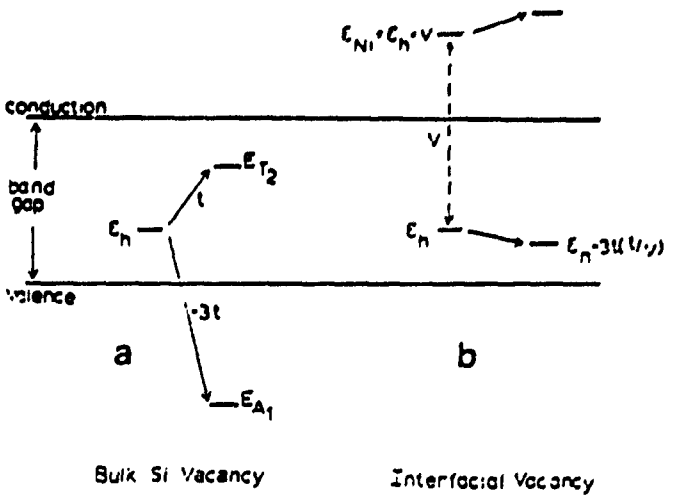


Fig. 2. Schematic energy level diagram of the a_1-symmetric levels of (a) the bulk Si vacancy and (b) the interfacial vacancy. In (a) the hybrid orbitals at ϵ_h lie in the lower part of the Si bandgap but interact strongly through t to produce the A_1 level resonant with the valence band and a T_2 level in the upper part of the bandgap. In (b) the hybrid orbitals of Si and Ni are no longer degenerate and their interaction is reduced by t/V. This brings E_{A_1} out of the valence band so that it now lies only slightly below the Si dangling bond energy ϵ_h. (The a_1-symmetric representation refers to the C_{3v} group appropriate for the interfacial vacancy. Since C_{3v} is a subgroup of T_d, the A_1 and one of the T_2 levels of the bulk Si vacancy are also a_1-symmetric.)

tight-binding calculation was performed for an embedded cluster of a vacancy and three Ni atoms (including d-orbitals on Ni) in an infinite Si host [8] using the Si tight-binding model of ref. [15]. The Si dangling-bond-like level is found at 0.4 eV. The d-orbitals are found to play only a minor role. Since the d-orbital energies lie well below the Fermi level, they tend to push up slightly on the Si dangling bond, but with a greatly reduced strength because the d-orbital is not a nearest neighbor to the dangling bond orbital and hence interacts with it either through a small second-neighbor interaction or indirectly via its interactions with the intervening Si atoms surrounding the vacancy. (The d-orbitals were taken to interact only with nearest neighbors.) More sophisticated calculations for an interface between semi-infinite slabs of $NiSi_2$ and Si have recently been completed [17] using the transfer-matrix technique [18]. The tight-binding bands of $NiSi_2$ have been fit to the bulk bands of Chabal et al. [19], and the tight-binding model of

Vogl et al. [20] has been used for bulk Si. Here the interfacial vacancy level is found to lie at 0.13 eV. Although the two calculations give slightly different results, and the estimates made for the dangling bond energy ϵ_h in table 1 differ by ~ 0.3 eV, they all show that the defect "pinning" level lies in the lower part of the Si bandgap. Measurements for a Si dangling bond quite similar to the one described here [21] at the SiO_2/Si interface show a level at 0.36 eV [22]. As mentioned earlier, the interfacial Fermi level for the silicides lie approximately in the range 0.23–0.55 eV.

Conclusion. The present theory is manifestly based on local atomic bonding and a localized defect, and is thus compatible with the experimental findings [1,9,10] that the observed Schottky barriers form before the completion of a complete metallic overlayer. Since in this model the barriers are determined mainly by Si, the barrier is affected to a lesser degree by the nature of the transition metal atom, stoichiometry, or crystal structure of the silicide. Furthermore, since the dangling bond is occupied by a single electron, it can act either as a donor or an acceptor – this leads to very nearly the same pinning position for both n- and p-Si, in agreement with the measurements. This is to be contrasted with Schottky barrier formation on III–V semiconductors, such as GaAs, where previous theoretical studies indicate that pinning is often due to surface antisite defect levels [6] which lead to different Fermi-level pinning positions for n- and p-type semiconductors.

O.F.S. wishes to thank the Arizona State Faculty Grants in Aid program for their support. We are grateful to the US Army Research Office (ARO-DAAG29-83-K-0122) and to the US Office of Naval Research (N00014-82-K-0447 and N00014-77-C-0537) for supporting this research.

References

[1] See, for example, P.S. Ho and K.N. Tu, Eds., Thin Films and Interfaces (North-Holland, New York, 1982). Reprinted from Thin Solid Films 93 (1983).
[2] J. Bardeen, Phys. Rev. 71 (1947) 717.
[3] J.M. Andrews and J.C. Phillips, Phys. Rev. Letters 35 (1975) 56.
[4] J.L. Freeouf, Solid State Commun. 30 (1980) 1059.
[5] W.E. Spicer, I. Lindau, P.R. Skeath, C.Y. Su and P.W. Chye, Phys. Rev. Letters 44 (1980) 420;
W. Monch and H. Gant, Phys. Rev. Letters 48 (1982) 512;
H.H. Wieder, Appl. Phys. Letters 38 (1980) 170.
R.H. Williams, V. Montgomery and R.R. Varma, J. Phys. C11 (1978) L735;
R.W. Grant, J.R. Waldrop, S.P. Kowadzyk and E.A. Kraut, J. Vacuum Sci. Technol. 19 (1981) 477.
[6] R.E. Allen and J.D. Dow, Phys. Rev. B25 (1982) 1423;
J.D. Dow and R.E. Allen, J. Vacuum Sci. Technol. 20 (1982) 659.
[7] C.A. Mead and W.G. Spitzer, Phys. Rev. 134 (1964) A713.
[8] O.F. Sankey, R.E. Allen and J.D. Dow, Solid State Commun. 49 (1984) 1, and to be published.
[9] G. Ottaviani, K.N. Tu and J.W. Meyer, Phys. Rev. B24 (1981) 3354.
[10] P.S. Ho and G.W. Rubloff, Thin Solid Films 89 (1982) 433.
[11] D. Cherns, G.R. Anstis, J.L. Hutchinson and J.C.H. Spence, Phil. Mag. A46 (1982) 849.
[12] G.A. Baraff and M. Schluter, Phys. Rev. Letters 41 (1978) 892.
[13] J. Bernholc, N.O. Lipari and S.T. Pantelides, Phys. Rev. Letters 41 (1978) 895.
[14] U. Lindefelt and A. Zunger, Phys. Rev. B26 (1982) 846.
[15] D.A. Papaconstantopoulos and E.N. Economou, Phys. Rev. B22 (1980) 2903.
[16] H.P. Hjalmarson, P. Vogl, D.J. Wolford and J.D. Dow, Phys. Rev. Letters 44 (1980) 810.
[17] O.F. Sankey, to be published.
[18] D.H. Lee and J.D. Joannopoulos, Phys. Rev. B23 (1981) 4997.
[19] Y.T. Chabal, D.R. Hamann, J.E. Rowe and M. Schluter, Phys. Rev. B25 (1982) 7598.
[20] P. Vogl, H.P. Hjalmarson and J.D. Dow, J. Phys. Chem. Solids 44 (1983) 365.
[21] E.H. Poindexter, P.J. Caplan, B.E. Deal and R.R. Razouk, J. Appl. Phys. 52 (1981) 879.
[22] M.C. Chen and D.V. Lang, Phys. Rev. Letters 51 (1983) 427.

Effects of diagonal Anderson disorder on x-ray emission and photoemission spectra

Marshall A. Bowen
Department of Physics, Western Illinois University, Macomb, Illinois 61455

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 30 January 1984)

The effects of diagonal Anderson disorder and of final-state interactions on x-ray emission and photoemission spectra are evaluated in a change of mean-field model. The disorder produces distinct sideband features on the x-ray photoemission spectra, but primarily broadens the emission spectra.

I. INTRODUCTION

In an effort to assess the combined effects of disorder and many-electron processes on x-ray emission and photoemission spectra, we have numerically computed ensemble-averaged spectra for a one-dimensional Anderson model[1] of a conduction band in a metal. In this tight-binding model, a single energy level α_n (having a random value between $-D/2$ and $D/2$) is associated with each site n, and there is a constant nearest-neighbor coupling β between adjacent sites. The effect of the core hole on the conduction electrons is simulated by a final-state (initial-state) interaction V localized at the site of the core hole for photoemission (emission). There are a finite number of sites, and the many-body states included in the calculation are the low-energy particle-hole pair states created by the sudden switching on or off of V. The model is identical to one employed earlier for a discussion of absorption spectra[2] and is a change of mean-field model.

We find that the emission spectra, like the absorption spectra,[2] are primarily broadened inhomogeneously by the disorder, but that distinct spectral features arise in the x-ray photoemission spectra (XPS) associated with specific environments of the core hole.

II. MODEL

For x-ray photoemission, the initial-state many-electron Hamiltonian is

$$H_I = \sum_{i=1}^{N} h_i \ ,$$

where h_i is the Anderson one-body Hamiltonian:

$$h = \sum_n \alpha_n |n)(n| + \beta|n)(n+1| + \beta|n-1)(n| \ .$$

Here, β is a constant and α_n is randomly distributed between $-D/2$ and $D/2$. The final-state many-electron Hamiltonian is

$$H_F = \sum_{i=1}^{N} h_i' \ ,$$

where h_i' is identical to h_i, except for the tight-binding matrix element at the site of the core hole—which has an additional electron-hole interaction term $V < 0$. The initial state $|I\rangle$ is a Slater determinant of the N lowest-energy single-particle orbital eigenfunctions $|\phi)$ of h, and the various final states $|F\nu\rangle$ are determinants of N eigenfunctions $|\psi)$ of h'. The x-ray photoemission spectrum for photoelectrons of energy E is

$$I(E) = \sum_\nu |\langle I|F\nu\rangle|^2 \delta(E + E_{F\nu} - E_I - \hbar\omega - \epsilon_{\rm core}) \ ,$$

where $E_{F\nu}$ and E_I are the final- and initial-state energies of the conduction electron gas and $\epsilon_{\rm core}$ the core-hole energy relative to the center of the conduction band.

Similarly, the emission spectrum for photons of energy E is

$$\kappa(E) = \sum_\nu |\langle F\nu|M|I\rangle|^2 \delta(E - E_I + E_{F\nu}) \ ,$$

where M is the dipole operator. Here, we have the initial state $|I\rangle$ being a determinant of $N+1$ $|\psi)$'s, and the various final states $|F\nu\rangle$ corresponding to configurations described by determinants of N $|\phi)$'s and one core orbital.

$$(F\nu|M|I) = M_0 \begin{vmatrix} \psi_1(\vec{R}) & \psi_2(\vec{R}) & \cdots & \psi_{N+1}(\vec{R}) \\ (\phi_{\nu,1}|\psi_1) & & & \\ (\phi_{\nu,2}|\psi_1) & & & \\ \cdots & & & \\ (\phi_{\nu,N}|\psi_1) & \cdots & & (\phi_{\nu,N}|\psi_{N+1}) \end{vmatrix} \ .$$

Here, $(\phi|\psi)$ is a scalar product, $\vec{R}$ the core-hole site, and we have assumed that the core radius is negligible,[3] so that M_0 is a constant.

We have the sum rules[3]

$$\int_{-\infty}^{\infty} \kappa(E)dE = M_0^2 \sum_i |\psi_i(\vec{R})|^2 \ ,$$

where the sum is over occupied orbitals of the initial state, and[4]

$$\int_{-\infty}^{\infty} I(E)dE = 1 \ .$$

The line shapes we display are ensemble averages (denoted by $\langle\langle\ \rangle\rangle$) over all core-hole sites (typically 1000 such sites), for example,

$$\chi(E) = \left\langle\left\langle \sum_\nu |\langle F\nu|M|I\rangle|^2 M_0^{-2} \delta(E - E_I - E_{F\nu})\right\rangle\right\rangle$$

 ©1984 The American Physical Society

III. RESULTS

Here, we display spectra for $V=-1$, and with $\beta=-\frac{1}{2}$. The amount of disorder is specified by the parameter D (recall that the on-site energies α_n range from $-D/2$ to $+D/2$). The calculations have been executed repeatedly for a single spin channel[4] and for finite lattices with 40 sites and 20 electrons (a half-filled band). The core hole has been confined to one of the 10 innermost sites. We have broadened the calculated spectra by convolving them with a Gaussian,[5] for ease of presentation. Ensemble averages of the spectra over typically $\simeq 10$ core-hole sites for each of $\simeq 100$ different Anderson lattices have been performed.

The peaked bottom portion of the emission spectrum (Fig. 1) for $D=0$ is a band-structure effect and corresponds to the band-minimum van Hove singularity[6] with its $(E-E_0)^{-1/2}$ behavior above the band bottom. The high-energy peak occurs at the Fermi energy and is an x-ray edge singularity[7]—a multielectron recoil effect. With increasing disorder D, the van Hove singularity is broadened, blurred, and cut off, and the edge singularity is weakened (as in the case of the absorption spectra[2]). Basically, increasing disorder makes the spectrum smooth, broad, and relatively featureless.

The XPS spectrum for $D=0$ has two peaks: a large one for recoil involving configurations with an electron in the bound-exciton state of the electron-hole interaction and a smaller one for configurations involving only band states (Fig. 2). (The separation of these peaks is approximately the energy of the Fermi level relative to the bound-exciton level.) Small disorder blurs this spectrum, until in the large-disorder limit (see Fig. 2 for $D=5$) the site of the core hole is completely decoupled from other sites. In this limit there are three types of transitions (Fig. 3). (i) Zero-recoil transitions that occur if the Anderson level on the core-hole site lies above the Fermi energy by more than $|V|$. In this case, the electron gas cannot recoil because the site is isolated and the electron-hole interaction cannot pull the empty Anderson level below the Fermi energy. Hence, the XPS spectrum in this limit produces a peak at $E-\hbar\omega-\epsilon_{\text{core}}\simeq 0$. (ii) V-recoil transitions in which single-particle Anderson states on the core-hole are initially occupied and move down in energy by $|V|$, the strength of the electron-hole interaction. This produces a peak at $E-\hbar\omega-\epsilon_{\text{core}}\simeq|V|$. (iii) Intermediate transitions, corresponding to an empty level above the Fermi level, but within $|V|$ of it, being pulled down below the Fermi level by the electron-hole interaction. These transitions produce recoil of the electron gas at energy $\alpha-|V|$, where α is the diagonal Anderson energy on the core-hole site.

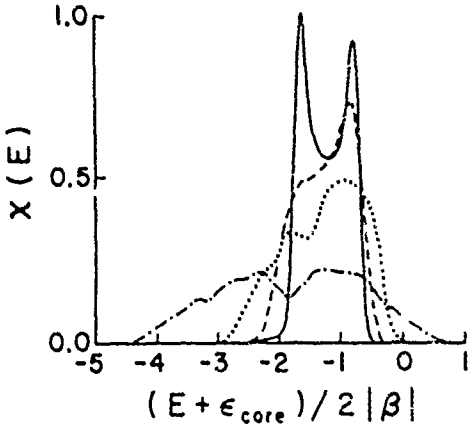


FIG. 1. X-ray emission spectra λ as a function of the emitted x-ray energy E for various disorder parameters D. The solid line is for $D=0$, the dashed line for $D=0.2$, the dotted line for $D=1$, and the chained line for $D=5$.

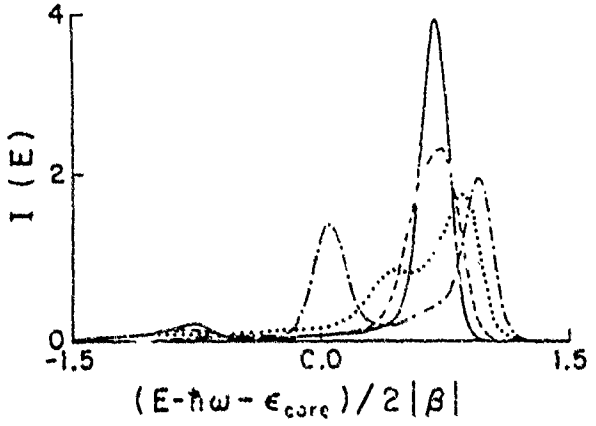


FIG. 2. X-ray photoemission spectra (XPS) as a function of the emitted electron energy E for various disorder parameters, as in Fig. 1.

The larger (smaller) peak for $D=1$ in Fig. 2 corresponds to configurations with an electron (no electron) occupying the core-hole-site one-electron level that lies below the Fermi level in the final state. Perhaps the most interesting case is the highly localized limit ($D\to\infty$), in which case the XPS spectrum provides information giving (i) the strength $|V|$ of the electron-hole interaction (which is the splitting between peaks), and an estimate of the number of states initially below and above the Fermi energy (the areas in the peaks at $E\simeq\hbar\omega+\epsilon_{\text{core}}-V$ and $\hbar\omega+\epsilon_{\text{core}}$, respectively).

Thus, we conclude that XPS spectra may be more sensitive to Anderson disorder than either absorption or emission spectra. Ordinary diagonal Anderson disorder corresponds to randomness in the on-site matrix elements α_n, and is not a realistic representation of the disorder occurring in amorphous metals. (Such disorder is better represented by randomness in β.) Perhaps the best physical realization of an Anderson alloy is a multicomponent crystalline alloy, with the number of components sufficiently large that a broad distribution of values for α_n is achieved. We hope that the present work will stimulate studies of such alloys.

The results of the present initial investigation are suffi-

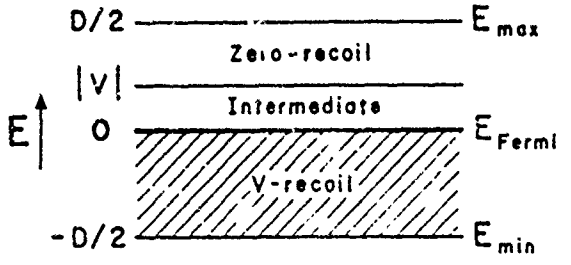


FIG. 3. Energy-level diagram for the conduction band, indicating the three regions—zero recoil, V recoil, and intermediate—appropriate to the large-disorder limit, $D\to\infty$. $E_{\min}$ and $E_{\max}$ are the band extrema and E_{Fermi} is the Fermi energy for the half-filled band. V is the electron-hole interaction.

ciently interesting that the authors have undertaken calculations of the combined effects of multielectron recoil and (i) off-diagonal Anderson disorder, and (ii) binary and ternary crystalline alloy disorder. The conclusions of those studies will be reported soon.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research for their generous support (Contract No. N00014-C-0537) and to S. Satpathy for useful discussions.

[1]P. W. Anderson, Phys. Rev. 109, 1492 (1958).
[2]S. Satpathy, J. D. Dow, and M. A. Bowen, Phys. Rev. B 28, 4255 (1983).
[3]C. A. Swarts, J. D. Dow, and C. P. Flynn, Phys. Rev. Lett. 43, 158 (1979); C. A. Swarts, and J. D. Dow (unpublished).
[4]J. D. Dow and C. P. Flynn, J. Phys. C 13, 1341 (1980).
[5]The broadening function is $(2\pi\Gamma^2)^{-1/2}\exp(-x^2/2\Gamma^2)$, with $\Gamma = 0.08|2\beta|$.
[6]L. van Hove, Phys. Rev. 89, 1189 (1953).
[7]J. Friedel, Comments Solid State Phys. 2, 21 (1969); P. Nozières and C. T. de Dominicis, Phys. Rev. 178, 1097 (1969).

J. Phys. Chem. Solids Vol. 46, No. 1, pp. 79–81, 1985
Printed in the U.S.A.

0022-3697/85 $3.00 + .00
© 1985 Pergamon Press Ltd.

OFF-DIAGONAL DISORDER AND X-RAY EDGES

MARSHALL A. BOWEN
Department of Physics, Western Illinois University, Macomb, IL 61455, U.S.A.

and

JOHN D. DOW
Department of Physics, University of Notre Dame, Notre Dame, IN 46556, U.S.A.

(*Received* 14 *May* 1984; *accepted* 21 *June* 1984)

Abstract—The combined effects of off-diagonal disorder and electron–hole pair production (the X-ray edge effect) are studied for a one-dimensional Anderson model of a liquid metal with a half-full band.

1. INTRODUCTION

In this paper we report model calculations of the X-ray absorption, emission, and photoemission spectra of a simple model for liquid metals. The model treats the many-body aspects of the X-ray transitions in a change-of-mean-field approximation [1, 2], while simulating the disorder of the liquid by a one-dimensional Anderson Hamiltonian with off-diagonal disorder [3]. We are interested in understanding the combined effects of multi-electron recoil and disorder on the spectra. A previous study [4], which considered only diagonal disorder, found dramatic effects on the X-ray lineshapes. However, this finding appears to be at odds with the experimental fact that X-ray spectra of metals change very little when the metals melt [5]. While diagonal disorder is more commonly treated theoretically, it is realized physically only in alloys made of many different types of atoms. The disorder found in a liquid metal is best simulated by an off-diagonal-disordered tight-binding model: the diagonal elements of the Hamiltonian are all the same (because the atoms are the same) but the off-diagonal matrix elements vary—because the bond-lengths vary [6].

2. THEORY

Here we consider the case of X-ray photoemission spectra (XPS). The treatment of absorption [4] and emission [7] is similar. The XPS lineshape is

$$I(E) = \sum_{\nu} |\langle I|F\nu\rangle|^2 \delta(E - \hbar\omega - \epsilon_{\text{core}} + E_{F\nu} - E_I) \quad (1)$$

where $\hbar\omega$ is the energy of the exciting photon, ϵ_{core} is the core energy level, and the sum is over all final states $|F\nu\rangle$ of the N-electron Fermi sea. In the initial state $|I\rangle$, the Fermi sea is quiescent and described by the tight-binding Hamiltonian

$$H_I = \sum_{i}^{N} h_i \quad (2)$$

where h is the Anderson single-particle Hamiltonian

$$h = \sum_{j=1}^{M-1} \beta_j [|j+1\rangle\langle j| + |j\rangle\langle j+1|] \quad (3)$$

and $|j\rangle$ is a localized single-particle state on the jth site. The $\{\beta_j\}$ are random numbers uniformly distributed in the interval $[-A, -B]$; M is the number of lattice sites in the system.

The final-state Hamiltonian is

$$H_F = \sum_{i=1}^{N} (h_i + v_i) \quad (4)$$

where v is the electron–hole interaction potential suddenly impressed upon the electron gas by the removal (to pseudo-infinite energy) of a core electron. We assume that v is localized at the lth site

$$v = V|l\rangle\langle l|. \quad (5)$$

The initial- and final-state Hamiltonians are both sums of one-electron Hamiltonians, and so the wavefunctions $|I\rangle$ and $|F\nu\rangle$ are Slater determinants of the single-particle eigenstates $|\phi\rangle$ and $|\psi\rangle$ of h and $h + v$, respectively. The initial-state energy of the electron gas is

$$E_I = \sum_{j=1}^{N} \epsilon_j \quad (6)$$

where the sum is over the N single-particle spin-orbitals j occupied in the initial state, and we have

$$h|\phi_j\rangle = \epsilon_j|\phi_j\rangle. \quad (7)$$

The final-state energies are

$$E_{F\nu} = \sum_{j=1}^{N} \epsilon_{\nu j} \quad (8)$$

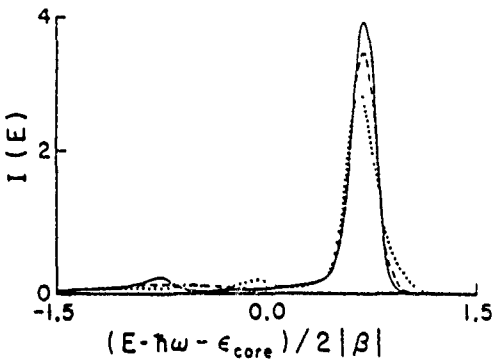


Fig. 1. The X-ray photoemission spectra as a function of the photoelectron's energy E. The solid line represents the calculated spectrum for $\beta = -1/2$, a constant. The range of β is $[-1/4, -3/4]$ for the dashed line and $[0, -1]$ for the dotted line. The photon energy is $\hbar\omega$ and the core electron energy is ϵ_{core}. The electron-hole interaction strength is $V = -1$.

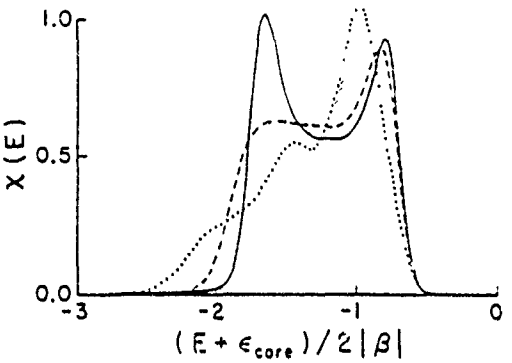


Fig. 2. The absorption lineshape as a function of the absorbed X-ray's energy E. The solid line is for β fixed at $-1/2$. The range of β is $[-3/4, -1/4]$ for the dashed line and $[-1, 0]$ for the dotted line. Note the X-ray edge singularity (for $\beta = -1/2$) at the left of the spectrum.

where we have

$$(h + v)|\psi_{\nu,l}\rangle = \epsilon'_{\nu,l}|\psi_{\nu,l}\rangle \tag{9}$$

and the νth final state is the configuration $\{\nu, 1; \nu, 2 \cdots; \nu, N\}$. Spin can be neglected, because the recoil profiles $I(E)$ for each spin channel are independent and produce the total recoil profile when convolved with one another [1]. The matrix element $\langle I|F\nu\rangle$ is

$$\langle I|F\nu\rangle = \begin{vmatrix} (\phi_1|\psi_{\nu,1}) & (\phi_1|\psi_{\nu,2}) \cdots (\phi_1|\psi_{\nu,N}) \\ (\phi_2|\psi_{\nu,1}) & (\phi_2|\psi_{\nu,2}) \cdots (\phi_2|\psi_{\nu,N}) \\ \vdots & \\ (\phi_N|\psi_{\nu,1}) & (\phi_N|\psi_{\nu,2}) \cdots (\phi_N|\psi_{\nu,N}) \end{vmatrix}. \tag{10}$$

The calculations of $I(E)$ proceed as follows: (i) using a random number generator, the set $\{\beta_1 \cdots \beta_{M-1}\}$ and the Hamiltonian h (eqn (3)) are generated for $M = 40$ sites; (ii) the eigenvalues ϵ and eigenvectors $|\phi\rangle$ are obtained by direct diagonalization of h; (iii) the core hole site l is specified and $h + v$ is diagonalized, generating eigenvalues ϵ' and states $|\psi\rangle$; (iv) various configurations $\{\nu, 1; \nu, 2; \ldots; \nu, N\}$ of N electrons are generated for the various final states, starting with the perturbed ground state (i.e. the ground state of the electron gas in the presence of the hole or v) and its low-lying excitations; (v) the excitation energies $E_\nu - E_l$ and matrix elements $\langle I|F\nu\rangle$ are calculated explicitly; (vi) the XPS spectrum $I(E)$ is then evaluated (for a finite system of electrons, it is a series of delta functions) and broadened to appear continuous†; (vii) this procedure is repeated for higher energy final states until the sum rule

$$\int_{-\infty}^{\infty} I(E)\, dE = 1 \tag{11}$$

is adequately exhausted; (viii) this procedure is repeated for new core hole sites l (restricted to one of the ten central sites of the lattice) and an average of $I(E)$ is performed; and (ix) the entire procedure is repeated for new random lattices $\{\beta_1 \cdots \beta_{M-1}\}$ and an ensemble average of $I(E)$ is taken. Convergence studies have shown that adequate convergence to identify all the major physical features of the spectra occurs for $N \simeq 20$ electrons, $M \simeq 40$ lattice sites, and $\simeq 100$ choices of $\{\beta_1 \cdots \beta_{M-1}\}$ [4]. A similar scheme is used for calculating absorption and emission spectral lineshapes $\chi(E)$.

† The broadening function is $(2\pi\Gamma^2)^{-1/2} \exp(-x^2/2\Gamma^2)$, with $\Gamma = 0.08$.

3. RESULTS

Our calculated results for the XPS, absorption, and emission lineshapes are shown in Figs. 1–3, respectively. Taking units such that the electron–hole interaction is $V = -1$, we consider, for each type of spectrum, three cases: (i) $\beta = -1/2$, which corresponds to an ordered one-dimensional lattice; (ii) β randomly distributed from $-1/4$ to $-3/4$ (moderate disorder), and (iii) β distributed in the interval $[-1, 0]$. This last case corresponds to extreme disorder, in that $\beta = 0$ corresponds to a completely broken bond—which, in one-dimension, completely disrupts the electronic wavefunction.

In studying the X-ray spectra, one should keep in mind the corresponding densities of one-electron

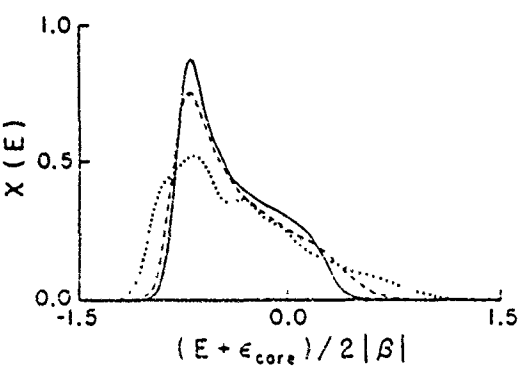


Fig. 3. The emission lineshape $\chi(E)$ as a function of the emitted X-ray's energy E. The solid line is for β fixed at $-1/2$. The range of β is $[-3/4, -1/4]$ for the dashed line and $[-1, 0]$ for the dotted line.

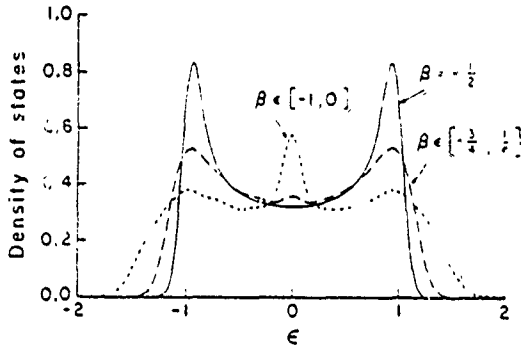


Fig. 4. Broadened† one-electron densities of states $D(\epsilon)$ vs energy ϵ for $\beta = -1/2$ (solid) and for the off-diagonally disordered, 10,000-atom, one-dimensional Anderson model with modest disorder $\beta\epsilon[-3/4, -1/4]$ (dashed) and extreme disorder $\beta\epsilon[-1, 0]$ (dotted). Any asymmetry with respect to $\epsilon = 0$ is associated with numerical noise. Note the appearance of a peak at $\epsilon = 0$, associated with delocalized states, for extreme disorder; all other states are localized.

states, Fig. 4, which exhibit first an amputation of the van Hove singularities [8] for moderate disorder ($\beta\epsilon[-3/4, -1/4]$), followed by a general flattening and the appearance (for $\beta\epsilon[-1, 0]$) of a peak at $\epsilon = 0$ corresponding to delocalized states at the band center. These delocalized states are important, especially in X-ray absorption and emission spectra, because they overlap the core-hole site, whereas most of the other states do not. To be sure, the overlap of a delocalized state with a core hole is not so large as that for a state localized on or near the core-hole site, but once thorough localization has set in for $\epsilon \neq 0$, the localized states significantly overlap only nearby sites—and much of the interesting physics involves only the core-hole site, its immediate neighbors, and the delocalized states.

The XPS spectrum is asymmetric [9] and blurred by disorder, which broadens the distribution of energies at which the electron gas can recoil in response to the shock associated with the sudden creation of the core hole. For zero disorder, the XPS spectrum exhibits two peaks: a large one corresponding to low-energy excitations of the conduction-electron Fermi sea in which the bound state due to the electron-hole interaction remains full, and a small one corresponding to an excited Fermi sea with no electron in the bound-state orbital.

With moderate disorder, $\beta\epsilon[-1/4, -3/4]$, the XPS spectrum is blurred. For extreme disorder, $\beta\epsilon[-1, 0]$, localization is apparent for all states except those at the Fermi energy (See the density of states, which features a peak at the Fermi energy $\epsilon = 0$ associated with these states, Fig. 4.) In the localized limit, a recoilless peak appears in the XPS spectrum, at $E = \hbar\omega - \epsilon_{core}$, where ϵ_{core} is the energy of the core electron. Perhaps such a peak can be detected in heavily disordered metals.

The absorption spectrum for zero disorder, $\beta = -1/2$, has an X-ray edge singularity [10] at the low-energy threshold and cuts off somewhat at high energy, due to the conduction-band maximum. This general lineshape persists even for quite considerable disorder, $\beta\epsilon[-1/4, -3/4]$, until for very large disorder, $\beta\epsilon[0, -1]$, structures associated with states localized in the disorder appear. (See Fig. 2.)

The emission spectrum for zero disorder, $\beta = -1/2$, has two characteristic peaks, one at low energy associated with the van Hove singularity in the band structure [8] and a second one corresponding to a divergent X-ray edge singularity [10, 11] at the Fermi surface (Fig. 3). The densities of states are given, for reference, in Fig. 4. Recall that the conduction band is half-full of electrons.

With increasing disorder $\beta\epsilon[-3/4, -1/4]$ the van Hove singularity in the emission spectrum is blurred, but the edge singularity remains. However, at extremely strong disorder, the states not at the band center become localized (see the density of states, Fig. 4) and the emission spectrum has a major peak associated with the delocalized Fermi-surface electrons (whose wavefunctions overlap the core hole significantly) plus additional structure associated with localization.

In summary, off-diagonal disorder does lead to localization of the one-electron states and to alteration of the X-ray spectra from those expected for ordered metals. In particular, both one-electron features (such as van Hove singularities) and many-electron features (such as the X-ray edge singularity) are affected. However, the disorder effects are modest even for considerable disorder, $\beta\epsilon[-3/4, -1/4]$, and do not lead to major new features until the disorder is so great, $\beta\epsilon[-1, 0]$, that bonds are rather effectively and irreversibly broken.

Thus we believe that these calculations, although executed for a simplified one-dimensional Anderson model of off-diagonal disorder, provide a modicum of theoretical justification for the fact that X-ray spectra of liquid metals are often similar to those of crystals. At the same time the calculations suggest that XPS experiments should search for the recoilless peak found in the theory for the limit of high localization.

Acknowledgments—We are grateful to the U.S. Office of Naval Research for their generous support (Contract No. N00014-77-C-0537).

REFERENCES

1. Dow J. D. and Flynn C. P. *J. Phys. C: Solid State* 13, 1341 (1980).
2. Swarts C. A., Dow J. D. and Flynn C. P. *Phys. Rev. Lett.* 43, 158 (1979); Swarts C. A. and Dow J. D. (1984), to be published.
3. Anderson P. W. *Phys. Rev.* 109, 1492 (1958).
4. Satpathy S., Dow J. D. and Bowen M. A., *Phys. Rev.* B28, 4255 (1983), and references therein.
5. Petersen H. and Baratoff C., *Phys. Rev. Lett.* 35, 663 (1975).
6. Harrison W. A., *Electronic Structure and the Properties of Solids*, p. 46. Freeman, San Francisco (1980).
7. Bowen M. A. and Dow J. D. (1984), *Phys. Rev. B* in press.
8. Hove L. van, *Phys. Rev.* 89, 1189 (1953).
9. Doniach S. and Sunjic M., *J. Phys. C: Solid State Chem.* 3, 285 (1970).
10. Friedel J., *Comments Solid State Phys.* 2, 21 (1969).
11. Nozieres P. and Dominicis C. T. de, *Phys. Rev.* 178, 1097 (1969).

Electronic states of the (100) (2×1) reconstructed Ge surface

David V. Froelich, Marshall A. Bowen,[a] and John D. Dow

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 2 February 1984; accepted 1 May 1984)

We present calculations of the surface electronic state dispersion curves $E(\bar{k})$ of the (100) (2×1) reconstructed surface of Ge, and compare them with recent angle-resolved photoelectron measurements by Nelson *et al.* We assumed Chadi's asymmetric dimer model of the surface reconstruction and performed our calculations using the analytic Green's function technique, with an empirical tight-binding Hamiltonian.

PACS numbers: 73.20.Cw, 79.60.Eq

Recently, some success was achieved in characterizing the geometry of the (2×1) reconstructed (100) surface of silicon: Using an energy minimization technique Chadi[1] determined that an asymmetric dimer, Fig. 1, was the most energetically favorable model of the surface geometry, and predicted its surface state dispersion relations $E(\bar{k})$. However, Himpsel *et al.*[2] and Uhrberg *et al.*[3] produced angle-resolved photoemission spectra in disagreement with these first predictions of Chadi,[1] which had been made assuming the asymmetric dimer geometry and using a simple sp^3 tight-binding model of the surface electronic structure. For a while it was believed that the asymmetric dimer model is not realized physically. However, Bowen *et al.*,[4] using the sp^3s^* model of Vogl *et al.*,[5] and Mazur *et al.*,[6] using a multineighbor tight-binding model, subsequently demonstrated that the asymmetric dimer (2×1) reconstruction is compatible with the photoemission data, and attributed the original failure of the theory to the oversimplified sp^3 model of the electronic structure rather than to any deficiency in the asymmetric dimer geometry itself.

Recently, Nelson *et al.*[7] used angle-resolved photoemission to obtain the electronic structure of the (2×1) reconstructed[8] (100) surface of germanium. Their analysis of the data led them to suggest that the reconstruction of the Ge surface is similar to that of Si, i.e., an asymmetric dimer. There are, however, no calculations of the surface state dispersion relations $E(\bar{k})$ for Ge comparable with those for Si. Thus, we have extended the theory of surface states to Ge, assuming the same asymmetric dimer reconstruction as for Si.

We used the sp^3s^* tight-binding Hamiltonian of Vogl.[5] The surface states were found by diagonalizing an effective Hamiltonian which was obtained by using the analytic Green's function technique[9] and an evanescent wave method.[10] We created a surface by orbital removal,[11] and accounted for the effects on the Hamiltonian of the reconstruction-related bond-length changes by rescaling the nearest-neighbor interatomic matrix elements according to the d^{-2} law,[12,13] where d is the bond length.

In comparing with the experimental results, it is necessary to realize that the Ge(100) (2×1) surface consists of domains of reconstruction in which the rectangular unit cells are oriented perpendicular to each other[7] (see Fig. 2). Therefore, there is some ambiguity in determining the surface wave vector for each experimentally determined state. Himpsel *et al.*[2] circumvented this problem when investigating Si by rotating their sample in a plane in which the two domains were degenerate, thereby avoiding any ambiguity. Nelson *et al.*[7] also rotated their sample, but in such a way that the resulting spectra do not correspond to a unique determination of the Brillouin zone.

Our results are displayed in Fig. 3, and are labeled "A" through "F" for the path $\bar{\Gamma}$ to $\bar{J}$ to $\bar{\Gamma}$ in one surface Brillouin zone (Fig. 2) and "1" through "6" for the path $\bar{\Gamma}$ to $\bar{J}'$ in the other Brillouin zone. The predicted surface states, as displayed, are shifted down in energy by 0.5 eV (a typical theoretical uncertainty) from their calculated values. Studies of Si indicate[14] that the Vogl tight-binding Hamiltonian produces surface states of this group-IV semiconductor too high by ≃0.5 eV, but otherwise provides a good semiquantitative description of the surface-state bands $E(\bar{k})$. (It is possible that a more sophisticated[15] theory could reduce this 0.5 eV uncertainty somewhat.)

The Ge data of Nelson *et al.*[7] are also given in Fig. 3. As

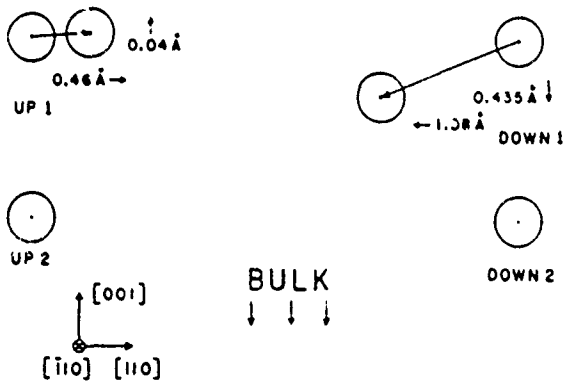


FIG. 1. The (2×1) asymmetric dimer reconstruction of Si proposed by Chadi (Ref. 1). In treating Ge, we have rescaled all lengths in proportion to 2.45 Å/2.35 Å, the ratio of Ge to Si bond lengths. The arrows indicate the relaxation of the surface layer atoms from their ideal bulk positions to their reconstructed positions. The plane of the figure is perpendicular to the surface. UP 1 and DOWN 1 are surface atoms. UP 1 being displaced away from bulk and DOWN 1 toward the bulk. UP 2 and DOWN 2 are subsurface atoms bonded to UP 1 and DOWN 1, respectively, and are assumed to be in their bulk positions. For Si they are 1.92 Å above the plane of the paper. If one were to build a ball and stick model of this Si surface these sites would be repeated every 3.84 Å in a direction perpendicular to the plane of the paper. In the [1̄10] direction the surface layer atoms alternate between UP 1 and DOWN 1, thus only half the surface unit cell is shown here.

 0734-211X/84/030390-03$01.00 © 1984 American Vacuum Society

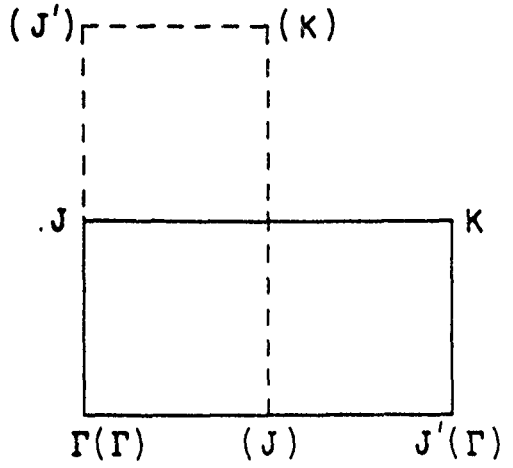


FIG. 2. The two perpendicular Brillouin zones on the (100) (2×1) reconstructed surface, as discussed in Ref. 7.

can be seen by their relative positions to the projected bulk states, all of the photoemission data are resonant with the valence band, except perhaps one point at the valence band maximum at $\bar{J}$. The following features of the data appear to be adequately explained by the theory:[16,17] The five high-energy points are associated with the resonance labeled F (which is degenerate with 6 at $\bar{\Gamma}$). The lowest-energy point is assigned to 5, D, or E. The ten mid-energy points are associated with resonance 6 and perhaps somewhat with E. The remaining point, at $\bar{\Gamma}$, is associated with either E and 3 or D and 5 (see below).

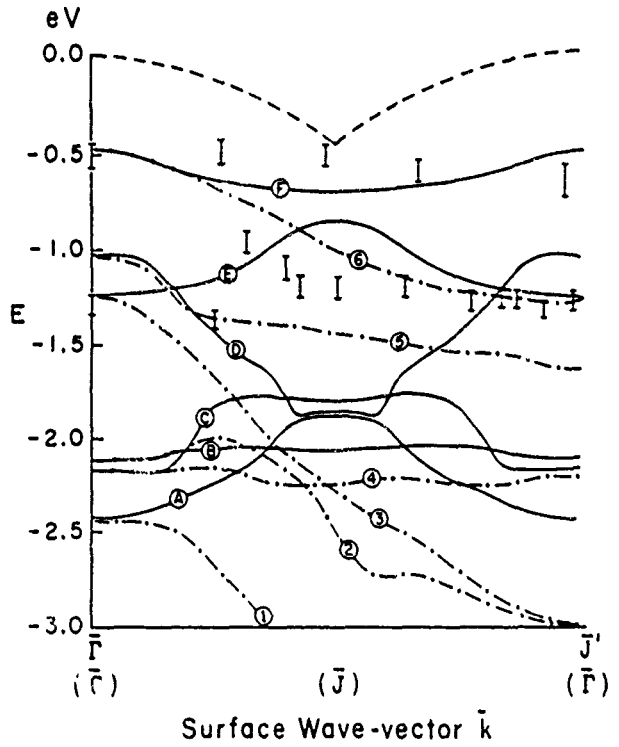


FIG. 3. The calculated surface state energies E (in eV) vs surface wave vector $\bar{k}$. (The states are shifted downward by the theoretical uncertainty, 0.5 eV, as discussed in the text.) Weak hybridizations at level crossings are neglected. The theory for the two perpendicular domains is superimposed for comparison with the photoemission data. The $\bar{\Gamma}$ to $\bar{J}$ states are shown as chained lines and are numbered. The $\bar{\Gamma}$ to $\bar{J}$ to $\bar{\Gamma}$ states are shown as solid lines and are lettered. The latter set of curves is symmetric about $\bar{J}$. The isolated points with the error bars are the photoemission data of Nelson *et al.* (Ref. 7). The maximum energy of the (unshifted) projected bulk valence band edge is displayed as a dashed line for reference purposes. (All states below this line in the figure are resonant with bulk states.)

We have also determined the eigenvectors and charge densities for each of the resonances. For the most part, in those regions of the Brillouin zone for which experimental data were obtained, the charge distribution is centered primarily on the "UP 1" site (see Fig. 1). Resonances F and 6 are primarily UP 1 everywhere in the Brillouin zone. The charge on E is mostly on the UP 2 site at $\bar{\Gamma}$ but shifts to the UP 1 site after crossing the D resonance. D is largely UP 1 everywhere except near its minimum energy where it is mostly "DOWN 2." 5 is primarily UP 1 near $\bar{\Gamma}$ and shifts to DOWN 2 about one-third of the way across the zone. 3 is mostly UP 2 in character throughout the region for which angle resolved photoemission spectra were obtained, shifting to predominantly UP 1 and other character below -1.7 eV. 1, 2, 4, A, B, and C have significant charge on all four sites UP 1, UP 2, DOWN 1, and DOWN 2, with the relative importance of the various sites changing with $\bar{k}$.

These results lead us to suspect that the resonance observed at -1.3 eV for $\bar{\Gamma}$ is due to the D or 5 resonances rather than E or 3, since their charge lies primarily on the surface, whereas the charge for E and 3 is on the subsurface layer. If this is so, then all the observed data can be associated with layer-1 surface resonances such that the charge is concentrated on the UP 1 site (which is farthest from the bulk and therefore is most easily observed).

Nelson *et al.*[7] gave a qualitative interpretation of their photoemission data that is similar to ours; however, they attributed the two bottom states near $\bar{\Gamma}$ as arising from the same resonance, whereas we find each arising from separate resonances. Both they and we agree that the photoemission data come primarily from two resonances passing through the zone, but our work indicates that these resonances are from differently oriented domains. Moreover, we predict other, perhaps weaker, resonances, as shown in Fig. 3.

The dispersion in our curves is qualitatively consistent with the photoemission data, but the gap near 6, E, and F at $\bar{J}$ is too small by a factor of 2. Given the simplicity of the model, this is adequate but not absolutely conclusive agreement between the theory and the data. However, we plan to explore the sensitivity of this gap to the amount of asymmetric dimer relaxation, in the hope that it can be used, in conjunction with experiments and total energy calculations, to determine precisely the surface geometry of the Ge (100) (2×1) surface.[18]

In conclusion our results lend support to the hypothesis that the (2×1) reconstructed (100) surface of germanium has the asymmetric dimer geometry, with relaxation roughly comparable with that of Si.

ACKNOWLEDGMENTS

We are grateful to the Office of Naval Research for their generous support (Contract No. N00014-77-C-0537 and a fellowship). We thank T. Chiang and L. Feldman for helpful

conversations and we are grateful to J. Nelson for discussion of his data.

[a]Permanent address: Department of Physics, Western Illinois University, Macomb, IL 61455

[1]D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979), J. Vac. Sci. Technol. **16**, 1290 (1979).

[2]F. J. Himpsel and D. E. Eastman, J. Vac. Sci. Technol. **16**, 1297 (1979).

[3]R. I. G. Uhrberg, G. V. Hansson, J. M. Nicholls, and S. A. Flodström, Phys. Rev. B **24**, 4684 (1981).

[4]M. A. Bowen, J. D. Dow, and R. E. Allen, Phys. Rev. B **26**, 7083 (1982), for a review of the Si(100) surface state literature, see the references cited therein.

[5]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

[6]A. Mazur, J. Pollmann, and M. Schmeits, Phys. Rev. B **26**, 7086 (1982).

[7]J. G. Nelson, W. J. Gignac, R. S. Williams, S. W. Robey, J. G. Tobin, and D. A. Shirley, Phys. Rev. B **27**, 3924 (1983).

[8]The reconstruction symmetry of the (100) surface of Ge is currently in dispute. Some research claim the surface reconstructs in a (4×2) pattern. This issue will be resolved only by further experimental investigation.

[9]R. E. Allen, Phys. Rev. B **20**, 1454 (1979).

[10]Y. C. Chang and J. N. Schulman, Phys. Rev. B **25**, 3975 (1982), **26**, 4400 (1982).

[11]J. Pollmann, in *Advances in Solid State Physics*, edited by J. Treusch (Vieweg, Braunschweig, 1979).

[12]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, 1980).

[13]The Hamiltonian of Ref. 5 is constructed with its parameters obeying the d^{-2} scaling law (Ref. 12). The accuracy of this law, which is highest for small bond length changes, is discussed in Ref. 5.

[14]J. P. Buisson, J. D. Dow, and R. E. Allen, Surf. Sci. **120**, L477 (1982); M. A. Bowen, J. D. Dow, and R. E. Allen, Phys. Rev. B **26**, 7083 (1982).

[15]The present surface state calculations are not self-consistent and do not allow for charge transfer. This can sometimes affect surface states. [See K. C. Pandey, Phys. Rev. Lett. **49**, 223 (1982); **47**, 1913 (1981).] More accurate surface states might result if the positions of the surface atoms were definitively known. Moreover, the approximations of the model Hamiltonian, such as the use of the d^{-2} law (Refs. 12 and 13) for treating reconstruction, are dictated by simplicity rather than quantitative accuracy—although any reasonable bond-length scaling argument would lead to qualitatively similar surface state dispersion relations. Therefore, the predictions here should be viewed as semiquantitative guides for interpreting data and for suggesting experiments, rather than as quantitatively reliable predictions of surface state energies. The model, despite its simplicity, is recommended by its history of successfully describing $E(\bar{k})$ for the (110) surfaces of III–V semiconductors (Ref. 19) and zinc chalcogenides (Ref. 20), and the (100)Si surface (Ref. 4), it also successfully predicted the surface state dispersion curves of InP (Ref. 21 and 22).

[16]In addition to the difficulties determining whether the reconstruction pattern is (2×1) or (4×2), it is sometimes difficult to determine unambiguously which states are surface states and which are bulk states, especially if both types of states lie at nearly the same energy. See Ref. 17.

[17]J. G. Nelson, W. J. Gignac, R. S. Williams, S. W. Robey, J. G. Tobin, and D. A. Shirley, Surf. Sci. **131**, 290 (1983).

[18]D. V. Froelich, M. A. Bowen, and J. D. Dow (unpublished). Preliminary results suggest that the gap between E and F near $\bar{J}$ is sensitive to the amount of asymmetry for the Ge–Ge surface dimerization, and, if taken literally (Ref. 15), indicate that the (100) surface geometry of Ge might be ≃20% closer than that of Si to a symmetric dimer model.

[19]R. P. Beres, R. E. Allen, and J. D. Dow, Solid State Commun. **45**, 13 (1983); R. P. Beres, R. E. Allen, J. P. Buisson, M. A. Bowen, G. F. Blackwell, H. P. Hjalmarson, and J. D. Dow, J. Vac. Sci. Technol. **21**, 548 (1982).

[20]R. P. Beres, R. E. Allen, and J. D. Dow, Phys. Rev. B **26**, 769 (1982).

[21]R. P. Beres, R. E. Allen, and J. D. Dow, Phys. Rev. B **26**, 5702 (1982).

[22]G. P. Srivastava, I. Singh, V. Montgomery, and R. H. Williams, J. Phys. C **16**, 3627 (1983).

Localized Perturbations in Semiconductors.

JOHN D. DOW

Department of Physics, University of Notre Dame - Notre Dame, IN 46556

1. – Introduction.

In these lectures we outline a simple but general theory of electronic states associated with localized perturbations in semiconductors. The basic problem we consider is the «deep-level problem», namely predicting the point-defect energy levels that lie near the middle of the band gap of a semiconductor. When we began work on this problem, a deep level was defined as a level that was not shallow, namely one more than 0.1 eV from the nearest band edge—a level that could not be thermally ionized at room temperature. (That definition has since been revised: see below.) Our own interest in the deep-level problem resulted from data of Wolford and Streetman for the N impurity in $GaAs_{1-x}P_x$ alloys [1]. This impurity appeared to be shallow in GaP, having a binding energy of only $\simeq 11$ meV, even smaller than the 35 meV effective-mass theory binding energy of the shallow donors S and Se. However, it became a genuine deep level in the alloy for $x \simeq 0.5$ and merged into the conduction band as a resonance for $x < 0.22$ (see fig. 1 [2, 3]). Thus the N impurity level was apparently shallow (for $x = 1$), deep (for $x \simeq 0.5$) and no level at all (for $x < 0.22$) as one varied alloy composition x continuously from GaP ($x = 1$) to GaAs ($x = 0$).

1'1. $GaAs_{1-x}P_x$. – The alloy host $GaAs_{1-x}P_x$ has a band structure that is well described by the virtual-crystal approximation [4] and varies continuously from the direct-gap band structure of GaAs (with the conduction band minimum at $\Gamma = (0, 0, 0)$ in the Brillouin zone) to the indirect-gap structure of GaP (with the conduction band minimum near the X-point: $(2\pi/a_x)(1, 0, 0)$) (see fig. 2). The band gap of GaAs is in the infra-red. Pure GaAs would emit such light because the band gap is direct, and the magnitude of the momentum of a thermalized electron-hole pair, $|k_e - k_h|$ can be equal to that of the emitted photon, $2\pi/\lambda$, which is essentially zero on the scale of the Brillouin zone. In contrast, GaP has an indirect-gap band structure, and so a thermalized conduction electron has a significantly different wave vector from a thermalized

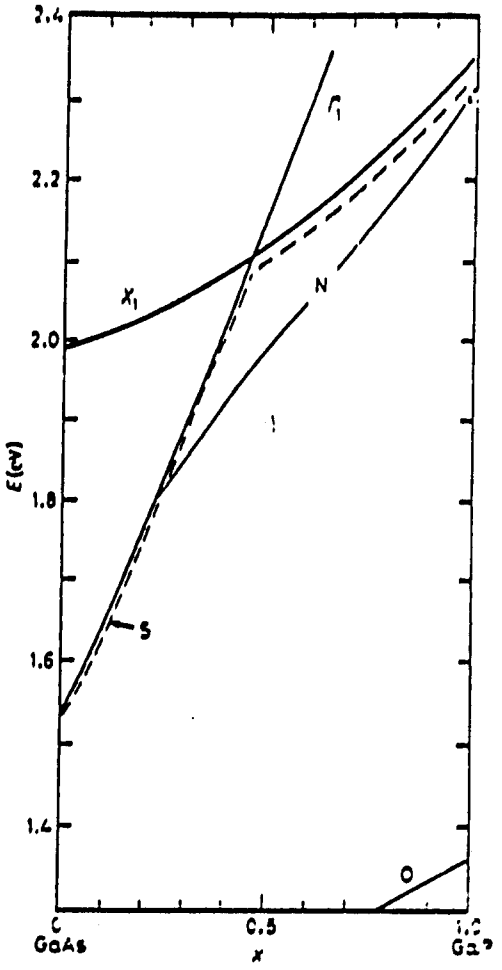


Fig. 1. – Illustration of the dependence of shallow and deep impurity levels on alloy composition x in $GaAs_{1-x}P_x$ alloys, after ref. [2, 3]. The zero of energy is the valence band maximum. The direct conduction band edge is Γ_1 and the indirect edge is X_1. The N and O deep levels are denoted by solid lines. The shallow levels of S (or Se) are denoted by dashed lines. Note that the direct-indirect cross-over occurs for $x \simeq 0.45$ and that the shallow-level binding energy is larger in indirect material (because the effective mass is larger).

hole. Hence pure GaP cannot emit light even though its band gap is in the green—a highly visible part of the spectrum. The alloy has become technologically important, because, for $x \simeq 0.4$, the band gap lies in the visible (red), but the band structure is still direct—hence this material is employed in red light-emitting diodes (LEDs).

1'2. *Column V site impurities* N, O, S *and* Se *in* $GaAs_{1-x}P_x$. – To fabricate a light-emitting diode that emits in the yellow or the green from these alloys, one needs a source or sink of crystal momentum, $\boldsymbol{K} \simeq (2\pi/a_L)(1, 0, 0)$, so that the selection rule $k_e - k_h \doteq K \simeq 0$ can be satisfied. Impurities can supply the

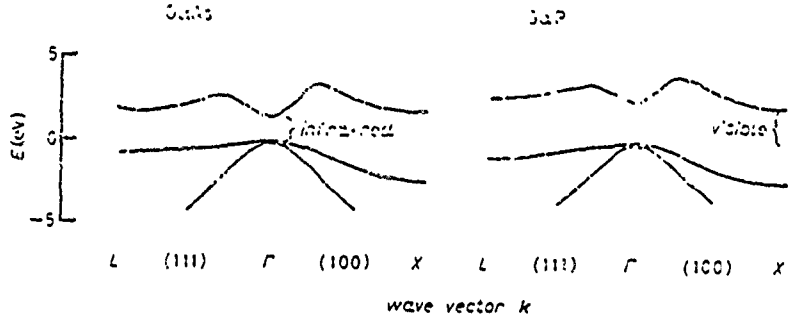


Fig. 2. – Electronic energy band structures $E(k)$ of GaAs and GaP from L to Γ to X along the (100) (*i.e.* Γ to X) and (111) (*i.e.* Γ to L) directions of the Brillouin zone, after M. L. COHEN and T. K. BERGSTRESSER: *Phys. Rev.*, 141, 789 (1966). Note that the band gap of GaAs is direct but in the infra-red; GaP has an indirect gap from the valence band maximum at Γ to the conduction band minimum at X, in the visible region of the spectrum.

needed momentum, with the impurities most likely to occur on the column V site of $GaAs_{1-x}P_x$, being N, O, S and Se. Ironically, two of these impurities, S and Se, produce *shallow* levels in the band gap of $GaAs_{1-x}P_x$ that lie close to the conduction band edge and follow the edge as the composition varies. But two do not. Oxygen lies several tenths of an electronvolt deep in the band gap of GaP and its energy level decreases linearly as x decreases—it is a genuine deep level by all definitions. The behavior of N (with respect to the valence band maximum) is especially interesting: in GaP it is *apparently shallow* with a 11 meV binding energy, and, with decreasing alloy composition x, its energy level decreases linearly, similar to the oxygen deep level, becoming a genuine deep trap (by the old definition: more than 0.1 eV from the conduction band edge) for $x \simeq 0.5$. At $x = 0.32$, however, the N level goes into the conduction band. In other words, N appears to be shallow energetically for $x = 1$, is deep for $x = 0.5$ and is a resonance for $x = 0$ (see fig. 1). Moreover, the N level is unattached to the conduction band edge, and dE/dx for this level is characteristic of a deep trap such as O. These facts led us to believe that N is, in fact, a deep level whose energy accidentally lies close to the conduction band edge in GaP and becomes resonant in GaAs—and focussed our attention on N as the prototypical deep trap [5].

2. – The Vogl model of electronic structure.

The foundation for much of what we shall discuss in these lectures is an empirical tight-binding theory of electronic structure developed by VOGL *et al.* [6]. This theory has three distinguishing features: it properly represents

the chemistry of the sp^3 bonding, because it has a basis which includes one s orbital and three p orbitals at each atomic site (as well as one additional s orbital: s^*). ii) It produces indirect-gap band structures for Si and GaP with a minimum number of basis functions, five per site. (Pure sp^3-basis tight-binding models do not: the extra s^* orbital produces the indirect band structure by pushing the indirect conduction band minimum down in energy). iii) The parameters of the Hamiltonian exhibit manifest chemical trends that are codified in scaling rules: the diagonal matrix elements are related to atomic energies, and the off-diagonal matrix elements are inversely proportional to d^2, the square of the bond length (Harrison's rule [7]).

This empirical Hamiltonian was arrived at by VOGL and HJALMARSON after a great deal of labor and represents an attempt to simultaneously describe the energy bands of sixteen semiconductors. The Vogl model drew much of its inspiration from Harrison's bond orbital model [8], which was one of the first successful attempts to develop a simple Hamiltonian for describing chemical trends for many semiconductors—it described valence band structures rather accurately. A distinguishing feature of the Vogl model is its ability to reproduce general features of the lowest conduction bands as well.

The scaling rules for chemical trends in the parameters of the Vogl model are very important. Because of them, the Vogl Hamiltonian can be generalized to treat *inhomogeneous* semiconductors—even though the information contained in the model's parameters comes exclusively from the known energy band structures of homogeneous semiconductors. For example, if one atom is changed (*e.g.*, one P in GaP is replaced by a N atom), the matrix elements for the changed Hamiltonian can be deduced by changing the host matrix elements according to the scaling rules.

The basic philosophy in $\mathbf{r}$ (R) space of empirical tight-binding theory is similar to the philosophy for ordinary pseudopotential theory (in k-space): remove the distant parts (in R-space) of the Hamiltonian and lump them into near-neighbor parameters that are determined empirically.

Tight-binding basis functions $|nbk\rangle$ are constructed from (unknown) localized quasi-atomic orbitals $|nbR_l\rangle$

$$|nbk\rangle = N^{-\frac{1}{2}} \sum_l |nbR_l\rangle \exp[ik\cdot R_l - ik\cdot v_b], \tag{1}$$

where $n = s, p_x, p_y, p_z$, or s^* specifies the basis orbital, we have $b =$ a (anion) or $b =$ c (cation), k is the Block wave vector, R_l is an anion site in a zincblende structure, we have $v_a = 0$, and v_c is the position of the cation relative to the anion. In this basis, the secular equation reduces to the 10×10 system

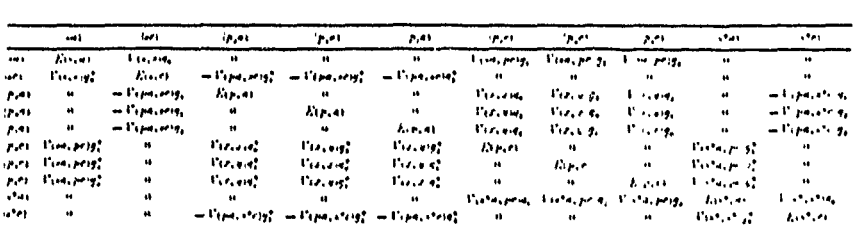
$$(H^{\alpha\beta} - \varepsilon(k\lambda))\, k\lambda\rangle = 0\,. \tag{2}$$

where we have (symbolically) the Bloch state

$$|k\lambda\rangle = \sum_{n,b} |nbk\rangle\langle nbk|k\lambda\rangle \,. \tag{3}$$

In the $|nbk\rangle$ tight-binding basis we have the 10×10 Hamiltonian matrix ($\lambda = 1, 2, \ldots, 10$)

(4)



Here we have for $k = (2\pi/a_L)(k_1, k_2, k_3)$ and a_L the lattice constant

$$g_0(k) = \cos(\pi k_1/2)\cos(\pi k_2/2)\cos(\pi k_3/2) - i\sin(\pi k_1/2)\sin(\pi k_2/2)\sin(\pi k_3/2)\,,$$

$$g_1(k) = -\cos(\pi k_1/2)\sin(\pi k_2/2)\sin(\pi k_3/2) + i\sin(\pi k_1/2)\cos(\pi k_2/2)\cos(\pi k_3/2)\,,$$

$$g_2(k) = -\sin(\pi k_1/2)\cos(\pi k_2/2)\sin(\pi k_3/2) + i\cos(\pi k_1/2)\sin(\pi k_2/2)\cos(\pi k_3/2)$$

and

$$g_3(k) = -\sin(\pi k_1/2)\sin(\pi k_2/2)\cos(\pi k_3/2) + i\cos(\pi k_1/2)\cos(\pi k_2/2)\sin(\pi k_3/2)\,.$$

We have

(5)

$$\begin{aligned}
&\langle saR|H|saR\rangle = E(s, a)\,,\\
&\langle paR|H|paR\rangle = E(p, a)\,,\\
&\langle scR|H|scR\rangle = E(s, c)\,,\\
&\langle pcR|H|pcR\rangle = E(p, c)\,,\\
&\langle s^*aR|H|a^*aR\rangle = E(s^*, a)\,,\\
&\langle s^*cR|H|s^*cR\rangle = E(s^*, c)\,,\\
&\langle saR|H|scR\rangle = V(s, s)\,,\\
&\langle p_xaR|H|p_xcR\rangle = V(x, x)\,,\\
&\langle p_yaR|H|p_ycR\rangle = V(y, y)\,,\\
&\langle saR|H|p_xcR\rangle = V(sa, pc)\,,\\
&\langle p_xaR|H|scR\rangle = V(sc, pa)\,,\\
&\langle s^*aR|H|p_xcR\rangle = V(s^*a, pc)\\
&\text{and}\\
&\langle p_zaR|H|s^*cR\rangle = V(pz, s^*c)\,.
\end{aligned}$$

This is the basic Hamiltonian to be used throughout the present work. The reader should become familiar with it by working the following three problems.

Problem 1. Compute the energy band structure at $k = 0$ of GaP, taking your zero of energy at the valence band maximum. Compare your results with fig. 2 of ref. [6]. For GaP the Vogl tight-binding matrix elements are $E(s, a) = -8.1124$, $E(p, a) = 1.1250$, $E(s, c) = -2.1976$, $E(p, c) = 4.1150$, $E(s^a, a) = 8.5150$, $E(s^a, c) = 7.1850$, $V(s, s) = -7.4709$, $V(x, x) = 2.1516$, $V(x, y) = 5.1369$, $V(sa, pc) = 4.2771$, $V(sc, pa) = 6.3190$, $V(s^a a, pc) = 4.6541$ and $V(pa, s^a c) = 5.0950$. (For other semiconductors, see ref. [6].)

Problem 2. Compute the band structure at the X-point, $k = (2\pi/a_L)(1, 0, 0)$, of the Brillouin zone.

Problem 3. Write down the change in the Hamiltonian matrix, ΔH, in the $|nbR\rangle$ basis for a N atom replacing P in GaP at $R = D$. Assume that the bond length does not change when N replaces P and that the matrix elements involving s^a remain unaltered (because s^a simulates nonlocal effects of distant neighbors). Neglect distinctions between the host basis orbitals $|nbR\rangle$ and the corresponding impurity orbitals (in subsequent work, we shall actually be using the impurity orbitals at the impurity site). Show that the matrix is 4×4 and diagonal. Suppose further that the diagonal matrix elements V_s and V_p of ΔH are given by the Vogl-Hjalmarson scaling rules [6]

$$V_s = 0.8(w(s, \mathrm{N}) - w(s, \mathrm{P})) \tag{6a}$$

and

$$V_p = 0.6(w(p, \mathrm{N}) - w(p, \mathrm{P})) , \tag{6b}$$

where the atomic-orbital energies w for N and P are [6] $w(s, \mathrm{N}) = -25.7130$, $w(p, \mathrm{N}) = -15.4388$, $w(s, \mathrm{P}) = -18.9425$ and $w(p, \mathrm{P}) = -10.6544$. Finally, using ref. [6], determine the numerical values of the defect potentials V_s and V_p for O, S, Se and Te and for B, C, N, O and F. If you have worked problem 3, you have set up the Hamiltonian for obtaining the deep levels of N in GaP.

3. – The Hjalmarson *et al.* theory of deep impurity levels.

3·1. *Qualitative remarks.* – In the late 1950's Kohn and collaborators developed the effective-mass theory of shallow impurities in semiconductors [9]. According to this theory, an impurity such as S substituting for As in GaAs produces a donor electron that orbits the extra nuclear charge of S (relative

to As) in a large hydrogenic orbit, the envelope wave function for which satisfies a Schrödinger equation

$$[(-\hbar^2/2m^*)\nabla - (Ze^2/\varepsilon r)]\psi = (E - E_c)\psi , \tag{7}$$

where m^* is the conduction band effective mass, Z is the excess valence of the impurity atom with respect to the host atom it replaces (unity for S on an As site in GaAs), ε is the GaAs static dielectric constant, and E_c is the energy of the conduction band edge (at $k = 0$ in GaAs). This effective-mass state has a total wave function that is primarily a product of this envelope function and the periodic part of a Bloch function evaluated at the wave vector of the conduction band minimum [9, 10]: it is made up primarily from one band (the GaAs conduction band, in this case). The effective-mass state is hydrogenic and virtually 100% hostlike. The impurity level is «attached» to the conduction band edge with a small binding energy of order 10 meV (13.6 eV$(m^*/m_0)^2/\varepsilon$, where m_0 is the free-electron mass) and follows the edge when the edge moves as a result of externally applied pressure or alloying (*e.g.*, alloying GaAs with GaP). The shallow levels control the electrical properties of the semiconductor, and, although the impurity potential in the central cell often deviates greatly (a few eV) from the Coulombic value, $-Ze^2/\varepsilon r$, only the long-ranged Coulombic potential seems to have a significant effect on the shallow states. (This should be bothersome, because central-cell potentials of order 1 eV must produce some effect on that scale.) Moreover, the shallow states are localized in k-space but delocalized in real space.

The effective-mass theory accounts for many of the data for impurity levels in the band gaps of semiconductors; however, it does not account for many facts, including the following: i) some isoelectronic impurities, such as N replacing P in GaP, produce levels in the gap despite the fact that their valence differences Z are zero, and ii) some levels in the gap lie far (more than 0.1 eV) from a band edge and are «deep levels». Early attempts to explain these facts attempted to modify the effective-mass theory to produce larger binding energies.

A central point of the Hjalmarson theory [11, 12] is that *every* heterovalent substitutional impurity produces *both* «deep levels» and shallow levels, and that *the «deep levels» do not necessarily lie in the fundamental band gap, but may be resonant with the host bands.* The deep and shallow states are two qualitatively different types of impurity states that coexist, but are rarely observed simultaneously. Deep levels are controlled by the central-cell potential, have wave functions that are linear combinations of wave functions from many (~ 10) host bands, are often antibonding in character and are largely hostlike. The deep-level energies are often unattached to nearby band edges and do not follow them when they move as a result of pressure or alloying. Deep impurity states are localized in real space and delocalized in k-space. The apparent

« binding energy » of a deep level relative to a nearby band edge is often large (tenths of eV) in magnitude and can be negative. When a deep level falls in the fundamental band gap, it can trap excitons or charge carriers, often enhancing the nonradiative recombination of electrons and holes. Thus deep levels tend to influence the optical properties of semiconductors even at concentrations as low as $10^{15}/cm^3$.

In these lectures, we limit ourselves to levels associated with sp^3-bonded substitutional impurities. (For discussions of interstitial sp^3-bonded impurities and transition metal impurities, see ref. [13] and [14], respectively.) Thus, in the energy vicinity of the band gap, for substitutional impurities in tetrahedral semiconductors, we expect exactly four deep levels to originate from the sp^3 bonding, three of which are degenerate: a s-like A_1 level and a p-like triply degenerate T_2 level. (A_1 and T_2 are irreducible representations of the tetrahedral group T_d.) If these « deep levels » due to the central-cell potential all happen to lie above the conduction band edge, as in the case of $GaP:S_p$ (S on a P site in GaP), then the only levels in the gap are the shallow levels associated with the long-ranged Coulomb potential—and S is termed a shallow impurity because only its shallow levels are observed in the gap. If one of the deep levels due to the central-cell potential falls within the fundamental band gap, as with $GaP:O_p$, then the impurity is termed « deep ». But a central point is that both shallow and deep levels of the same impurity coexist (fig. 3): they are distinct (although deep levels near a band edge may hybridize with shallow levels). Isoelectronic defects, such as $GaP:N_p$, have no long-ranged Coulomb potential and hence no shallow levels: all of their defect levels (except possibly levels associated with a strain field surrounding them) are « deep ».

3·2. *Energy scales and the nature of the theory.* – Before constructing a theory of deep impurity levels, one should first determine the important physics.

To begin with, the bonding in semiconductors is sp^3 in character, and a proper treatment of a localized defect state must account for this. The spectral distribution of the sp^3 bonds covers $\simeq 20$ eV, the combined widths of the valence bands and the lowest conduction bands.

The defect potential in the central cell can be crudely estimated as the difference between the atomic energies of the defect and the host atom it replaces —and is typically *several electronvolt* in magnitude—of order 5 eV, 7 eV and 15 eV for S, N and O (all on the P site) in GaP and 4 eV for P in Si. The fact that the central-cell defect potential is so large should be extremely puzzling, especially in the case of the shallow donors P in Si and S_p in GaP—because Nature requires that a perturbation of several electronvolt exhibit itself on a scale of order of electronvolt, and the shallow impurities appear at first glance to exhibit consequences of the central-cell defect potential on only the millielectronvolt scale. The resolution of this dilemma lies in the fact that the shallow donors also produce « deep resonances », quasi-localized states at energies of order 1 eV

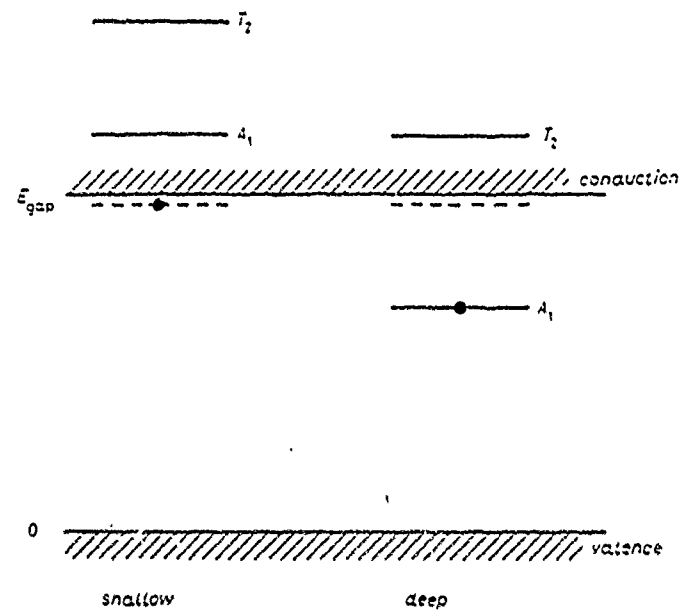


Fig. 3. – Schematic illustration of the difference between «shallow» and «deep» sp^3-bonded substitutional (donor) impurities, after ref. [3]. The shallow energy levels in the band gap are dashed. The deep levels of A_1 (s-like) and T_2 (p-like) symmetry are denoted by heavy lines. In the case of a «shallow impurity» the deep levels are resonances and lie outside the fundamental band gap; for a «deep impurity» at least one deep level lies within the gap. The lowest level is occupied by the extra electron (dark circles) if the impurity has a valence one greater than the host atom it replaces (*e.g.*, S or O on a P site in GaP).

above the conduction band minimum: the existence of these resonances has been appreciated only recently. This notion, that impurities produce «deep» levels above the conduction band minimum, requires a new definition of «deep». The old definition was that any level in the fundamental band gap more than 0.1 eV from the nearest band edge was «deep». Now, following HJALMARSON *et al.*, we *define* a deep level as one whose physics is controlled by the central-cell potential: as a result, «deep» levels now may have very small (< 0.1 eV) binding energies (such as the N level in GaP) or may lie above the conduction band edge with «negative binding energies», resonant with the host bands («deep resonances»). They may also lie resonant with the valence bands (binding energies greater than the band gap).

In addition to the central-cell potential, several other physical effects influence deep levels on a scale of a few tenths of an eV. These include lattice relaxation around the defect [15] and charge state splitting [16, 17] of the defect levels (*e.g.*, the difference due to electron-electron interactions in the S^0 and S^- one-electron energy levels in Si). Furthermore, the Coulomb potential outside the central cell, $-Ze^2/\varepsilon r$, is also of order 0.1 eV for $r \gtrsim a_L$. In much of the work we

discuss, we neglect all such effects of order 0.1 eV and concentrate on predicting the *chemical trends* in deep energy levels, from one impurity to another or from one host to another. By neglecting these effects, we obtain a very simple theory in which the defect potential matrix is diagonal (in a localized basis) and is localized to the central cell of the defect—because of the scaling rules of the Vogl Hamiltonian.

The band gap of a typical semiconductor is $\simeq 1$ eV and studies of deep levels are directed toward determining the levels with an accuracy of < 0.1 eV. However, the band gap energy is *not* a scale of physical relevance to the deep-impurity problem, because deep levels (whether «bound» in the gap or resonant) are unattached to band edges. Instead the band gap energy determines the scale of *observability* of most deep levels. A simple way to think about deep impurity levels is that they lie throughout the $\simeq 20$ eV range of the sp^3 bond, but that only the small fraction of these levels that lies within the «window» of the band gap is observable by conventional means. Hence a complete description of deep-level experiments on the scale of observability of deep levels requires a theory with an accuracy of 0.1 eV out of 20 eV, or 0.5 %. No contemporary theory is capable of such an accuracy: the best accuracy achievable is a few tenths of an eV. Therefore, the goal of theory should not be to predict the absolute energies of deep levels in the band gap, because this goal is presently unattainable. Rather, theories should be constructed with the intent of simply displaying the physics of deep levels, predicting chemical trends in data and predicting qualitative phenomena—such as suggesting the conditions under which a deep resonance should descend into the band gap and become a bound deep level.

Because of the intrinsic limitations of contemporary theory, HJALMARSON *et al.* constructed a theory of deep levels that considered only the central-cell impurity potential of the defect. The theory can be and has been modified to include lattice relaxation, charge state splittings and the long-ranged part of the screened Coulomb interaction, but the requirement of simplicity is best met with the Hjalmarson model. In fact, the model's predictions have turned out to be in remarkably good agreement with the data.

In 1969 LANNOO and LENGLART [18] predicted the deep energy level of the diamond vacancy using a simple tight-binding model. Their approach predicts deep levels in good agreement with the most recent calculations [19-26], and their tight-binding ideas provided an essential guide for the development of the Hjalmarson theory. The two elements that are missing from that early work are i) an accurate treatment of both the sp^3 character of the chemical bond and the indirect conduction band structure and ii) a quantitative prescription for predicting the deep levels of impurities as well as vacancies, that is, a scheme for determining the defect potential of an impurity. (For a vacancy the defect potential is infinitely positive, as shown by LANNOO and LENGLART, causing the defect «atom» to be decoupled from the host by virtue of the

infinite-energy denominators in perturbation theory.) Subsequent tight-binding theories of deep impurity levels met with varying degrees of success, but a significant improvement occurred as a result of the work of Vogl *et al.*, which produced the sp^3s^* Hamiltonian with manifest *chemical trends* in its parameters and with adequate conduction bands.

3·3. *Hjalmarson theory of deep levels.* – The Hjalmarson theory of deep impurity levels is a Green's function theory of the type proposed originally by KOSTER and SLATER [27]. The host Hamiltonian matrix H^0 is the Vogl Hamiltonian, eq. (4). Because lattice relaxation and changes of bond length have been neglected, the defect potential matrix V, in the sp^3s^* basis localized at each site, is zero except at the impurity site and is diagonal at the impurity site:

$$(nbR|V|n'b'R') = \delta_{n,n'}\delta_{b,b'}\delta_{R,R'}v_n. \tag{8}$$

Here δ is the Kronecker delta-function and we have $v = (V_s, V_p, V_p, V_p, 0)$; V_s and V_p are given by eqs. (6). The Hamiltonian is

$$H = H^0 + V. \tag{9}$$

Because the defect potential matrix is localized, a Green's function method is useful. Normally to find the impurity levels in a crystal of N unit cells with two atoms per cell and five orbitals per atom, one must solve by brute force a $10N \times 10N$ matrix equation. With a Green's function method, one need only solve a 4×4 matrix: the size of the defect rather than the size of the crystal. In fact, tetrahedral point-group symmetry reduces this 4×4 matrix to four 1×1 matrix equations.

The Green's function matrix for the perturbed crystal is

$$G(E) = (E - H)^{-1} \tag{10}$$

and is related to the unperturbed Green's function

$$G^0(E) = (E - H^0)^{-1} = \sum_k |k\lambda\rangle (E - E_{k\lambda})^{-1} \langle k\lambda| \tag{11}$$

by Dyson's equation

$$G = G^0 + G^0 V G, \tag{12}$$

(Verify this by multiplying on the left by $E - H^0$ and on the right by $E - H$.) The formal solution of Dyson's equation is the matrix

$$G = [1 - G^0 V]^{-1} G^0, \tag{13}$$

which has nontrivial solutions at the energies determined by the secular equation

$$\det[1 - G^0 \Gamma] = 0 . \tag{14}$$

Taking matrix elements of the secular equation in the $|nlR\rangle$ basis, we find the eigenvalue equations for the degenerate energy E

$$\Gamma_s^{-1} = (sD|G^0|sD) = I^{-1} \sum_{kl} |(sD|kl\Gamma)|^2 [E - E_{kl}] \tag{15}$$

and

$$\Gamma_p^{-1} = (p_\nu D|G^0|p_\nu D) . \tag{16}$$

Rewriting these results in terms of the local host spectral densities $D^0_{A_1}(E')$ and $D^0_{T_1}(E')$, we have [28]

$$\Gamma_s^{-1} = P\int_{-\infty}^{\infty} dE' D^0_{A_1}(E')/[E - E'] \tag{17}$$

and

$$\Gamma_p^{-1} = P\int_{-\infty}^{\infty} dE' D^0_{T_1}(E')/[E - E'] \tag{18}$$

with

$$D^0_n(E') = (nD|\delta(E' - H^0)|nD) = \sum_{kl} |(nD|kl\Gamma)|^2 \delta(E' - E_{kl}) . \tag{19}$$

To see how these results, eqs. (17) and (18), are obtained more directly, it is useful to take matrix elements of Dyson's equation (12) in the $|nlR\rangle$ basis and to consider the impurity site $lR = D$: we have

$$(nD|G|n'D) = (nD|G^0|n'D) - \sum_\nu (nD|G^0|\nu D)\Gamma_\nu (\nu D|G|n'D) , \tag{20}$$

where n, n' and ν range over s, p_x, p_y and p_z. Because H^0 is invariant under the operations of the tetrahedral point group T_d, H^0 and $(E - H^0)^{-1}$ are invariant operators. Since the s-state transforms according to the A_1 irreducible representation of T_d and the p-states transform according to the x, y and z rows of the T_2 representation, $(nD|G^0|n'D)$ is diagonal in n and eqs. (17) and (18) follow.

The energy E is always to be interpreted as having an infinitesimal positive imaginary part $i0$; this gives the correct boundary conditions for the Green's function. Because of this and the identity $(x - i0)^{-1} = \mathcal{P}(1/x) - i\pi\delta(x)$, we have [28]

$$G(E) = \mathcal{P}(E - H)^{-1} - i\pi\delta(E - H) ,$$

where $\mathcal{P}$ denotes a principal value and $\delta(x)$ is Dirac's delta-function.

Equations (17) and (18) are the central equations of the Hjalmarson theory. To solve them, one needs to first evaluate the spectral densities $D_n^0(E')$ for the host. When the Hamiltonian H^0 is diagonalized to find the eigenvalues $E_{kλ}$, the overlap integrals $\langle nbk'k\lambda'\rangle$ are the components of the normalized eigenvectors. In terms of these quantities, we have

$$\langle nbR'k\lambda'\rangle = N^{-\frac{1}{2}} \exp[ik\cdot(R + \tau_b)]\langle nbk|k\lambda'\rangle .$$

Hence we have

$$D_n^0(E') = N^{-1} \sum_{k\lambda} |\langle nbk|k\lambda'\rangle|^2 \delta(E' - E_{k\lambda}) ,$$

which can be summed either using the Lehmann-Taut method [29] or (for the deep-level energy E in the fundamental band gap) using the special-point method [30].

In practice, the deep levels $E_{A_1}(\Gamma_1)$ and $E_{T_2}(\Gamma_{25}')$ are calculated by computing the functions $V_1(E_{A_1})$ and $V_2(E_{T_2})$ as follows: i) a value of E is selected, ii) the spectral densities $D_n^0(E')$ are evaluated, iii) the right-hand sides of eqs. (17) and (18) are evaluated, and iv) V_1 and V_2 are determined from those equations. Plots of E *vs.* V for A_1 and T_2 states then give predictions of deep levels *vs.* defect atomic energy (see eqs. (6)). The vacancy levels are the asymptotes $E(V = \infty)$ of these curves.

Problem 4. Compute $V(E)$ for energies E outside the host band, in the case of a defect in a one-dimensional nearest-neighbor tight-binding crystal. (Hint: $H^0 = \delta \sum_R [|R\rangle(R+1| - |R+1\rangle(R|]$, $V = V_0|D\rangle(D|$. Compute the band structure E_k. Then compute $\langle R|G^0|R'\rangle = N^{-1} \sum_k (E - E_k)^{-1} \exp[ik(R - R')]$. To evaluate $\langle R|G^0|R'\rangle = \langle R - R'|G^0|0\rangle$ analytically for energies outside of the band, use a contour integral over the unit circle.) Repeat this calculation for the defects on the P site in GaP, using first one special point [30] and then ten special points to evaluate the sums over k. You will obtain good results with ten special points.

4. – Qualitative physics.

The *qualitative* physics determining deep levels is depicted for the case of $GaP:N_p$ in fig. 4, after [11]. In this figure we consider, for simplicity, only the *s*-states of the atoms (and the A_1-symmetric defect level) and note that the Ga atom's energy ε_{s_a} lies above the P energy ε_{p_c}. When these two widely separated atoms are brought together into a molecule, the levels repel—resulting in a bonding-antibonding splitting that, in lowest order of perturbation theory about the infinite-lattice-constant limit, is proportional to $v^2/(\varepsilon_{s_a} - \varepsilon_{p_c})$, where

v is the Ga-to-P transfer matrix element (and is about the same for all semiconductors [7]). The important point is that the bonding-antibonding splitting is inversely proportional to the energy denominator $\varepsilon_{Ga} - \varepsilon_P$. When these molecules are brought together into a solid, the antibonding states produce the conduction band and the bonding states yield the valence band, with the fundamental band gap in between.

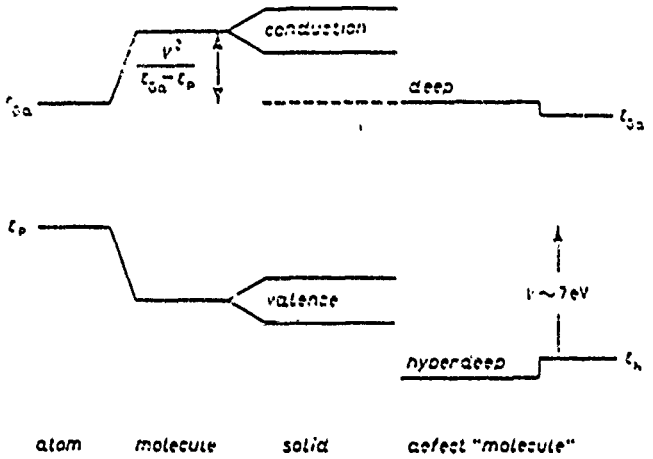


Fig. 4. – Schematic illustration of the qualitative physics governing deep levels, after ref. [3, 11]. See text.

Now imagine a «defect molecule» with a N impurity replacing one of the P atoms. Its atomic energy is $\simeq 7$ eV lower than that of P, and so, when it interacts with a Ga atom to form a molecule, the resulting bonding-antibonding splitting is smaller than for Ga and P, because the energy denominator $\varepsilon_{Ga} - \varepsilon_N$ is $\simeq 7$ eV larger than the denominator $\varepsilon_{Ga} - \varepsilon_P$. As a result, the level in the gap (the deep level) can lie below the conduction band edge in the gap. In fact, fig. 4 illustrates that the issue of whether a deep level is «bound» in the gap or «resonant» with the host bands depends primarily on whether the bands are broad enough to cover up the deep level; that is, it depends on the amount of bonding-antibonding splitting.

Several features of fig. 4 are worthy of special mention: i) The N deep level is derived from the Ga dangling-bond energy ε_{Ga} and is antibonding and *hostlike* (Ge-like), not impuritylike (N-like). The N-like level is the bonding *hyperdeep level* lying below the valence band; it is electrically inactive, being full of electrons, and is normally unobserved. ii) The deep level is orthogonal to the hyperdeep level. iii) The deep level is repelled upward by the hyperdeep level by means of the bonding-antibonding level repulsion. iv) The deep level is «pinned» to the Ga dangling-bond level and cannot be pulled below it: Imagine decreasing the energy of the N level ε_N relative to ε_P from ~ -7 eV to $\simeq -15$ eV

(oxygen) and then to $\simeq -1000$ eV (an ideal vacancy): the deep-level energy will move down only slightly, never becoming deeper than the Ga dangling bond or ideal P vacancy energy of ε_{Ga}. This is the meaning of «deep-level pinning»: *Major changes in the deep-level potential result in only minor changes of the deep-level energy*, or $|dE'/dV| \ll 1$.

The deep-level pinning can be illustrated by plotting the deep level in the gap *vs.* the defect potential V, which (according to the Vogl model's scaling rules) is proportional to the difference in atomic energies of the defect and the host (P atom). This is done in fig. 5 for the s-like A_1 states of defects substituting for P in GaP. The curve $E(V)$ is similar to a hyperbola, having the energy of a Ga dangling bond or a P vacancy as its asymptote. One can see that $E(V=\infty)$ corresponds to a vacancy, because, as the magnitude of the defect potential increases, the defect atom becomes less and less coupled to the host (recall that in perturbation theory the coupling is inversely proportional to an energy denominator of order V) until for $V=\infty$ the defect is totally uncoupled, namely a vacancy [18]. Once one recognizes that the physics of deep levels results in a hyperbolalike curve $E(V)$, the problem of predicting deep

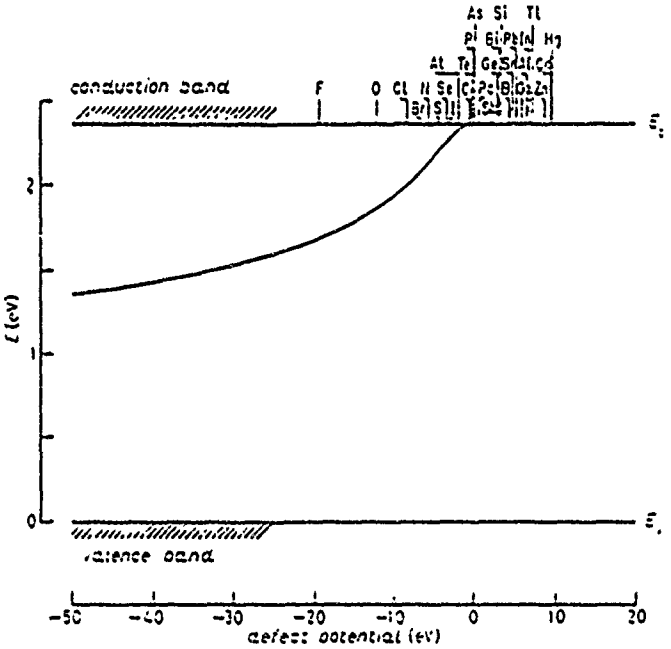


Fig. 5. – Energy levels in the band gap, as calculated by HJALMARSON *et al.* [11], *vs.* defect potential for A_1 symmetric states of defects on the P site in GaP, after ref. [3]. Note that, if the theory is taken literally with no allowances for a theoretical uncertainty, the S defect is predicted to have a deep level in the gap, just slightly below the conduction band edge. Experimentally it is known that S is a shallow donor; hence, one must make allowances for the uncertainty in the theory and recognize that in fact the deep level for S must lie slightly above the conduction band minimum in GaP.

levels reduces to predicting two numbers: the ideal-vacancy energy $E(\infty)$ and the threshold potential V_T at which the resonant deep level passes into the gap and ceases being a resonance.

Figure 6 illustrates very schematically the wave functions of deep levels, using only s-states, for simplicity. The host valence band of GaP has a bonding wave function that is largely P-like, but with a significant Ga-like component.

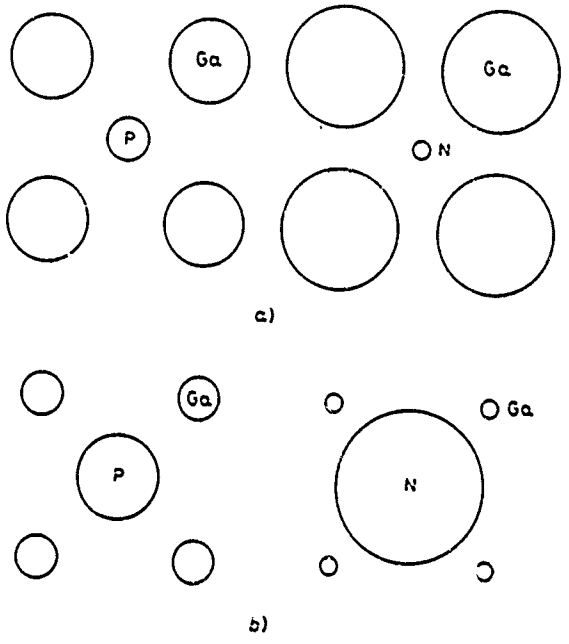


Fig. 6. – Schematic illustration of the wave functions in the (bonding, *a*)) valence and (antibonding, *b*)) conduction bands of GaP (using a s-state model) and in the bonding (hyperdeep) and antibonding (deep) levels of N on a P site in GaP, after ref. [3].

When N replaces P, N is more electronegative and attracts electrons to it. As a result, the bonding hyperdeep level has a wave function that is overwhelmingly N-like, with just a small Ga component. The antibonding deep level is orthogonal to the hyperdeep level, and so has a wave function that is almost exclusively Ga-like, with only a small N component.

All of these ideas have been abstracted from the Hjalmarson *et al.* theory.

5. – Evidence supporting the theory.

The theory of Hjalmarson *et al.* [11, 12] has made literally thousands of predictions of deep levels and can account for an extremely large body of data. Here we review a few representative predictions of the theory.

5·1. *Wave functions.* – Figure 7 shows the magnitude of the deep-level wave function of substitutional S^- in Si, as a function of distance R from the S center, in comparison with electron nuclear double resonance (ENDOR) data [31, 32]. The ENDOR measurements give the charge density $|\psi(R)|^2$ at Si sites adjacent to and nearby the S defect. The data do not reveal the phase of the wave function, and so we have plotted only the magnitude, even though the wave function itself oscillates rather rapidly. Note that the simple Hjalmarson theory is in excellent agreement with the data, for distances R out to the sixth nearest-neighbor shell. Beyond this distance, the effective-mass wave function (which fails badly for small R) describes the data well—indicating that the present theory would have been in even more dramatic agreement with the data if a Coulomb tail, $-e^2/\varepsilon r$ for $r \geq a_c$, had been added to the defect potential. This success is by no means trivially obtained, since some theories predict quite inaccurate wave functions for the S^- deep level [31, 33].

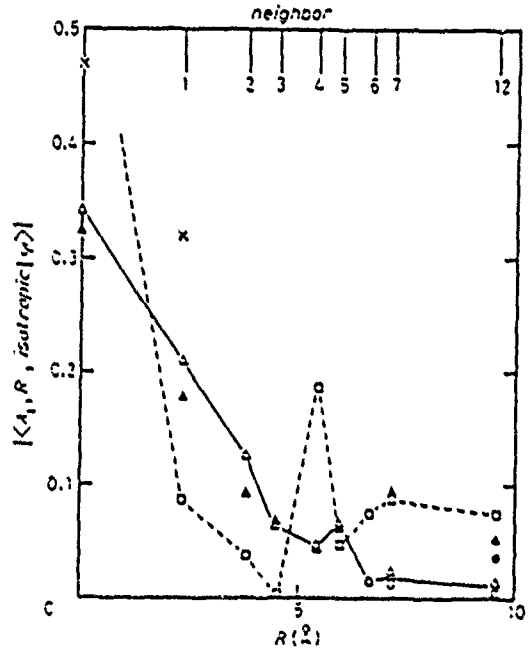


Fig. 7. – The magnitude of the isotropic part of the wave function of a S^- impurity in Si, as a function of the distance R (in Å) from the impurity site, after ref. [31]. The solid triangles and circles are derived from ENDOR data of ref. [32]; the open triangles and circles are the calculations of ref. [31]. Effective-mass theory is denoted by the dashed line and open squares. The ×'s are the theory discussed in ref. [33].

5·2. *Deep levels in* II-VI *compound semiconductors.* – The sort of predictions the theory produces are illustrated in fig. 8, where the levels associated with column V defects on a S site in CdS are given. These impurities, which one

might naively expect to be shallow acceptors, are, with the exception of N, deep levels predicted to lie well within the band gap of CdS. This is undoubtedly one reason (but not the only reason) why CdS and many other II-VI materials cannot be doped *p*-type—the expected shallow acceptors are, in fact, deep. The exception to this rule is N, which is predicted to yield no deep levels in the gap and to be a shallow acceptor level. Interestingly enough, in ZnSe (another II-VI host), WU *et al.* [34] have ion-implanted N, to find that it does produce a shallow acceptor level—as predicted by KOBAYASHI *et al.* [35].

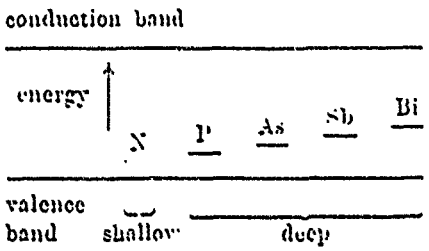


Fig. 8. – Theoretical predictions of ref. [35] showing that the expected standard *p*-type shallow dopants, except N, on the S site in CdS produce deep levels in the gap, after ref. [3].

5·3. *Sankey's theory of paired defects.* – SANKEY *et al.* [36] have extended the theory of deep levels associated with substitutional point defects to pairs of sp^3-bonded substitutional defects—with physically transparent results. The paired defect is a «molecule» that has «molecular orbitals» corresponding to π-like (or c_{2v}-symmetric) states and σ-like (or a_1-symmetric) states (see fig. 9). The relevant (near) π-like molecular orbitals are composed of T_2 single-defect orbitals polarized perpendicular to the spine of the molecule and have the same energies as the single-defect *p*-like T_2 states. The two *p*-like T_2 states of the single defect that are polarized parallel to the spine of the molecule hybridize

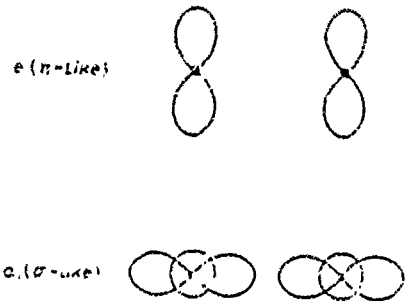


Fig. 9. – Schematic illustration of the π-like (c_{2v}-symmetric) molecular orbitals and the σ-like a_1 states of a defect pair. The dots indicate defects, circles denote s-states, and protuberances denote p-states, after ref. [3].

with the two s-like A_1 states centered on the different defects to form σ-like c-symmetric molecular states. In contrast with the π-like states, these hybridizing σ-like states are significantly perturbed by the presence of the second defect. Their energies, however, obey an approximate interlacing theorem, which requires that the «molecular» levels interlace the «atomic» A_1 and T_2 levels. (Recall that the p-like T_2 isolated-defect level decomposes into a σ-like a_1 molecular level polarized along the molecular spine plus two π-like e-levels. Hence both A_1 and T_2 isolated-defect levels produce the same molecular a_1-symmetry σ-like states of the pair. The interlacing theorem states that levels of the same symmetry, namely a_1, when perturbed, do not cross the unperturbed levels.) Because of this interlacing theorem, it is often possible to estimate the energies of the paired-defect levels relative to the isolated-defect levels to within a few tenths of an eV—without executing a calculation.

SANKEY has developed these ideas and applied them to the nearest-neighbor (spectator, oxygen) pairs in GaP (fig. 10). He has shown that the isolated oxygen

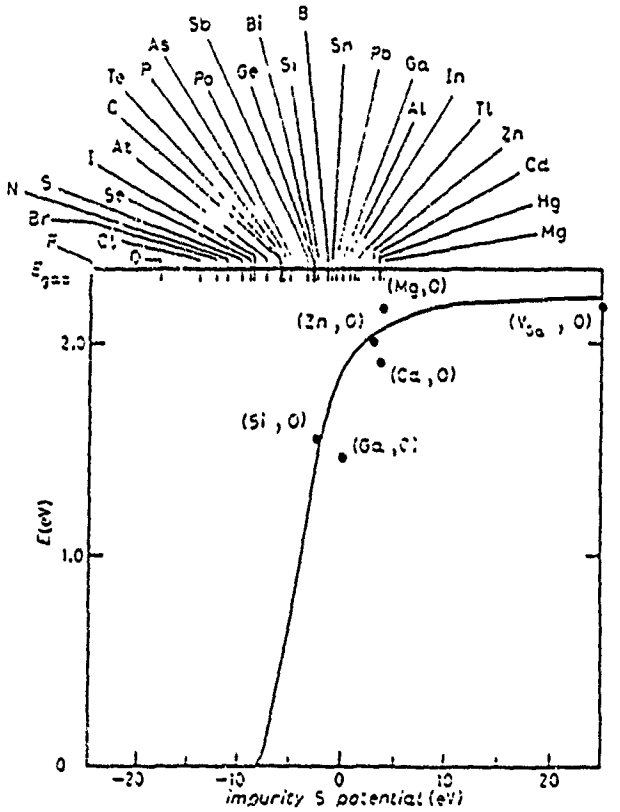


Fig. 10. – Trends in the energy levels for nearest-neighbor paired complexes of a spectator impurity on a Ga site and an O defect on a P site in GaP, after ref. [36]. The dots are data, and the spectator atoms label the abscissa.

level (*i.e.* the (Ga, O) pair) cannot be driven into the conduction band by any electropositive defect on an adjacent cation site, but that it can be driven into the valence band by any one of the following impurities on a neighboring Ga site: F, O, Cl, Br, N, S, Se, or I.

5·4. *Surface defects and Schottky-barrier heights.*

5·4.1. Core excitons at surfaces. Some of the best evidence supporting the theory comes from core exciton experiments, because, by the optical alchemy approximation, or $Z+1$ rule [37], a core exciton is identical to an impurity atom [38]: for example, core-excited Ga is Ga plus a core hole plus an electron and (because the core hole has almost the same charge distribution as a proton) is virtually identical to unexcited Ge, the atom immediately to

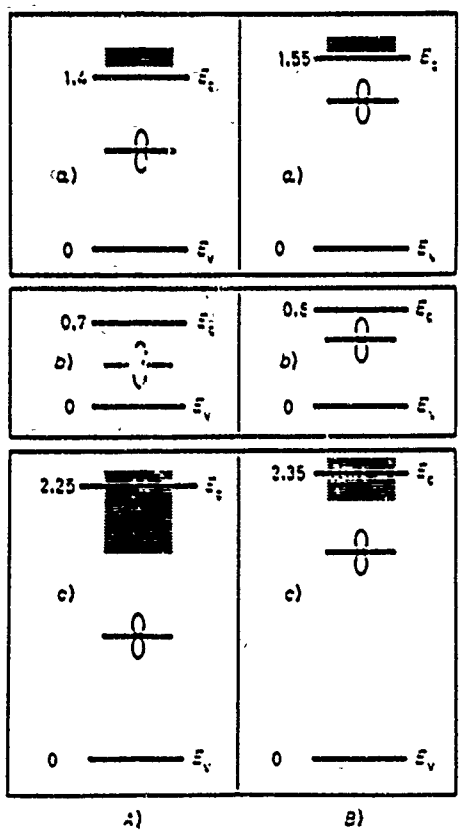


Fig. 11. – Comparison of *A*) experiment and *B*) theory for Ga site (110) surface core excitons in *a*) GaAs, *b*) GaSb and *c*) GaP, after ref. [39]. E_v and E_c denote valence and conduction band edges. The propeller denotes the core exciton level. The horizontal lines denote the lower portion of the intrinsic surface state bands. The theoretical band gaps are appropriate for 4 K, and hence are larger than the experimental gaps obtained at room temperature.

its right in the periodic table. Thus core-excited Ga is a Ge impurity and core-excited In is Sn. Figures 11 and 12 show that the predicted spectra for core-excited Ga at the (110) surface of Ga-group V compounds and for excited In at the surface of In-group V semiconductors account for the data for core

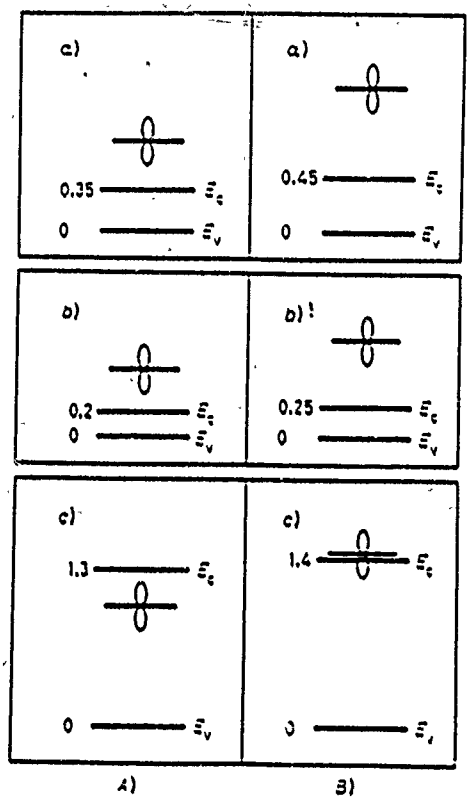


Fig. 12. – Comparison of *A*) experiment and *B*) theory for the In site (110) surface core excitons in *a*) InAs, *b*) InSb and *c*) InP, after ref. [39].

excitons at the relaxed (110) surfaces of these materials [39]. The theory also predicts a transition from shallow effective-mass exciton behavior to deep exciton behavior for the Si $2p$ core exciton in Si_xGe_{1-x} alloys. Evidence of this has been reported very recently by BUNKER *et al.* [38].

5·4.2. Defects at surfaces. An impurity at a surface has energy levels very similar to those of an (impurity, vacancy) pair, because the surface (in a nearest-neighbor tight-binding model) can be created by inserting a sheet of vacancies into the bulk, so that the impurity and the vacancy next to it form a pair whose energy levels are only slightly perturbed by the more distant vacancies of the sheet. To be sure, one must account for lattice relaxation at

the surface; nevertheless, the essential qualitative physics of defects at surfaces is the same as for (defect, vacancy) pairs. Therefore, the deep levels of surface defects can be computed using eq. (14), and the basic ideas used for deep levels associated with substitutional defects carry over to the surface defect problem, where the secular equation is formally the same as eq. (14), but its evaluation is considerably more complicated due to the reduced symmetry of the defect.

One point that should be emphasized is that a surface is a large perturbation, and the *deep levels associated with a surface impurity are likely to lie several tenths of an eV distant from the corresponding bulk impurity levels*. Indeed the number of deep levels bound in the gap may be different for a surface defect from the number for the same defect in the bulk. In particular, impurities that are «shallow» in the bulk often produce one, or even two, deep levels in the gap when they reside at the surface. For example, a nearest-neighbor pair of P impurities in Si is predicted to produce a deep level, although isolated P is a classic shallow donor [36].

5'4.3. Schottky-barrier heights.

5'4.3.1. Bardeen's model of Fermi-level pinning. In 1947 BARDEEN [40] proposed that the Schottky barriers that occur at metal/semiconductor interfaces are due to Fermi-level pinning by states at the interface. Stated simply

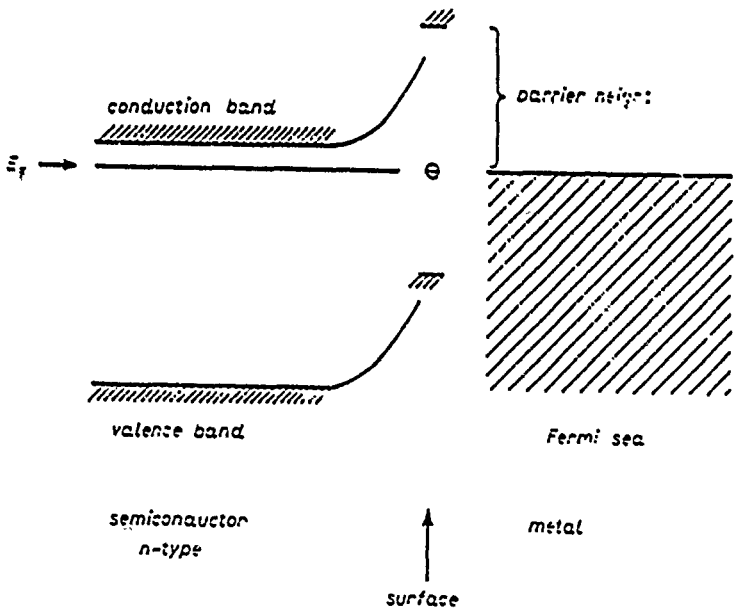


Fig. 13. – Schematic illustration of band bending and Schottky-barrier formation in the Bardeen model, after ref. [3].

for a degenerate *n*-type semiconductor, the Fermi levels of the bulk semiconductor, the metal and the semiconductor surface align in electronic equilibrium. At zero temperature, the Fermi level of the semiconductor surface is the level into which the next electron falls, and, if this is a bound deep level in the gap associated with a surface defect, charge will flow and the host energy bands will bend until this level aligns with the Fermi level of the bulk semiconductor (fig. 13). This causes a Schottky barrier to form.

5'4.3.2. Spicer's native-defect model. SPICER and co-workers [41] have championed the notion that the Bardeen states responsible for Fermi-level pinning of III-V semiconductors are deep levels associated with native surface defects. In this model, the Schottky-barrier height for *n*-type material is the binding energy of the surface deep level with respect to the conduction band edge (see fig. 13).

5'4.3.3. Allen's theory of Schottky-barrier heights in III-V semiconductors. ALLEN *et al.* [42, 43] have calculated the binding energies of deep levels produced by various defects, native antisite defects in particular, at the (110) surfaces of III-V compounds and ternary III-V alloys. This approach followed an earlier suggestion by DAW *et al.* [44] that Fermi-level pinning by deep levels associated with surface vacancies might account for many Schottky-barrier

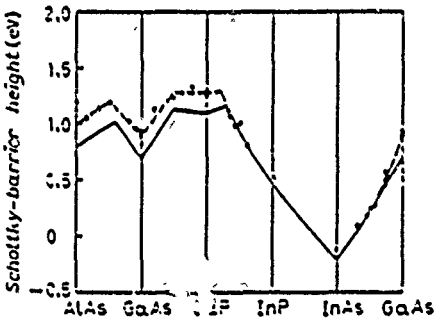


Fig. 14. – Predicted and observed Schottky-barrier heights in III-V alloys *vs.* alloy composition, after ref. [46]: · · · theory (GaAs), ——— experiment (Au). The theory assumes Fermi-level pinning by a cation on an anion site at the Au/semiconductor contact.

height data. The results of Allen's calculations for deep levels associated with surface cation-on-anion-site impurities [45, 46] are given in fig. 14, where they are in remarkable agreement with the data. This simple theory, which has been discussed in detail elsewhere [45], is capable of explaining numerous once-puzzling experimental facts. Fermi-level pinning by antisite defects accounts for the Schottky barriers between nonreactive metals and most III-V

semiconductors. Dangling bonds or extrinsic defects become important when reactive metals or special surface preparations are involved [43]. SANKEY *et al.* [47] have shown recently that Bardeen's ideas can be applied to Si as well as to the III-V's: Schottky-barrier data for Si/transition metal silicide interfaces can be understood in terms of Fermi-level pinning by interfacial dangling bonds. This work unifies the understanding of Schottky-barrier heights of III-V semiconductors and Si and also explains why free-surface calculations give good estimates of the Fermi-level pinning positions of interfacial defects. Hence it appears likely that the Schottky barriers on the common semiconductors have heights determined by surface deep levels (most associated with various native defects).

5'5. *Intrinsic surface states.* – In a nearest-neighbor tight-binding model, a simple way to create a surface is to insert a sheet of vacancies into a bulk semiconductor. The deep levels associated with this sheet defect are surface states. Therefore, the basic theoretical approach to the deep-level problem applies to predicting surface states as well. There are, of course, many technical problems associated with efficiently solving the secular equation (14) for surface states at relaxed and reconstructed surfaces. Those difficulties are beyond the

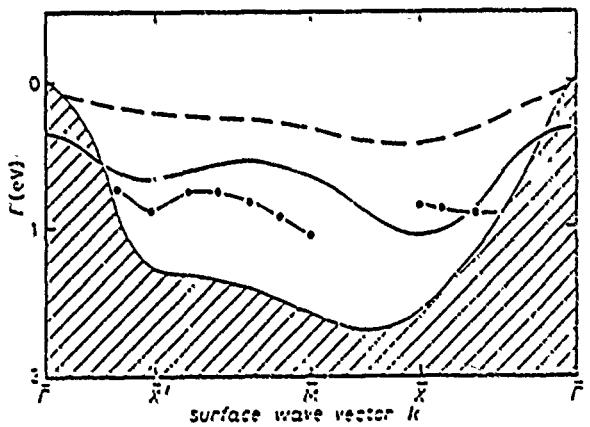


Fig. 15. – Surface state dispersion relations as measured [49] and predicted [48-50] for InP, after ref. [49] and [3]: – – – self-consistent pseudopotential. —— tight binding. • data.

scope of the present work [45]. Nevertheless, calculations of intrinsic surface state dispersion curves have been executed based on this model [48]. Typical results are given in fig. 15 for the relaxed (110) surface of InP [49]. As with the predictions of deep levels associated with point defects, the surface state predictions are comparably accurate with the best local-density theories [49, 50].

6. – Metastable alloys.

The ideas developed for treating band structures and impurity levels can be applied to treating large concentrations of «impurities» and to predicting phase transitions in alloys. Recently GREENE *et al.* [51] have grown metastable, sustitutional, crystalline $(GaAs)_{1-x}Ge_{2x}$ and $(GaSb)_{1-x}Ge_{2x}$ alloys, even though the constituents, GaAs and Ge or GaSb and Ge, are immiscible at equilibrium. The $(GaAs)_{1-x}Ge_{2x}$ alloys exhibit a V-shaped bowing of the fundamental band gap as a function of alloy composition x (fig. 16). This bowing cannot be

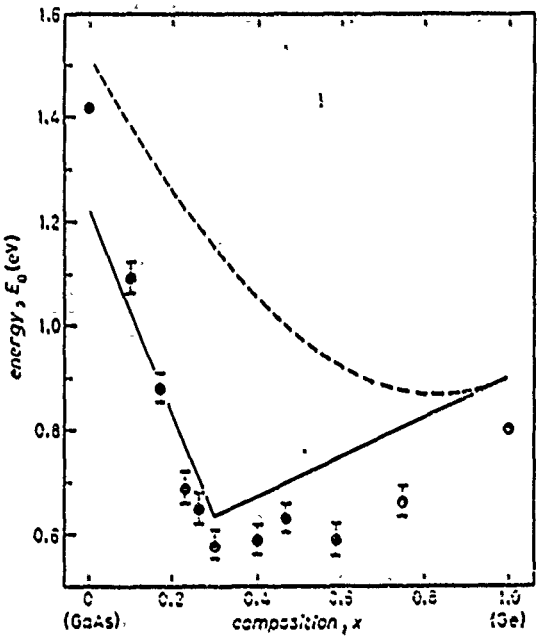


Fig. 16. – Direct energy gap of $(GaAs)_{1-x}Ge_{2x}$ alloys *vs.* alloy composition x, after ref. [53]. The dashed line represents the ordinary virtual-crystal approximation; the solid line represents the theory of ref. [52, 53]. The data are indicated by circles. The theory uses parameters appropriate to 4 K, whereas the data are for room temperature (at which the band gaps are different).

explained by the conventional virtual-crystal approximation, which assumes that each alloy is a crystal whose tight-binding parameters are interpolated between those of GaAs and Ge. NEWMAN *et al.* [52, 53] have shown that the V-shaped bowing is due to an order-disorder transition between a zincblende and a diamond phase (see fig. 17).

In modeling the phase transition in these alloys, NEWMAN was faced with data that required a reasonably accurate theory of electronic structure valid

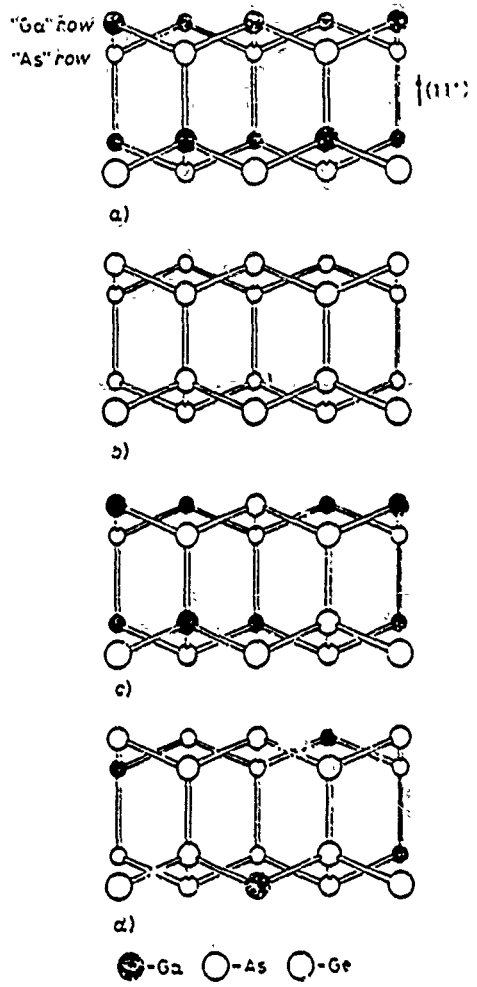


Fig. 17. – A schematic model of the phases of $(GaAs)_{1-x}Ge_{2x}$, after ref. [52, 53]: *a)* the GaAs (ordered) zincblende phase, *b)* the Ge (disordered) diamond phase, *c)* the GaAs-rich ordered zincblende phase of the alloy and *d)* the Ge-rich disordered diamond phase of the alloy. In the ordered phase, the great majority of Ga atoms occupy nominal Ga sites, but in the disordered phase Ga atoms show no site preference.

in both the ordered zincblende and the disordered diamond phases. Yet few theories are capable of accurately treating both a phase transition and electronic structure. NEWMAN circumvented this problem by adopting a two-Hamiltonian model.

The first Hamiltonian involves a pseudospin formalism, in which occupation of a site R by Ga, As, or Ge is represented by a pseudospin S_R that is «up» (+1), «down» (−1), or «zero» (0):

$$H = \sum_{R \neq R'} [J S_R S_{R'} - K S_R^2 S_{R'}^2] - \sum_R [h S_R + \Delta S_R^2] .$$

Here J, K, h and Γ are parameters defined in ref. [52], and R and R' are nearest-neighbor sites. In this pseudospin language, crystalline GaAs is a « pseudoantiferromagnet », because alternating sites are occupied by Ga (« up ») and As (« down »); the Ge atoms (« spin-zero ») dilute the « antiferromagnetism ». The « spin-spin » interaction between nearest neighbors causes As atoms to preferentially surround Ga and discourages the formation of clusters of only Ga or only As. Details of the Hamiltonian, which is similar to the Blume, Emery, Griffiths Hamiltonian [54] for ^{3}He-^{4}He solutions, are given in ref. [52]. NEWMAN solved this Hamiltonian, in a mean-field approximation, to obtain the equation for the order parameter $M(x, x_c)$:

$$M/(1-x) = \tgh\,[M'(1-x_c)]\,.$$

Here x_c is the alloy composition at which the minimum band gap is observed, 0.3 for $(GaAs)_{1-x}Ge_{2x}$. The order parameter is related to the probability that a Ga atom will be found on a nominal cation site of an imagined zincblende lattice; if this probability is the same as that for finding Ga on a nominal anion site, then the crystal structure is « diamond » and not « zincblende ». That is, the alloy is « disordered ».

The second Hamiltonian is the Vogl empirical tight-binding theory, but with the alloy's parameters determined by a new virtual-crystal approximation that depends on the order parameter of the first Hamiltonian: The diagonal matrix elements are interpolated assuming, for example, that the average cation is $[(1-x+M)/2]\,Ga + [(1-x-M)/2]\,As + [x]\,Ge$. The energy band structure computed using this model yields a direct band gap $E_0(x)$ *vs.* alloy composition x (fig. 16) and explains the observed V-shaped bowing.

7. – Summary.

In summary, the simple ideas originating in the work of Hsu, Wolford and Streetman and quantified in both the Vogl empirical tight-binding scheme and the Hjalmarson theory of defects have proven to have widespread applicability to defect states, interface states, surface states, Schottky barriers and alloy theory. The most reassuring feature of the theory is that it is simple enough to be used by a nonexpert, and yet it produces rather good predictions for the « deep » electronic states associated with almost any localized perturbation in a semiconductor.

* * *

We are grateful to the U. S. Office of Naval Research and the U. S. Army Research Office for their generous support of this research (Contract Nos. N00014-77-C-0537, N00014-81-K-0357 and DAAG29-83-K-0122). We have

benefitted greatly from many stimulating conversations with colleagues and students throughout the period of this research. The author apologizes to the many authors whose work he has not cited adequately; his sole excuse for this omission is a desire to present lectures from a uniform viewpoint and notation. He would like to thank his co-authors cited below for teaching him so much. In the preparation of this manuscript, he benefitted from conversations with D. FROELICH, D. JENKINS, G. KIM, K. NEWMAN, J. SHEN and D. WOLFORD.

REFERENCES

[1] D. J. WOLFORD and B. G. STREETMAN: *Phys. Rev. Lett.*, 36, 1400 (1976); D. J. WOLFORD: Ph.D. Thesis, University of Illinois (1979). The actual binding energy of the S impurity in GaP is 107 meV, larger than the 85 meV effective mass value due to valley-orbit splitting.
[2] D. J. WOLFORD, W. Y. HSU, J. D. DOW and B. G. STREETMAN: *J. Lumin.*, 18/19, 863 (1979).
[3] J. D. DOW: figures to be published.
[4] The virtual-crystal approximation treats the alloy as a crystal whose anion is the average anion $(1-x)\,As + xP$. The off-diagonal matrix elements of a tight-binding model are obtained by linearly interpolating Vd^2, where d is the bond length obtained by linear interpolation.
[5] W. Y. HSU, J. D. DOW, D. J. WOLFORD and B. G. STREETMAN: *Phys. Rev. B*, 16, 1597 (1977).
[6] P. VOGL, H. P. HJALMARSON and J. D. DOW: *J. Phys. Chem. Solids*, 44, 365 (1983). Table I of this reference should be corrected as follows: For ZnTe, we have $E(s, c) = -0.9350$ and $\Gamma(pa, s^*a) = 1.2381$. For SiC we have $\Gamma(pa, sc) = 4.2218$ $(= \Gamma(sc, pa))$ and $\Gamma(pa, s^*c) = 3.9605$. The 4's in eq. (9) should be 2's.
[7] W. A. HARRISON: *Electronic Structure and the Properties of Solids* (San Francisco, Cal., 1980).
[8] W. A. HARRISON: in *Festkörperprobleme*, Vol. 17, edited by J. TREUSCH (Wiesbaden, 1977), p. 135.
[9] W. KOHN: *Solid State Phys.*, 5, 257 (1957).
[10] F. BASSANI and P. PARRAVICINI: *Electron States and Optical Transition in Solids* (Elmsford, 1975).
[11] H. P. HJALMARSON, P. VOGL, D. J. WOLFORD and J. D. DOW: *Phys. Rev. Lett.*, 44, 810 (1980), and to be published.
[12] H. P. HJALMARSON: Ph. D. Thesis, University of Illinois (1979) (unpublished).
[13] O. F. SANKEY and J. D. DOW: *Phys. Rev. B*, 27, 7641 (1983).
[14] L. A. HEMSTREET: *Phys. Rev. B*, 15, 834 (1977); P. VOGL: to be published.
[15] M. SCHEFFLER, J. P. VIGNERON and G. B. BACHELET: *Phys. Rev. Lett.*, 49, 1765 (1982).
[16] J. BERNHOLC, N. O. LIPARI, S. T. PANTELIDES and M. SCHEFFLER: *Phys. Rev. B*, 26, 5706 (1982).
[17] S. LEE, J. D. DOW and O. F. SANKEY: *Phys. Rev. B*, to be published.
[18] M. LANNOO and P. LENGLART: *J. Phys. Chem. Solids*, 30, 2409 (1969).
[19] E. KAUFER, P. PECHER and M. GERL: *J. Phys. C*, 9, 2319 (1976).

[20] J. D. Joannopoulos and E. J. Mele: *Solid State Commun.*, 20, 729 (1976).
[21] M. Jaros and S. Brand: *Phys. Rev. B*, 14, 4494 (1976).
[22] U. Lindefelt: *J. Phys. C*, 12, L419 (1979); D. A. Papaconstantopoulos and E. N. Economou: *Phys. Rev. B*, 24, 2993 (1980).
[23] G. A. Baraff and M. Schlüter: *Phys. Rev. Lett.*, 41, 892 (1978); *Phys. Rev. B*, 19, 4965 (1979).
[24] J. Bernholc and S. T. Pantelides: *Phys. Rev. B*, 18, 1780 (1978); J. Bernholc, N. O. Lipari and S. T. Pantelides: *Phys. Rev. Lett.*, 41, 895 (1978); *Phys. Rev. B*, 21, 3545 (1980); J. Bernholc, S. T. Pantelides, N. O. Lipari and A. Baldereschi: *Solid State Commun.*, 37, 705 (1981).
[25] C. A. Swarts, D. L. Miller, D. R. Franceschetti, H. P. Hjalmarson, P. Vogl, J. D. Dow, D. J. Wolford and B. G. Streetman: *Phys. Rev. B*, 21, 1708 (1980).
[26] W. A. Harrison: *Phys. Rev. B*, 8, 4487 (1973); D. J. Chadi and M. L. Cohen: *Phys. Status Solidi B*, 68, 405 (1975); K. C. Pandey and J. C. Phillips: *Phys. Rev. B*, 13, 750 (1976); D. J. Chadi: *Phys. Rev. B*, 16, 790 (1977).
[27] G. F. Koster and J. C. Slater: *Phys. Rev.*, 96, 1208 (1954).
[28] For example, the density of states of the perturbed Hamiltonian is

$$D(E') = -\pi^{-1}\,\mathrm{Im}\,\mathrm{Tr}G(E')$$

and the density of states of the host is

$$D^0(E') = -\pi^{-1}\,\mathrm{Im}\,\mathrm{Tr}G^0(E') = -\pi\sum_{kl}\delta(E'-E_{kl})\,.$$

[29] G. Lehmann and M. Taut: *Phys. Status Solidi B*, 54, 469 (1972); J. Rath and A. J. Freeman: *Phys. Rev. B*, 11, 2109 (1975).
[30] D. J. Chadi and M. L. Cohen: *Phys. Rev. B*, 8, 5747 (1973).
[31] S. Y. Ren, W. M. Hu, O. F. Sankey and J. D. Dow: *Phys. Rev. B*, 26, 951 (1982). See also S. Y. Ren: *Sci. Sin.*, 27, 443 (1984).
[32] G. W. Ludwig: *Phys. Rev.*, 137, A 1520 (1965).
[33] T. Shimizu and K. Minami: *Phys. Status Solidi B*, 48, K181 (1971).
[34] Z. L. Wu, J. J. Metz, C. J. Werkhoven, B. J. Patrick and R. N. Bhargava: *Appl. Phys. Lett.*, 40, 345 (1982).
[35] A. Kobayashi, O. F. Sankey and J. D. Dow: *Phys. Rev. B*, 28, 946 (1983).
[36] O. F. Sankey, H. P. Hjalmarson, J. D. Dow, D. J. Wolford and B. G. Streetman: *Phys. Rev. Lett.*, 45, 1650 (1980); O. F. Sankey and J. D. Dow: *Appl. Phys. Lett.*, 38, 685 (1981); *J. Appl. Phys.*, 52, 5139 (1981); *Phys. Rev. B*, 26, 3243 (1982); *Solid State Commun.*, in press. See also J. D. Dow and O. F. Sankey: *Am. Inst. Phys. Conf. Proc.*, Vol. 73, edited by R. A. Street, D. K. Biegelsen and J. C. Knights (New York, N. Y. 1981), p. 141.
[37] See, for example, J. D. Dow, D. R. Franceschetti, P. C. Gibbons and S. E. Schnatterly: *J. Phys. F*, 5, L211 (1975).
[38] H. P. Hjalmarson, H. Büttner and J. D. Dow: *Phys. Rev. B*, 24, 6010 (1981). Experimental evidence supporting this theory has recently been reported by B. A. Bunker, S. L. Hulbert, J. P. Stott and F. C. Brown: *Phys. Rev. Lett.*, 53, 2157 (1984).
[39] R. E. Allen and J. D. Dow: *Phys. Rev. B*, 24, 911 (1981).
[40] J. Bardeen: *Phys. Rev.*, 71, [illegible].
[41] W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su and I. Lindau: *J. Vac. Sci. Technol.*, 16, 1422 (1979), and references therein; W. E. Spicer, I. Lindau, P. R. Skeath and C. Y. Su: *J. Vac. Sci. Technol.*, 17, 1019 (1980), and references therein.

[42] R. E. ALLEN and J. D. DOW: *Phys. Rev. B*, 25, 1423 (1982).
[43] J. D. DOW and R. E. ALLEN: *J. Vac. Sci. Technol.*, 20, 659 (1982). See also ref. [45] and references therein.
[44] M. S. DAW and D. L. SMITH: *Phys. Rev. B*, 20, 5150 (1979); *J. Vac. Sci. Technol.*, 17, 1028 (1980); *Appl. Phys. Lett.*, 36, 690 (1980); *Solid State Commun.*, 37, 205 (1981); M. S. DAW, D. L. SMITH, C. A. SWARTS and T. C. MCGILL: *J. Vac. Sci. Technol.*, 19, 508 (1981).
[45] For a review of the theory of intrinsic and extrinsic surface states, see J. D. DOW, R. E. ALLEN and O. F. SANKEY: *Chemistry and Physics on Solid Surfaces*, Vol. 5, edited by R. VANSELOW and R. HOWE, Springer Series in Chemical Physics, 35 (Berlin, Heidelberg, New York, N.Y., 1984).
[46] R. E. ALLEN, T. J. HUMPHREYS, J. D. DOW and O. F. SANKEY: *J. Vac. Sci. Technol. B*, 2, 449 (1984).
[47] O. F. SANKEY, R. E. ALLEN and J. D. DOW: *Solid State Commun.*, 49, 1 (1984).
[48] R. P. BERENS, R. E. ALLEN and J. D. DOW: *Phys. Rev. B*, 26, 5702 (1982).
[49] G. P. SRIVASTAVA, I. SINGH, V. MONTGOMERY and R. H. WILLIAMS: *J. Phys. C*, 16, 3627 (1983).
[50] F. MANGHI, E. MOLINARI, C. M. BERTONI and C. CALANDRA: *J. Phys. C*, 15, 1099 (1982).
[51] J. E. GREENE: *J. Vac. Sci. Technol. B*, 1, 229 (1983), and references therein.
[52] K. E. NEWMAN and J. D. DOW: *Phys. Rev. B*, 27, 7495 (1983).
[53] K. E. NEWMAN, A. LASTRAS-MARTINEZ, B. KRAMER, S. A. BARNETT, M. A. RAY, J. D. DOW, J. E. GREENE and P. M. RACCAH: *Phys. Rev. Lett.*, 50, 1466 (1983).
[54] M. BLUME, V. J. EMERY and R. B. GRIFFITHS: *Phys. Rev. A*, 4, 1071 (1971).

Reprinted From
Highlights of
Condensed-Matter Theory
© 1985, LXXXIX Corso
Soc. Italiana di Fisica - Bologna - Italy

PHYSICAL REVIEW B VOLUME 32, NUMBER 8 15 OCTOBER 1985

X-ray spectra of model binary alloys $A_{1-x}B_x$

Marshall A. Bowen
Department of Physics, Western Illinois University, Macomb, Illinois 61455

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 14 January 1985)

The x-ray photoemission, absorption, and emission spectra of a one-dimensional tight-binding model for a binary metallic alloy $A_{1-x}B_x$ are evaluated in a change of mean-field model. The combined effects of disorder and multielectron excitations are included. The extent of the asymmetric tails of the x-ray photoemission lines depend on the local character of the Fermi-energy states in the vicinity of the core hole; this effect could explain the long-standing mystery of why Na 1s lines in Na_xWO_3 are symmetric. Features in the absorption and emission spectra reminiscent of the anomalous ramplike thresholds observed for absorption by rare-gas atoms in alkali-metal hosts are also found.

I. INTRODUCTION

As a first step toward understanding the x-ray absorption, emission, and photoemission spectra of binary alloys, we present here the results of model calculations for a one-dimensional, substitutional, crystalline, binary alloy $A_{1-x}B_x$. The model treats a single orbital and a single electron per site in a nearest-neighbor tight-binding approximation. Many-electron effects due to the final-state interactions of the electrons with the core hole are treated in a change-of-mean-field approximation.[1] Hence, the model exhibits features associated with both the "x-ray edge anomalies"[2–4] and binary-alloy disorder.[5] To our knowledge, this is the first study of the combined effects on x-ray spectra of disorder in a binary-alloy and many-electron recoil.

II. MODEL

The one-electron Hamiltonian governing the behavior of the alloy is

$$h=\sum_{n=1}^{M}\epsilon(n)|n)(n|+\beta|n)(n+1|+\beta|n+1)(n|\ . \quad (1)$$

Here we have M sites, $|n)$ refers to the one-electron orbital centered at the nth site, β is the nearest-neighbor transfer matrix element, and $\epsilon(n)$ is a random variable which takes on the values ϵ_A with probability $1-x$, and ϵ_B (with probability x).

In the case of x-ray photoemission the band has N electrons in both the initial and final states, but the initial-state N-electron Hamiltonian

$$H_I=\sum_{i=1}^{N}h_i \quad (2)$$

changes suddenly as a result of the removal of the core electron to the final-state Hamiltonian

$$H_F=\sum_{i=1}^{N}h_i' \ , \quad (3)$$

where we have

$$h'=h+V_0|\mathbf{R})(\mathbf{R}| \ , \quad (4)$$

with $|\mathbf{R})$ referring to the orbital centered on the core-hole site; V_0 is the electron-core interaction strength (and a negative number). The initial many-electron state of the electron gas $|I\rangle$, in this model, is a Slater determinant of the lowest-energy single-particle eigenstates $|\phi)$ of h; the final states $|Fv\rangle$ are all the various determinants of the eigenstates $|\psi)$ of h'.

The x-ray photoemission spectrum is

$$I(E)=\sum_{v}|\langle I|Fv\rangle|^2\delta(E+E_{Fv}-E_I-\hbar\omega-\epsilon_{core}) \ , \quad (5)$$

where the summation is over all possible final-state configurations. The photoemission recoil energy is

$$E_{Fv}-E_I\equiv\sum_{i}^{N}\epsilon_i'-\sum_{i}^{N}\epsilon_i \ , \quad (6)$$

where the sums are over all occupied one-electron states in the electronic configurations $|Fv\rangle$ and $|I\rangle$, respectively. The photoemission line shape has contributions from both spin-up and spin-down channels. It can be shown, however, that one can evaluate the line shapes for each of these channels independently and that the two-channel line shape is a convolution of the single-channel shapes.[1] Hence, for simplicity of presentation we consider here only the spin-up channel, and we have $M=2N$.

X-ray emission of a photon of energy E can be treated in a manner completely analogous to photoemission, and has a line shape

 ©1985 The American Physical Society

$$\kappa(E)=\sum_{\nu}|\langle F\nu|M|I\rangle|^2\delta(E=E_I+E_{F\nu})\,, \tag{7}$$

where the final states $|F\nu\rangle$ are Slater determinants of N $|\phi\rangle$'s and one core orbital and the initial state is a determinant of $N+1$ $|\psi\rangle$'s. A similar expression holds for x-ray absorption.[6] We assume that the core hole has a negligible radius, in which case the dipole matrix element M can be simplified, and we have

$$\langle F\nu|M|I\rangle=M_0\begin{vmatrix}\psi_1(\mathbf{R}) & \psi_2(\mathbf{R}) & \cdots & \psi_{N+1}(\mathbf{R})\\ (\phi_{\nu,1}|\psi_1) & & & \\ (\phi_{\nu,2}|\psi_1) & & & \\ \cdots & & & \\ (\phi_{\nu,N}|\psi_1) & \cdots & & (\phi_{\nu,N}|\psi_{N+1})\end{vmatrix}\,. \tag{8}$$

Here $\psi(\mathbf{R})$ is proportional to the overlap of $|\psi\rangle$ with the core hole at site $\mathbf{R}$, M_0 is a constant, $(\phi|\psi)$ is a scalar product, and we have assumed that the core hole has a negligible radius.[6]

The line shapes $I(E)$ and $\chi(E)\equiv\kappa(E)/M_0^2$ are calculated as follows: The eigenstates $|\phi\rangle$ and $|\psi\rangle$ and the corresponding eigenvalues are evaluated for a one-dimensional lattice with $M=40$ sites occupied by a specific configuration of atoms A and B, as determined by a random-number generator. The core hole is confined to one of the ten innermost sites. The matrix elements between determinants, such as $\langle I|F\nu\rangle$, are evaluated for many electronic configurations ν and the spectra are calculated. The calculation of a given spectrum is terminated (i.e., no more configurations ν are included) when the sum rules for x-ray photoemission spectra[1] (XPS)

$$\int_{-\infty}^{\infty}I(E)dE=1 \tag{9}$$

and for emission[6]

$$\int_{-\infty}^{\infty}\chi(E)dE=\sum_{i}|\psi_i(\mathbf{R})|^2 \tag{10}$$

are adequately exhausted (the sum is over occupied initial-state orbitals ψ_i). The calculations are repeated for (typically $\simeq$100) different atomic configurations and (typically $\simeq$10 per atomic configuration) different core-hole sites, Gaussian broadened[7] and ensemble averaged. The ensemble-averaged spectra are displayed in Figs. 1 to 6 for $\beta=-\frac{1}{2}$, $V_0=-2|\beta|$, $\epsilon_B=2|\beta|$, and $\epsilon_A=-2|\beta|$ for $x=0.2$, 0.4, 0.6, and 0.8.

III. RESULTS

The results can best be understood in terms of the broadened[7] densities of states displayed in Fig. 7. In all cases the Fermi surface lies within a band and the system is metallic.

A. X-ray photoemission spectra

For a core hole created at an A site (Fig. 1), the x-ray photoemission spectra exhibit long tails associated with low-energy excitation of Fermi-surface electrons for $x=0.2$ and 0.4, but not for $x=0.8$ or 0.6. The reason for this is that the one-electron states near the Fermi surface are A-like for $x=0.2$ and 0.4, but are B-like for $x=0.8$ or 0.6 (see Fig. 7). Only the A-like states are efficiently excited as the shock wave due to the A-site core-hole creation propagates outward. The B-like electron states at the Fermi energy for $x=0.6$ and 0.8 do not thoroughly overlap with and couple to the A-site core hole, and are not so easily excited as a result of the core-hole creation. Hence, the A-site XPS lines for $x=0.6$ and 0.8 do not have long tails for negative $E-\hbar\omega-\epsilon_{\text{core}}$, but the lines for $x=0.2$ and 0.4 do. (Similarly, creation of a B-site core hole produces a long XPS tail for $x>0.5$ but not for $x<0.5$, as shown in Fig. 2.) Thus we have a clear dependence of the shape of the XPS line on the char-

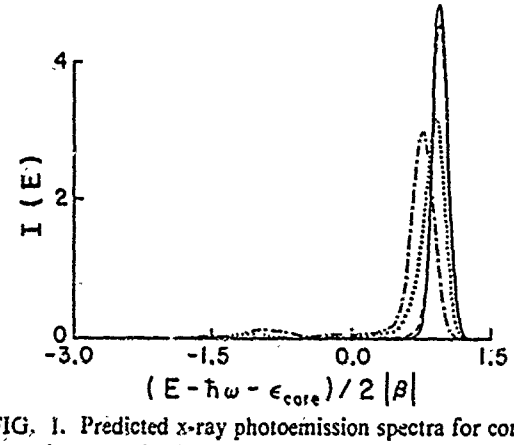


FIG. 1. Predicted x-ray photoemission spectra for core excitation of the A site in $A_{1-x}B_x$ as a function of the emitted electron's energy E, for $x=0.2$ (dash-dotted line), $x=0.4$ (dotted line), $x=0.6$ (dashed line), and $x=0.8$ (solid line).

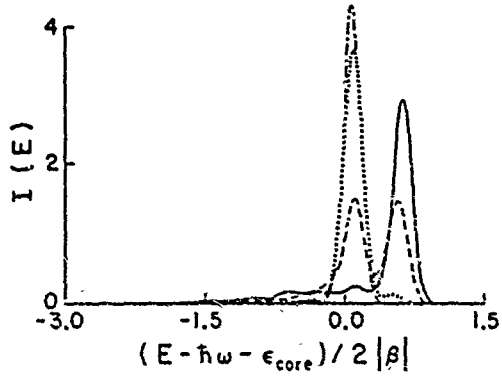


FIG. 2. Predicted B-site x-ray photoemission spectra, as in Fig. 1.

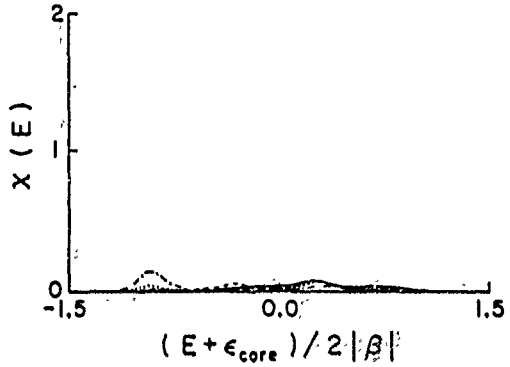


FIG. 3. Predicted x-ray absorption spectra $\chi(E)$ for excitation of a core level at the A site, with notation as in Fig. 1.

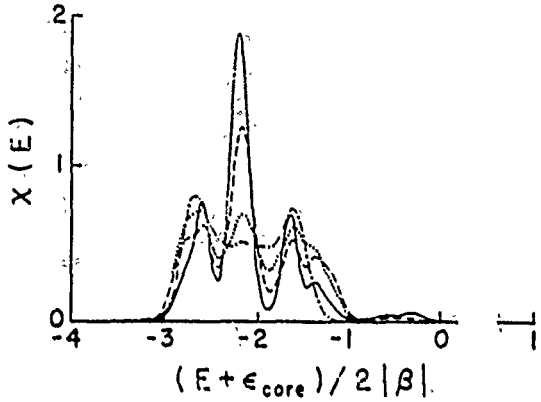


FIG. 5. Predicted A-site x-ray emission spectra.

acter of the Fermi-energy electrons at the site of the core hole.

The asymptotic theory of Doniach and Sunjic[4] for excitation of a free-electron gas is valid for electron energies E near the photoemission threshold energy E_T and gives an XPS line shape

$$I(E) \propto (E_T - E)^{-1+\Delta} \Theta(E_T - E) , \tag{11}$$

where Θ is the unit step function and the exponent Δ is expressible in terms of δ_l, the change of Fermi-energy phase shifts of an electron as a result of the potential of the core hole:

$$\Delta = \sum_{l=0}^{\infty} 2(2l+1)(\delta_l/\pi)^2 . \tag{12}$$

This asymptotic line shape does not depend on the character or density of states near the Fermi energy, except through the phase shifts δ_l. Nevertheless, our calculations, which solve a Doniach-Sunjic type of model for all energies (not just for $E \to E_T$), show that the extent of the XPS tail does indeed depend on the character of the Fermi-energy states.

This behavior may have been observed in sodium-tungsten bronzes: Campagna *et al.*[8] and Chazalviel *et al.*[9] have reported both an asymmetric W XPS line and an excessively symmetric Na 1s XPS line in Na_xWO_3; the latter cannot be explained by the asymptotic theory, [Eq. (11)].[10] It is noteworthy that in the simplest model the Na states do not contribute to the conduction band;[8,9,11] hence the Na-like character of the Fermi-energy states in Na_xWO_3 should be small—and by analogy with the present results we expect the Na XPS line to be quite symmetric. Hence, the present theory indicates that the large asymmetric tail predicted on the basis of the asymptotic Doniach-Sunjic theory should not necessarily be expected when the amplitude of the Fermi-energy one-electron states at the core-hole site is not large — because the electron-hole pair excitations of those states (which are responsible for the long tail) cannot be efficiently achieved.

A second interesting feature of the A-site XPS spectra is the small bump for $x = 0.2$ and 0.4 near $E - \hbar\omega - \epsilon_{core} \simeq -2|\beta|$ ($= |V_0|$ here), which we associate with transitions of the electron gas that leave a hole in the bound state below the A band. (This bound state always occurs in one dimension and is caused by the attractive electron-hole interaction; it lies of order $\simeq |V_0|$ below the band bottom.)

The XPS spectra at the B site are especially interesting. The lines for $x = 0.2$ and 0.4 are nearly symmetric because the Fermi-energy states are largely A-like and not efficiently excited by a B-site core hole. They are also almost recoilless (viz., at zero energy) because the B-like states that are perturbed by the core hole are unoccupied and therefore do not contribute to the recoil energy [Eq. (6)]. The $x = 0.4$ spectrum has, in addition to its recoilless peak, a weak high-energy peak associated with recoil: The on-site level at ϵ_B is pulled below the Fermi level by

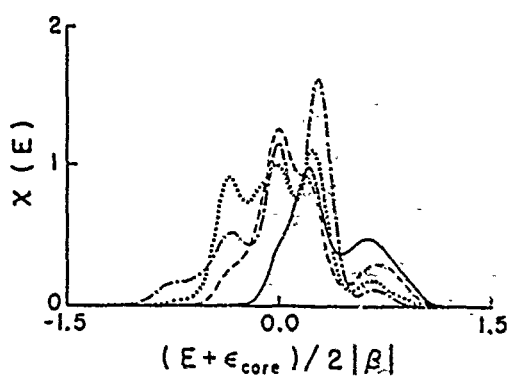


FIG. 4. Predicted B-site x-ray absorption spectra.

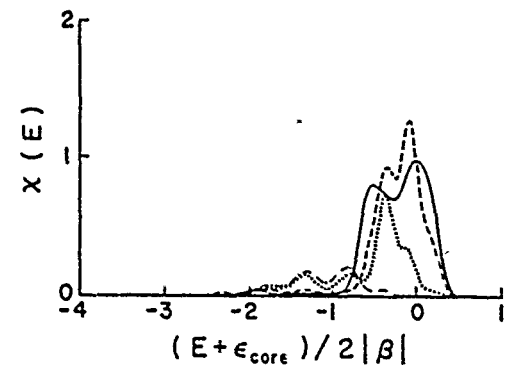


FIG. 6. Predicted B-site x-ray emission spectra.

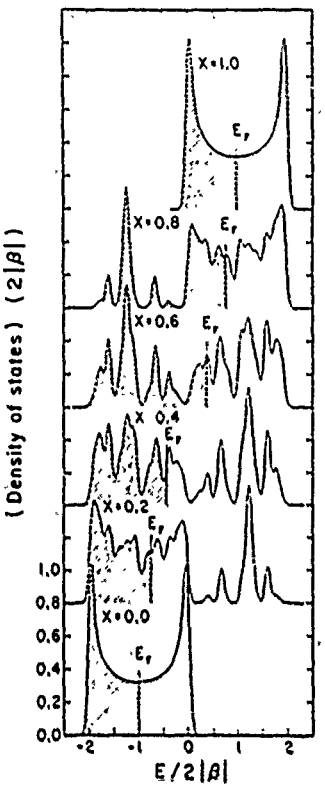


FIG. 7. One-electron density of states (times $2|\beta|$) of $A_{1-x}B_x$ vs E (in units of $2|\beta|$) for $x=0$, 0.2, 0.4, 0.6, 0.8, and 1.0. The Fermi energies are denoted by E_F and occupied one-electron states are shaded. Note that the one-electron states at the Fermi energy are A-like for $x<0.5$ and B-like for $x>0.5$.

the electron-hole interaction, becomes occupied, and contributes to the recoil energy. (This effect was predicted first for d states by Kotani and Toyozawa[12] and then, in a different context, by Mehreteab and Dow.[13]) As the alloy composition x and the Fermi energy increase, local alloy configurations which lead to recoil become more probable and the recoil peak grows as the recoilless peak decreases in amplitude. For $x\simeq 0.8$ the Fermi energy lies within $|V_0|$ of ϵ_B and the recoil peak is dominant. Bumps below the main peak are associated with the alloy disorder.

B. Absorption spectra

The absorption spectra for the core hole on an A site (Fig. 3) are generally weak because of the predominantly B-like character of the unoccupied one-electron states. The strongest spectrum is for $x=0.2$ and corresponds to a case in which there is a reasonable amount of A character to the final state.

For a core hole at a B site the absorption spectra (Fig. 4) are stronger because the B hole couples strongly to the B-like unoccupied electron states. Even for $x=0.2$ there is some B character to the Fermi-energy states, and the absorption edge (at the left of Fig. 4) becomes more abrupt as x and the B character of the Fermi-energy states increases. In general the spectra exhibit low-energy absorption edges and bumps at higher energy that are derived from the hole-perturbed bumps in the densities of states. [The van Hove singularity[14] at the band maximum is weakened both by the matrix element, Eq. (8), and the alloy disorder.]

An especially interesting feature of the calculated B-site spectrum for $x=0.2$ is its low-energy edge — which does not show the expected[2,3] peaked threshold behavior (predicted for free-electron metals):

$$\gamma(E)\propto(E-E_T)^{-\alpha_0}\Theta(E-E_T)\,, \tag{13}$$

where E_T is the threshold energy and α_0 is the x-ray edge exponent.[15] Instead, the threshold line shape is very nearly a linear function of energy, $\alpha_0\simeq -1$. Such behavior is what has been observed for rare-gas atoms in alkali-metal hosts by Flynn and co-workers,[16] and has remained a major unexplained anomaly for years.[17] The present work suggests that the "anomaly" may be a consequence of the non-B character of the Fermi-surface states.

C. Emission spectra

The emission spectra for an A-site core hole evolve in an interesting fashion as a function of composition x. For small x, $x=0.2$, the spectrum exhibits a low-energy peak associated with a band-bottom van Hove singularity[14] (see Figs. 5 and 7) that has been partially amputated by disorder; it also has a high-energy edge with a peak reminiscent of an x-ray edge anomaly [Eq. (13), with E and E_T reversed]. For $x=0.4$, 0.6, and 0.8, additional features associated with alloy disorder as manifested in the densities of states (Fig. 7) are reflected in the spectra. In addition, the $x=0.6$ and 0.8 spectra have weak high-energy edges (that are more or less ramplike functions of energy) because the Fermi-energy states are B-like and do not couple effectively to an A hole.

On the B site the spectra are dramatically different, showing bumps associated with the alloy disorder (shifted by the electron-hole interaction), high-energy x-ray edges [Eq. (13)], for $x=0.8$ and $x=0.6$ that weaken as the B character of the Fermi-energy states is lost with decreasing x. For $x=0.8$ there is a high-energy x-ray edge (near $E+\epsilon_{\text{core}}\simeq 0$), the remnants of a van Hove singularity in the density of states (near $E+\epsilon_{\text{core}}\simeq -|\beta|$), and a weak low-energy peak (near $E+\epsilon_{\text{core}}\simeq -4|\beta|$) associated with the density of states: B character is mixed into the A-like states by the alloy. The prominent x-ray edge occurs because the B hole efficiently excites the B-like Fermi-surface particle-hole excitations. For $x=0.6$ to $x=0.2$, qualitatively similar structures appear, most of which are peaks shifted by the electron-hole interaction, but associated with the disorder as reflected also in the densities of states. The strength of the emission weakens as the B character of the Fermi surface is lost (as x decreases). Also, the high-energy x-ray edge weakens and becomes ramplike for $x=0.2$.

D. Summary

In summary the predicted x-ray spectra of one-dimensional $A_{1-x}B_x$ substitutional alloys are rich in

features that are associated with alloy disorder, x-ray edge effects, and van Hove singularities. The effects of alloy disorder found here are probably more pronounced than one would find in three-dimensional alloys, owing to the lack of multiple paths circumventing any highly disordered region. Two particularly interesting features of the calculated spectra appear to hold promise for explaining some old mysteries. (i) The anomalously symmetric XPS lines of sodium-tungsten bronzes appear to be related to the fact that the character of the one-electron states at the Fermi energy is such that these states are not efficiently excited by a Na-site core hole, and (ii) the ramplike linear absorption thresholds of rare-gas atoms in alkali-metal hosts appear to be related to the fact that the Fermi-surface states at the rare-gas site do not have sufficient alkali-metal character. An interesting prediction of the model is that the emission spectrum from rare-gas atoms in alkali-metal hosts should also have ramplike thresholds rather than edge anomalies. It would be gratifying if this prediction were verified experimentally.

ACKNOWLEDGMENT

We are grateful to the U.S. Office of Naval Research for their generous support (under Contract No. N00014-84-K-0352).

[1]J. D. Dow and C. P. Flynn, J. Phys. C 13, 1341 (1980).
[2]J. Friedel, Comments Solid State Phys. 2, 21 (1969).
[3]P. Nozières and C. T. de Dominicis, Phys. Rev. 178 1097 (1969).
[4]S. Doniach and M. Sunjic, J. Phys. C 3, 285 (1970).
[5]P. Dean, Rev. Mod. Phys. 44, 127 (1972); Proc. Soc. London, Ser. A 260, 263 (1961); Proc. Phys. Soc. London 84, 727 (1964); D. N. Payton and W. M. Visscher, Phys. Rev. 154, 802 (1967); 156, 1032 (1967); 175, 1201 (1968).
[6]C. A. Swarts, J. D. Dow, and C. P. Flynn, Phys. Rev. Lett. 43, 158 (1979); C. A. Swarts and J. D. Dow (unpublished).
[7]The broadening function is $(2\pi\Gamma^2)^{-1/2}\exp(-x^2/2\Gamma^2)$, with $\Gamma=0.08|2\beta|$. For Fig. 7 we used $\Gamma=0.05|2\beta|$.
[8]M. Campagna, G. K. Wertheim, H. R. Shanks, F. Zumsteg, and E. Banks, Phys. Rev. Lett. 34, 738 (1975).
[9]J. N. Chazalviel, M. Campagna, G. K. Wertheim, and H. R. Shanks, Phys. Rev. B 16, 697 (1977).
[10]J. D. Dow, Philos. Mag. 35, 837 (1977).
[11]J. B. Goodenough, Prog. Solid State Chem. 5, 145 (1971); Bull. Soc. Chim. Fr. 1965, 1200 (1965); L. F. Mattheiss, Phys. Rev. B 6, 4718 (1972).
[12]A. Kotani and Y. Toyozawa, J. Phys. Soc. Jpn. 37, 912 (1974).
[13]E. Mehreteab and J. D. Dow, Phys. Rev. B 26, 2261 (1982); Solid State Commun. 43, 837 (1982).
[14]L. van Hove, Phys. Rev. 89, 1189 (1953).
[15]In terms of the Fermi-energy phase shift associated with the electron-hole interaction, we have $\alpha_0=2\delta_0/\pi-\delta_0^2/\pi^2$.
[16]R. A. Tilton, D. J. Phelps, and C. P. Flynn, Phys. Rev. Lett. 32, 1006 (1974); R. A. Tilton and C. P. Flynn, *ibid.* 34, 20 (1975); D. J. Phelps, R. Avci, and C. P. Flynn, *ibid.* 34, 23 (1975).
[17]M. A. Bowen and J. D. Dow, Nuovo Cimento 1D 587 (1982); M. A. Bowen, D. V. Froelich, and J. D. Dow, Phys. Lett. 102, 73 (1984).

Electronic structure of $Pb_{1-x}Sn_xTe$ semiconductor alloys

Seongbok Lee and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 20 April 1987)

The electronic structures of the pseudobinary alloy semiconductors $Pb_{1-x}Sn_xTe$ are analyzed using a tight-binding model with spin-orbit interaction. The densities of states and the band gaps at the L point are computed for both the effective media using the virtual-crystal approximation and the realistic media employing the recursion method, and the results are compared. Both theories exhibit alloying effects such as band broadening, energy shifts, and Dimmock's band-crossing phenomenon. However, significant deviations from the virtual-crystal approximation are found for the cation-derived s-like deep valence-band states.

I. INTRODUCTION

The narrow-gap IV-VI semiconductor compounds and their pseudobinary alloys have unique properties. They have on the average five valence electrons per atom, small direct band gaps at the L point, and high static dielectric constants of order 10^3. They often show a variety of anomalous thermodynamic, acoustic, and electronic properties.[1,2] $Pb_{1-x}Sn_xTe$ is an especially interesting semiconductor alloy because the symmetry of valence- and conduction-band edges of SnTe is reversed compared to PbTe and other IV-VI semiconductors: The conduction- and the valence-band edges have L_6^- and L_6^+ symmetry, respectively, in PbTe and most other IV-VI semiconductors, while the ordering is "Dimmock reversed" in SnTe.[1-3] This has an interesting consequence: the fundamental band gap closes to zero at an intermediate composition x in $Pb_{1-x}Sn_xTe$.[3] This property of the fundamental energy-band gap vanishing for a selected composition means that alloys with compositions near this composition exhibit small band gaps that satisfy the special needs for infrared sources[4] and detectors[5] in modern technology. Therefore it is very important to understand the effects of alloy disorder on the electronic structures of these technologically important materials.

Recently, Spicer *et al.*[6] have reported experiments indicating the selective breakdown of the virtual-crystal approximation for deep valence bands in $Hg_{1-x}Cd_xTe$ [which is a covalent semiconductor alloy containing "light" (Cd) and "heavy" (Hg) atoms], and have identified that phenomenon as resulting from the Hg $6s$ atomic levels being significantly below the Cd $5s$ levels. Also, Hass *et al.*[7] have obtained similar disorder effects theoretically, in $Hg_{1-x}Cd_xTe$ using the coherent-potential approximation. Davis[8] has also found large deviations from virtual-crystal behavior theoretically in $Pb_{1-x}Sr_xS$ where the cations Pb (configuration $6s^26p^2$) and Sr (configuration $5s^2$) differ so much that an average cation potential is meaningless.

The present work analyzes the effects of alloy disorder on the electronic structures of the random alloys $Pb_{1-x}Sn_xTe$ using the recursion method with a tight-binding model. $Pb_{1-x}Sn_xTe$ is an interesting material for this purpose because its constituent semiconductor compounds PbTe and SnTe have very similar overall electronic structures, except for the Dimmock reversal of the valence- and conduction-band edges; the alloy contains light (Sn) and heavy (Pb) cations. Moreover, the electronic band structures of these materials have large spin-orbit splittings, and the fundamental gaps are not at the center of the Brillouin zone, $k=0$. Indeed, some authors believe that PbTe and SnTe are ionic rather than covalent materials.[9] Therefore the usual criteria[10] for the validity of the virtual-crystal approximation may not apply.

In Sec. II, the tight-binding model for the parent semiconductors PbTe and SnTe is discussed, and the recursion method is outlined. In Sec. III, the results of the calculations are presented and discussed. Section IV summarizes the conclusions.

II. CALCULATIONAL PROCEDURES

A. Tight-binding model

It is well known that $Pb_{1-x}Sn_xTe$ forms a single-phase pseudobinary alloy over the entire composition range x, with about 2% of lattice-constant change from PbTe to SnTe. Both compounds crystallize in the rock-salt structure with lattice constant 6.443 Å for PbTe and 6.327 Å for SnTe (Ref. 11) at 300 K. The electronic structures of PbTe and SnTe (and other IV-VI compounds) have been extensively investigated theoretically and experimentally.[1-3] A variety of computational techniques such as the relativistic augmented-plane-wave (APW) method,[12-14] the orthonormalized-plane-wave (OPW) method,[15] the empirical pseudopotential method,[16-19] and the relativistic Green's function or Korringa-Kohn-Rostoker method (KKR) (Ref. 20) have been used to calculate the electronic band structures of these materials. More recently, a self-consistent relativistic APW calculation for SnTe (Ref. 21) and first-principles pseudopotential total-energy calculation for

 ©1987 The American Physical Society

the ground-state properties and electronic structures of PbTe and SnTe (Ref. 22) have been reported. Although considerable differences may exist concerning some details, such as the parity assignments at the L point[23] and gap structures at critical points (for example, some calculations[13,16,21] showed a "hump structure," i.e., the L point is not a minimum- or maximum-energy point, but a saddle point in SnTe), the general features of the various band structures mentioned above are quite similar. Concentrating on this point and the fact that the recursion method takes its most convenient form in a tight-binding model, we shall use in this work the empirical tight-binding Hamiltonian matrix elements of Lent *et al.*,[24] which are obtained by fitting the eigenvalues of the tight-binding Hamiltonian matrix to the experimental band gap at the L point and to band energies at symmetry points, as calculated by Herman *et al.*[15]

Since the relativistic corrections to the energies of heavy materials, particularly those including Pb, are significant,[25,26] the Hamiltonian used for band calculations should include these effects. The relativistic Hamiltonian which produces the energy-band structure has the following form:[12]

$$H=(p^2/2m)+V+H_{so}+\hbar^2\nabla^2 V/8m^2c^2-p^4/8m^3c^2 , \tag{1}$$

where V is the periodic crystal potential. The spin-orbit term which may split degenerate levels is

$$H_{so}=\hbar\sigma\cdot(\nabla\times\mathbf{p})/4m^2c^2 , \tag{2}$$

and the remaining terms are the Darwin and mass-velocity terms, respectively.

Employing the ideas of Harrison,[27] Chadi,[28] and Vogl *et al.*,[29] the nearest-neighbor tight-binding Hamiltonian can be constructed,

$$\begin{aligned} H_0=&\sum_{\mathbf{R},\sigma,i}(|a,i,\sigma,\mathbf{R}\rangle E_{i,a}\langle a,i,\sigma,\mathbf{R}| \\ &+|c,i,\sigma,\mathbf{R}+\mathbf{d}\rangle E_{i,c}\langle c,i,\sigma,\mathbf{R}+\mathbf{d}|) \\ &+\sum_{\mathbf{R},\mathbf{R}',\sigma,i,j}(|a,i,\sigma,\mathbf{R}\rangle V_{i,j}\langle c,j,\sigma,\mathbf{R}'+\mathbf{d}| \\ &+\text{H.c.})+H_{so} , \end{aligned} \tag{3}$$

where H.c. means Hermitian conjugate, R are the lattice vectors, i and j are the localized quasiatomic orbitals for the cation and anion, σ is the spin index (up or down), a and c refer to the anion and cation, respectively, and d is the position of the cation relative to the anion in any unit cell: $\mathbf{d}=(a_L/2,0,0)$. The spin-orbit interaction term can be described by the following Hamiltonian:[24]

$$H_{so}=\sum_{\mathbf{R},i,j,\sigma,\sigma'}[|a,i,\sigma,\mathbf{R}\rangle(\lambda_a/2)\mathbf{L}_a\cdot\boldsymbol{\sigma}_a\langle a,j,\sigma',\mathbf{R}|+|c,i,\sigma,\mathbf{R}\rangle(\lambda_c/2)\mathbf{L}_c\cdot\boldsymbol{\sigma}_c\langle c,j,\sigma',\mathbf{R}|] . \tag{4}$$

As a basis set, we used 18 quasiatomic orbitals localized on each atomic site which are assumed to be mutually orthonormalized by the method of Löwdin:[30] s, p_x, p_y, p_z, $d_{x^2-y^2}$, $d_{3z^2-r^2}$, d_{xy}, d_{yz}, and d_{zx} for each spin-up and -down state. The parameters of this model are given in Ref. 24, and reproduce the experimental band gaps at the L point[3] (0.186 eV for PbTe and 0.3 eV for SnTe) as well as the calculated band energies of Ref. 15 at the high-symmetry points Γ, X, and L. The resulting band structures are given in Ref. 24. In particular, the Dimmock reversal of the band structure from PbTe to SnTe is correctly reproduced by the model.

B. Recursion method

To obtain the densities of states of $Pb_{1-x}Sn_xTe$ alloys, we require a theory that is capable of predicting the spectra characteristic of pairs and clusters of minority atoms, namely a theory that goes beyond the virtual-crystal approximation (VCA) (Ref. 31) and the coherent-potential approximation (CPA) (Refs. 6, 7, 10, and 32–36). We use the recursion method,[8,37] which exploits the fact that the Hamiltonian matrix for the alloy can be transformed into a real symmetric matrix by unitary transformation from the old basis $|\xi\rangle$ with $\xi=0,1,\ldots,N$, where ξ stands for $b,i,\sigma,\mathbf{R}$ to a new basis $|\nu\rangle$ ($\nu=0,1,2,\ldots,N$). Thus we have

$$H|\nu|=b_\nu|\nu-1|+a_\nu|\nu|+b_{\nu+1}|\nu+1| . \tag{5}$$

With an initial choice of $|0|$ and $b_0=0$, this equation can be iterated to determine the recursion coefficients a_ν and b_ν ($\nu=0,1,\ldots,N$) and the Green's function:

$$G_{0,0}(E)=\cfrac{1}{E-a_0-\cfrac{b_1^2}{E-a_1-\cfrac{b_2^2}{E-a_2-\cdots}}} , \tag{6}$$

where E has an infinitesimal positive imaginary part. In practice this expansion is cut off at some finite level L ($=51$ here), and the remainder is neglected. Then the local densities of states for a specific site b and symmetries i are obtained from $G_{0,0}$ by taking the imaginary part: $(-1/\pi)\mathrm{Im}G_{0,0}$. The choice of initial state $|0|$,

$$0|=\sum_{\mathbf{R},\sigma}\pm|b,i,\sigma,\mathbf{R}\rangle , \tag{7}$$

where $\pm$ means that each term is given a randomly chosen sign, yields the local density of states projected onto the b th cation or anion site and the symmetry i in a random alloy, and the sum of these local-state densities is the total density of states. Details of the method can be found in Ref. 37; computer programs for executing the recursion method are available.[38]

III. RESULTS AND DISCUSSION

We first calculate the density of states for the perfect crystals PbTe and SnTe, employing the nearest-neighbor tight-binding model discussed in the previous section. The results are shown in Figs. 1 and 2. The dot-dashed curve is the density of states obtained by the Lehmann-Taut method.[39] In this method, the Brillouin zone is decomposed into a set of tetrahedra, and the integration over the Brillouin zone is evaluated using an analytic expression. The solid curve is from the recursion method. A $12\times12\times12$-atom cluster was generated to simulate the perfect infinite crystal, and the local density of states for each orbital i,σ was calculated with periodic boundary conditions.

The overall agreement between the two methods is very good, except for some minor details such as the peak structures and the band-gap smearing; the differences between the results of the recursion method and the Lehmann-Taut method are within the tolerable range. The δ-function-like peaks are associated with van Hove singularities[40] due to the long-range order. The more or less smooth peaks in the upper valence bands given by the recursion method (solid curve) are partly due to the finite size of the cluster and partly due to the limited resolution of the present method because of the finite cutoff at $L=51$. (We determined this by varying the size of the cluster and L.) Another difference is that while the Lehmann-Taut method clearly shows the band gap to contain no states, the band edges are smeared in the recursion method. The main reasons for this are the limited resolution of the method and the incomplete cancellation of the off-diagonal elements of the Green's function due to the choice of randomly phased initial state. The band edges can be sharpened by choosing an initial state $|0\rangle$ localized at the center of a cluster or by investigating the spectral density of states (as will be discussed below). In Fig. 3, the contribution of each orbital to the density of states of PbTe is displayed. The lowest valence band is predominantly anion s-like, and the middle valence band is cation s-like. The upper valence bands have dominant anion p-like character, while the lower conduction bands are p-like and cation derived. This can be visualized by the following simple picture. The Pb atom has four valence electrons ($6s^26p^2$) with free-atomic orbital energies -12.42 and -6.97 eV (relative to vacuum) for s and p orbitals, respectively, and the Te atom has six valence electrons ($5s^25p^4$) with orbital energies -19.05 eV ($5s$) and -9.79 eV ($5p$).[41] The two $5s$ electrons of Te, which have the lowest orbital energies, form an isolated valence band deep in energy, and the two $6s$ electrons of Pb form a middle valence band. The two $6p$ electrons of Pb and the four $5p$ electrons of Te interact with each other to form bonding (valence band) and antibonding (conduction band) bands. There-

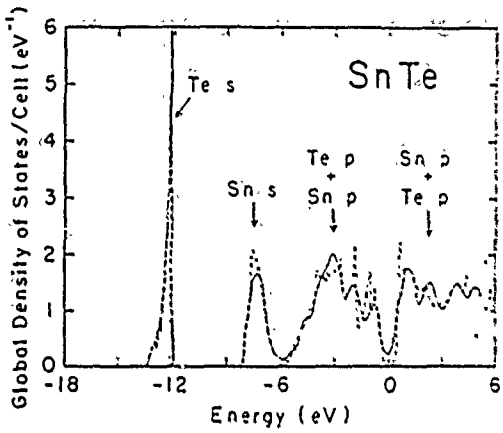


FIG. 2. The virtual-crystal-approximation (dot-dashed curve) and the recursion-method (solid curve) density of states in SnTe. A $12\times12\times12$-atom cluster with periodic boundary conditions was used in the recursion method.

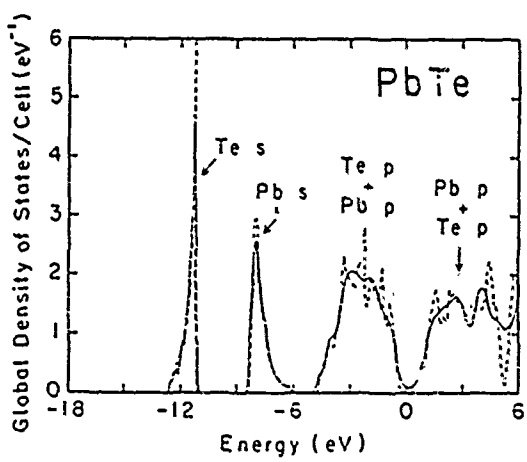


FIG. 1. The virtual-crystal-approximation (dot-dashed curve) and the recursion-method (solid curve) density of states in PbTe. A $12\times12\times12$-atom cluster with periodic boundary conditions was used in the recursion method. The zero of energy is the valence-band maximum.

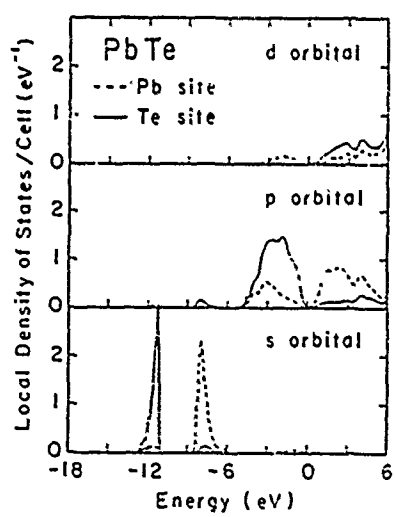


FIG. 3. Local density of states for cation (dot-dashed curve) and anion (solid curve) calculated by the recursion method in PbTe.

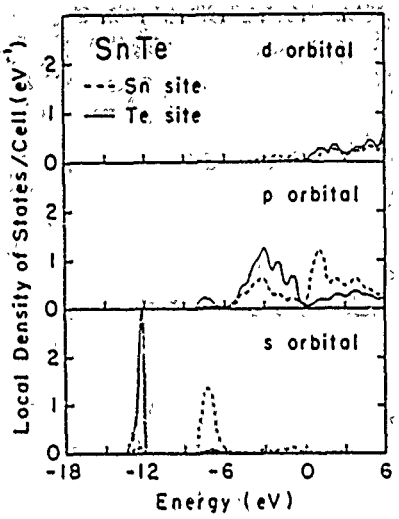


FIG. 4. Local density of states for cation (dot-dashed curve) and anion (solid curve) calculated by the recursion method in SnTe.

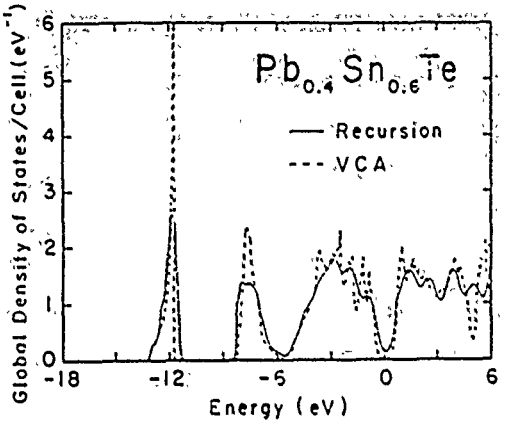


FIG. 6. The virtual-crystal-approximation (dot-dashed curve) and the recursion-method (solid curve) density of states in $Pb_{0.4}Sn_{0.6}Te$. A $12\times12\times12$-atom cluster with periodic boundary conditions was used in the recursion method.

form, alloying PbTe and SnTe, which is equivalent to distributing Pb and Sn atoms randomly on cation sites, has the largest effect on the cationlike middle valence band. The characteristics of the local density-of-states structure in SnTe are similar to those of PbTe (see Fig. 4); the $5s$ and $5p$ free-atomic orbital energies of Sn are at -12.97 and -7.21 eV, respectively.

We generate a model of the random alloy $Pb_{1-x}Sn_xTe$ by randomly occupying cation sites by either Pb (with probability $1-x$) or Sn (with probability x), while all anion sites are occupied by Te. The matrix elements of the alloy Hamiltonian are derived from those of PbTe and SnTe as follows: On cation sites, we use either PbTe or SnTe matrix elements, depending on whether the site was occupied by Pb or Sn. On Te sites, we average the PbTe and SnTe matrix elements, weighting the average in proportion to the number of neighboring Pb and Sn atoms to the Te. Then the densities of states for $Pb_{1-x}Sn_xTe$ are calculated using both the virtual-crystal approximation and the recursion method for a number of compositions x. Again, the density of states is obtained by the use of the Lehmann-Taut method in the virtual-crystal approximation, and a $12\times12\times12$-atom cluster is used in the recursion method with periodic boundary conditions. In order to avoid sample-dependent results, we repeated the calculations for five different alloy configurations of 12^3 atoms, and averaged the densities of states. The results are shown in Figs. 5–8. The solid curves represent the recursion density of states, and the dot-dashed curves are for the virtual-crystal approximation (VCA) results. Both the virtual-crystal approximation and the recursion density of states show the alloying effects, i.e., energy shifts and width changes of the density-of-states peaks. However, analysis of the middle valence band near -7 eV, which has the greatest alloying effects, clearly reveals the differences between the predictions of the two

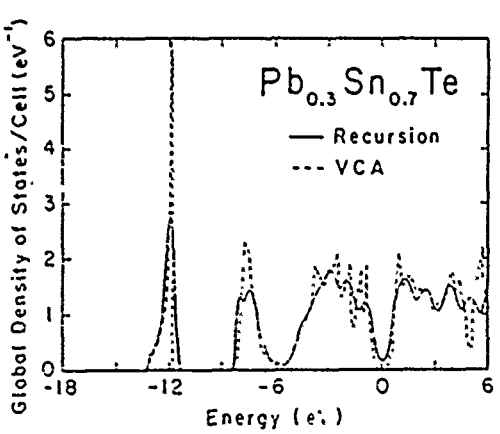


FIG. 5. The virtual-crystal-approximation (dot-dashed curve) and the recursion method (solid curve) density of states in $Pb_{0.3}Sn_{0.7}Te$. A $12\times12\times12$-atom cluster with periodic boundary conditions was used in the recursion method.

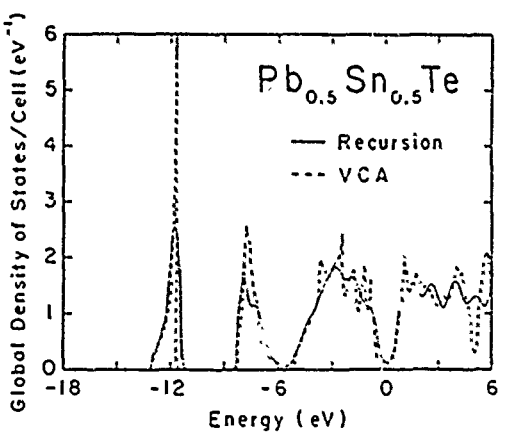


FIG. 7. The virtual-crystal-approximation (dot-dashed curve) and the recursion-method (solid curve) density of states in $Pb_{0.5}Sn_{0.5}Te$. A $12\times12\times12$-atom cluster with periodic boundary condition was used in the recursion method.

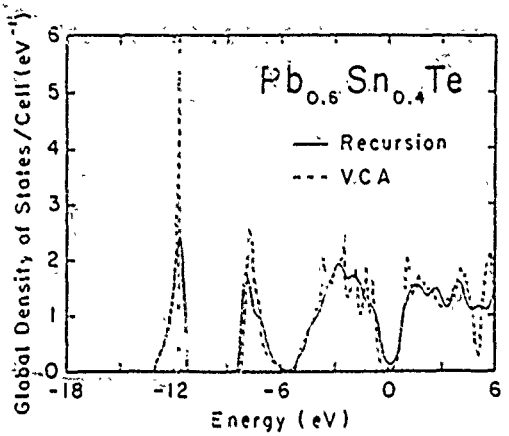


FIG. 8. The virtual-crystal-approximation (dot-dashed curve) and the recursion-method (solid curve) density of states in $Pb_{0.6}Sn_{0.4}Te$. A $12\times12\times12$-atom cluster with periodic boundary condition was used in the recursion method.

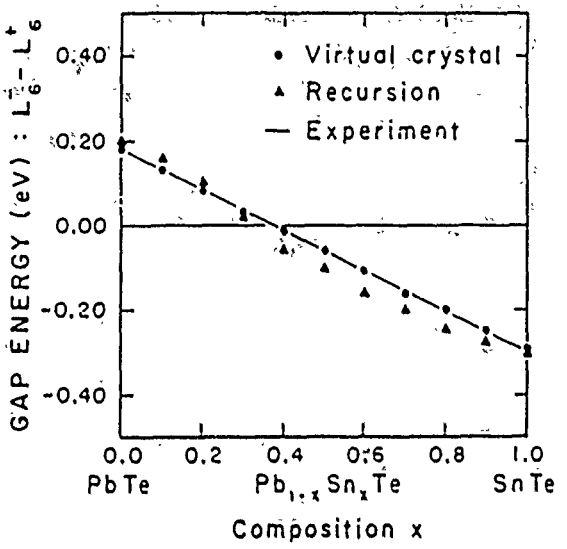


FIG. 9. The band gap $E(L_6^-)-E(L_6^+)$ of $Pb_{1-x}Sn_xTe$ vs composition x. The solid circles (triangles) are obtained using the virtual-crystal approximation (the recursion method), and the solid line represents the interpolation of PbTe and SnTe experimental results of Ref. 3.

methods—the effect of disorder. This band is a doublet, with its low- and high-energy components due to Pb and Sn s states, respectively.

Since the bands that exhibit the alloy effects which are beyond the virtual-crystal approximation are cation s-like in character, spin-orbit coupling does not produce any novel features in the spectra of $Pb_{1-x}Sn_xTe$, beyond the spin-orbit features found in PbTe, SnTe, and a virtual-crystal theory of $Pb_{1-x}Sn_xTe$.

Fortunately the results we find agree rather well with what is expected, based on the Onodera-Toyozawa theory of alloys[10]—despite the fact that theory, to our knowledge, has not been applied previously to alloys with fundamental band gaps at the L point of the Brillouin zone. The density-of-states spectra of the alloys exhibit some features that are "persistent" and others that are "amalgamated" in the terminology of Ref. 10. The persistent features are associated with the cationlike middle valence bands: the Pb $6s$-like and Sn $5s$-like bands that retain their characters in the alloy because the perfect-crystal bands do not overlap in energy. The remaining bands are amalgamated and tend to form hybrids of the PbTe and SnTe bands rather than exhibit separate PbTe- and SnTe-like bands. This amalgamation occurs because the PbTe and SnTe bands overlap in energy, and hence mix in the alloy.[10] Bands that fall within this amalgamated regime can generally be described, in a first approximation, by the virtual-crystal approximation.

Although it is straightforward to include a valence-band offset in the calculation by adding a constant energy to all of the diagonal matrix elements of either PbTe or SnTe (by construction, the matrix elements of Ref. 24 place the zero of energy at the valence-band maximum), we have not done so here because the offset is thought to be small (of order 60 meV),[42] almost negligible on the scale of the figures.

It is well-known that the fundamental band gap of $Pb_{1-x}Sn_xTe$ closes at some intermediate composition because of the inverted band structure of SnTe. We calculated $E(L_6^-)-E(L_6^+)$ of $Pb_{1-x}Sn_xTe$ as a function of composition x by diagonalizing the virtual-crystal empirical tight-binding Hamiltonian (solid circles in Fig. 9). Also the corresponding quantity can be calculated using the recursion method. In alloys, the translational symmetry is broken, thus the wave vector $\mathbf{k}$ is not a good quantum number. However, we still can define the spectral-density functions analogous to those of the perfect crystal by the following:[8]

$$A(\mathbf{k},E)=-(1/\pi)\lim_{\epsilon\to0}\mathrm{Im}\langle b,i,\sigma,\mathbf{k}\,|\,G(E+i\epsilon)\,|\,b,i,\sigma,\mathbf{k}\rangle, \tag{8}$$

where $|b,i,\sigma,\mathbf{k}\rangle$ is a normalized Bloch sum over all unit cells of orbital i with spin σ on each atomic site b (anion or cation). Then the position and broadening of the peak represents the energy shift and damping of a particular quasiparticle state of energy E and wave vector $\mathbf{k}$. Since L_6^- (L_6^+) has anion (cation) p-like character, a Bloch sum of $(p_x+p_y+p_z)/\sqrt{3}$ on each anion (cation) site at the L point is chosen as the initial state $|0\rangle$ for L_6^+ (L_6^-), and the spectral density of states $A(\mathbf{k},E)$ is calculated. Then the gap is defined by the differences in the peak values of $A(\mathbf{k},E)$, i.e., $E(L_6^-)-E(L_6^+)$. The theoretical predictions are shown also in Fig. 9 (solid triangles) in comparison with a linear interpolation of the experimental band gaps of PbTe and SnTe (Ref. 3) (solid line). The theoretical uncertainty in $E(L_6^-)-E(L_6^+)$ is $\sim\pm0.02$ eV for $0<x<1$. The calculated band gap is almost a linear function of composition x and compares well with the experimental results.

IV. SUMMARY

The electronic structures of $Pb_{1-x}Sn_xTe$ alloys, including their parent semiconductor compounds, have

been analyzed using the tight-binding model with spin-orbit interaction. The densities of states were computed for both the effective media using the virtual-crystal approximation and the realistic media employing the recursion method, and the results were compared. As expected, both theories exhibited alloying effects such as band broadening and energy shifts. However, the two methods differed in their predictions for the cation-derived s-like states, which experienced the greatest alloying effect. The alloy composition dependence of the band gap at the L point was analyzed, and exhibits Dimmock's band-crossing phenomenon. The above facts show that the recursion method is a useful tool for the study of the electronic structure of random $Pb_{1-x}Sn_xTe$, and in particular for the cationlike middle valence band. However, they also show that the virtual-crystal approximation provides a remarkably good description of the electronically important top valence and bottom conduction bands. Finally, they demonstrate that the Onodera-Toyozawa criteria can be applied to $Pb_{1-x}Sn_xTe$, even though these alloys have their fundamental band gaps at L: the cationlike s-like middle valence bands are persistent while the top valence band and lowest conduction band are amalgamated.

ACKNOWLEDGMENTS

We are grateful to the Office of Naval Research (Contract No. N00014-84-K-0352), the Air Force Office of Scientific Research (Contract No. AFOSR-85-0331), the University of Notre Dame, and the University of Illinois at Urbana-Champaign Physics Department for their support of this work. We benefited greatly from stimulating conversations with M. Bowen, E. Ho, C. Lent, and O. Sankey.

[1]W. Jantsch, in *Dynamical Properties of IV-VI Compounds*, Vol. 99 of *Springer Tracts in Modern Physics* (Springer-Verlag, Berlin, 1983), p. 1; A. Bussmann-Holder, H. Bilz, and P. Vogl, *ibid.*, p. 51.

[2]G. Nimiz and B. Schlicht, in *Narrow-Gap Semiconductors*, Vol. 98 of *Springer Tracts in Modern Physics* (Springer-Verlag, Berlin, 1983), p. 1.

[3]J. O. Dimmock, I. Melngailis, and A. J. Strauss, Phys. Rev. Lett. 16, 1193 (1966).

[4]H. Preier, Appl. Phys. 20, 189 (1979).

[5]H. Holloway, J. Appl. Phys. 50, 1386 (1979).

[6]W. E. Spicer, J. A. Silberman, M. Morgan, I. Lindau, J. A. Wilson, A.-B. Chen, and A. Sher, Phys. Rev. Lett. 49, 948 (1982).

[7]K. C. Hass, H. Ehrenreich, and B. Velický, Phys. Rev. B 27, 1088 (1983).

[8]L. C. Davis, Phys. Rev. B 28, 6961 (1983).

[9]M. Schlüter, G. Martinez, and M. L. Cohen, Phys. Rev. B 11, 3808 (1975).

[10]Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968).

[11]R. F. Bis and J. R. Dixon, J. Appl. Phys. 40, 1918 (1969).

[12]J. B. Conklin, L. E. Johnson, and G. W. Pratt, Phys. Rev. 137, A1282 (1965).

[13]S. Rabii, Phys. Rev. 182, 821 (1969).

[14]S. Rabii and R. H. Lasseter, Phys. Rev. Lett. 33, 703 (1974).

[15]F. Herman, R. L. Kortum, I. B. Ortenberg, and J. P. van Dyke, J. Phys. (Paris) Colloq. 29, C4-62 (1968).

[16]Y. W. Tung and M. L. Cohen, Phys. Rev. 180, 823 (1969).

[17]R. L. Bernick and L. Kleinman, Solid State Commun. 8, 569 (1970).

[18]S. E. Kohn, P. Y. Yu, Y. Petroff, Y. R. Shen, Y. Tsang, and M. L. Cohen, Phys. Rev. B 8, 1477 (1973).

[19]G. Martinez, M. Schlüter, and M. L. Cohen, Phys. Rev. B 11, 651 (1975).

[20]H. Overhof and U. Rössler, Phys. Status Solidi 37, 691 (1970); J. Korringa, Physica 13, 392 (1947); W. Kohn and N. Rostoker, Phys. Rev. 94, 1111 (1954).

[21]J. S. Melvin and D. C. Hendry, J. Phys. C 12, 3003 (1979).

[22]K. M. Rabe and J. D. Joannopoulos, Phys. Rev. B 32, 2302 (1985).

[23]See for example, G. Martinez and M. L. Cohen, Phys. Rev. Lett. 35, 1746 (1975), and references therein.

[24]C. S. Lent, M. A. Bowen, J. D. Dow, R. S. Allgaier, O. F. Sankey, and E. S. Ho, Superlatt. Microstruct. 2, 491 (1986).

[25]L. E. Johnson, J. B. Conklin, and G. W. Pratt, Phys. Rev. Lett. 11, 538 (1963).

[26]F. Herman, C. D. Kuglin, K. F. Cuff, and R. L. Kortum, Phys. Rev. Lett. 11, 541 (1963).

[27]W. A. Harrison, Phys. Rev. B 8, 4487 (1973).

[28]D. J. Chadi, Phys. Rev. B 16, 790 (1977).

[29]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[30]P.-O. Lödin, J. Chem. Phys. 17, 365 (1950).

[31]L. Nordheim, Ann. Phys. (N.Y.) 9, 607 (1931); 9, 641 (1931); F. Bassani and D. Brust, Phys. Rev. 131, 1524 (1963); H. Amar, K. H. Johnson, and C. D. Sommers, *ibid.* 153, 655 (1967); M. M. Pant and S. K. Joshi, *ibid.* 184, 635 (1969).

[32]P. Soven, Phys. Rev. 156, 809 (1967).

[33]D. W. Taylor, Phys. Rev. B 156, 1017 (1967).

[34]B. Velický, S. Kirkpatrick, and H. Ehrenreich, Phys. Rev. 175, 747 (1968).

[35]A.-B. Chen and A. Sher, Phys. Rev. B 17, 4726 (1978).

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

[37]R. Haydock, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35, p. 215.

[38]S. Lee, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1986 (unpublished).

[39]G. Lehmann and M. Taut, Phys. Status Solidi B 54, 469 (1972).

[40]L. van Hove, Phys. Rev. 89, 1189 (1953).

[41]C. F. Fisher, At. Data 4, 301 (1972).

[42]K. E. Ambosch, H. Clemens, E. J. Fantner, G. Bauer, M. Kriechbaum, P. Kocevar, and R. J. Nicholas, Surf. Sci. 142, 571 (1984).

Dangling bonds and Schottky barriers

Otto F. Sankey
Department of Physics, Arizona State University, Tempe, Arizona 85287

Roland E. Allen and Shang-Fen Ren
Department of Physics, Texas A&M University, College Station, Texas 77843

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 30 January 1985; accepted 24 April 1985)

We review theoretical interpretations of Schottky barriers and Fermi-level pinning, which result when metals and other chemical species are deposited on semiconductor surfaces. Experiments indicate that these two phenomena are closely connected, so a theory of Schottky barriers must also explain Fermi-level pinning for submonolayer coverages of both metallic and nonmetallic species. Proposed mechanisms include the following: (a) *Intrinsic surface states.* For GaAs and several other materials, there are no intrinsic surface states within the band gap; GaP, e.g., does have surface states in the gap, but they are not at the correct energy to explain Schottky barrier formation. (b) *Metal-induced gap states.* These states, which require a thick metal overlayer, cannot explain Fermi-level pinning at submonolayer metallic coverages. They also cannot explain why a single semiconductor (*n*-type InP) exhibits two distinct Schottky barrier heights. Furthermore, they cannot explain why the Schottky barrier persists when there is an oxide layer between semiconductor and metal. Metal-induced states can in principle give rise to Schottky barriers at defect-free interfaces, but they fail to explain much of the existing experimental data for III–V semiconductors and Si. (c) *The classic Schottky model.* This model is not in agreement with experiment for III–V and Group IV semiconductors, but does appear to account for the measurements involving nonreactive metals on GaSe—a layered material expected to be relatively free of defects. (d) *The Spicer defect model.* This phenomenological model, now supported by microscopic theoretical studies, appears to account for many of the observations regarding Schottky barrier and Fermi-level pinning. We review our theoretical investigations within the framework of the defect model, which provide a satisfactory explanation of the principal observations for both III–V and Group IV semiconductors. We conclude that the levels responsible for Schottky barriers and Fermi-level pinning arise from two sources: (1) bulk-derived deep levels (e.g., the deep donor level for the antisite defect As_{Ga}, which persists when this defect is present at the surface, but which is shifted in energy), and (2) dangling-bond deep levels (which are also shifted in energy according to the environment of the dangling bond). Most of the observed Schottky barriers—for both III–V and Group IV semiconductors—are attributed to dangling bonds.

I. INTRODUCTION

Schottky barriers have long been of interest, and many microscopic mechanisms and phenomenological pictures have been hypothesized to explain them. Nowadays, however, the theorist is considerably more constrained by the body of experimental findings than he was only a few years ago. It is no longer sufficient to offer a theory that fits only a limited set of Schottky barrier data. A theory now must also explain the related data on Fermi-level pinning at submonolayer coverages of various metals and other chemical species.[1–15] For example, the experiments of Williams and co-workers[5–7] indicate that *n*-InP exhibits two distinct Fermi-level pinning positions, even for thick metal depositions—one just below the conduction band edge E_c and the other about 0.5 eV below E_c. Recently Mönch and Gant[11] discovered that the Fermi-level pinning for *p*-GaAs anneals out at the annealing temperature of an antisite defect. A satisfactory theory of Schottky barriers must explain such observations in addition to successfully predicting Schottky barrier heights.

Here we review various microscopic mechanisms that have been proposed to explain Schottky barriers and Fermi-level pinning, and we examine to what extent these mechanisms are compatible with the experimental data.

II. POSSIBLE MECHANISMS

A. The classic Schottky mechanism

The original interpretation of Schottky barriers—that which is associated with Schottky himself (and other early workers)—is that charge transfer between semiconductor and metal results in a dipole layer at the interface and an equilibration of Fermi energies. According to this interpretation, we have

$$\phi_B = W - \chi, \tag{1}$$

where ϕ_B is the barrier height, W is the work function of the metal, and χ is the electron affinity of the semiconductor.[16] Equation (1) is violated for III–V and Group IV semiconduc-

 0734-211X/85/041162-05$01.00 © 1985 American Vacuum Society

tors,[2–15] and this is essentially the reason that Bardeen originally proposed Fermi-level pinning by surface states of some kind. However, Eq. (1) may be consistent with the measurements for some more ionic systems, such as some II–VI's, and for relatively nonreactive metals (the noble metals, Au and Ag) on GaSe. Since GaSe is a layered material expected to have a low concentration of defects, this observation indicates that nearly defect-free surfaces can exhibit "Schottky" type Schottky barrier formation.

B. Fermi-level pinning

Bardeen proposed that Schottky barrier heights are due to Fermi-level pinning by surface states, either intrinsic or extrinsic.[1] The essential feature of the Bardeen model is that the Fermi energies of the bulk semiconductor, the semiconductor surface, and the bulk metal must align—and that charges diffuse, causing band bending, until this alignment is achieved. Much of the current interpretation of Schottky barrier formation is based on this Fermi-level pinning model—with many of the current controversies centered on the issue of what agents do the Fermi-level pinning at specific metal/semiconductor contacts prepared under specific well-defined conditions.

1. Intrinsic surface states

Since the mid-1970's, it has been known that there are no intrinsic surface states in the band gap for the (110) surface of GaAs,[17–19] and so such states cannot account for the observed Schottky barriers and Fermi-level pinning on this surface. This same conclusion appears to apply to other direct-gap III–V semiconductors like InP. In the case of GaP, there are intrinsic (110) surface states within the band gap,[20] but these states are just under the conduction-band edge, whereas the observed Fermi-level pinning position for *n*-GaP (110) is much lower in the band gap—at an energy in good agreement with the predicted acceptor level for the surface antisite defect P_{Ga}.[21]

2. Metal-induced gap states

The possibility of metal-induced gap states at semiconductor/metal interfaces has been discussed by a number of workers.[16,22–30] Such states should be properly treated through a microscopic calculation in which semiconductor atoms are bonded to metal atoms. Only a calculation of this kind—with the interface electronic states obtained by solving Schrodinger's equation at the semiconductor/metal interface in an atomistic picture—constitutes a proper microscopic theory. Some treatments of metal-induced gap states, however, have been based on plausibility arguments or very crude approximations, with a consequent loss of reliability.

In principle, metal-induced gap states can produce Schottky barriers and Fermi-level pinning. In practice, however, they do not seem to explain the observations on those systems that have been studied up until now. There are several difficulties in trying to explain the observations with metal-induced states. (1) There is a clear connection between Schottky barriers for thick metal contacts and Fermi-level pinning for submonolayer coverages of metal[2–15], i.e., a theory of the former is also required to be a theory of the latter. Metal-induced gap states, however, are supposed to extend into the bulk of the metal, and so *they are not even defined for submonolayer coverages.* (2) Theories of metal-induced states predict a single barrier height for each semiconductor, almost independent of the metal in the metal/semiconductor contact and insensitive to surface treatments. They, therefore, fail to explain why two distinct Schottky barrier heights are observed for *n*-InP.[5–7] (3) Schottky barrier heights are often not very much affected when an oxide layer is inserted between the semiconductor and the metal (see, e.g., Ref. 5), whereas metal-induced gap states require an intimate semiconductor/metal contact; i.e., metal-induced gap states do not appear to explain real-world Schottky barriers, which often involve an oxide layer at the interface. (4) Metal-induced states do not apply to semiconductor/nonmetallic interfaces, like the GaAs/oxide interface, which exhibits the same Fermi-level pinning as metals.[2–4,8–14] (5) The states associated with any adsorbate—metal or otherwise—can be expected to depend strongly on the chemical identity of the adsorbate. Fermi-level pinning positions and Schottky barrier heights, however, are often about the same for a variety of adsorbates.[1–15] (6) Estimates indicate that the dipole produced by metal-induced gap states will often be at least an order of magnitude too small to account for the Schottky barrier height.[31] (7) The Fermi-level pinning for *p*-GaAs (110) anneals out at the annealing temperature of an antisite defect[11]—a fact that is difficult to explain in a metal-induced gap-state picture.

3. The Spicer defect model

The fact that measurements on several III–V semiconductors (GaAs, InP, and GaSb) provided strong evidence for Fermi-level pinning—together with the fact that experiments also indicate that these materials have no intrinsic surface states within the band gap—led Spicer and coworkers[2–4] to propose the defect model. In this model, chemisorption (or cleavage) produces defects of some kind, and these defects give rise to deep (or shallow) levels that account for the observed Fermi-level pinning and Schottky barrier formation.

The original Spicer defect model was phenomenological, and did not specify the precise nature of the defects responsible for barrier formation. Over the past several years, we have attempted to provide a microscopic theory by performing detailed calculations for particular defects.[32–44] We find that the principal experimental observations are well explained by the theoretical results for a few simple defects.

For example, Fig. 1 shows the interpretation of the observed Fermi-level pinning positions and Schottky barrier heights for various metals and other chemical species deposited on the (110) surface of *n*-InP.[5] The fact that there are two observed barrier heights for this single semiconductor finds a natural explanation in the defect model—namely, there are two different native defects involved (in addition to possible surface impurities). Nonreactive metals produce antisite defects, and reactive metals produce vacancies[37–44] as the dominant defects. The observed "switching" of barrier heights

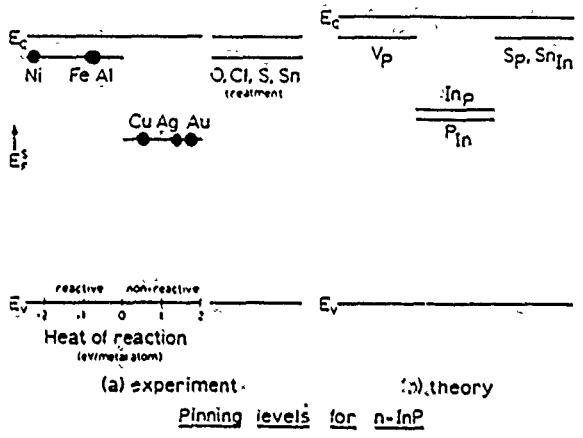


FIG. 1. Surface Fermi energy of *n*-type InP vs heat of reaction of InP with the metals Ni, Fe, Al, Cu, Ag, and Au, according to Ref. 5. The theoretical surface defect levels for the P vacancy (V_P), the native antisite defects (In_P and P_{In}), and the extrinsic impurities S on a P site (S_P) and Sn on an In site (Sn_{In}), are given at the right. The *n*-InP data can be interpreted as follows: Nonreactive metals produce antisite defects; reactive metals and treatment of the surface with oxygen and Cl produce P vacancies. Treatments with Sn and S produce surface Sn_{In} and S_P. After Refs. 37 and 38.

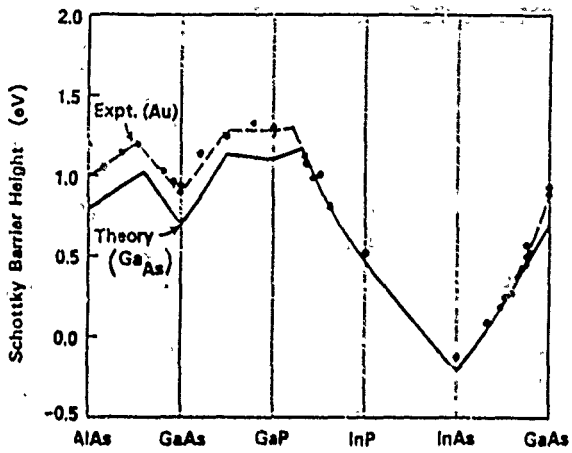


FIG. 3. Schottky barrier height ϕ_B as function of alloy composition x for Au contacts to *n*-type $Al_{1-x}Ga_xAs$, $GaAs_{1-x}P_x$, $Ga_{1-x}In_xP$, $InP_{1-x}As_x$, and $In_{1-x}Ga_xAs$, after Ref. 39, where the source of the experimental data are cited. The theoretical barrier heights correspond to the antisite defect cation-on-anion-site (e.g., Ga_{As}).

shown in Fig. 1 provides an initial but crucial test for any theory of Schottky barriers.

In Fig. 2, the experimental[2–4,9–11] and theoretical[33] levels for GaAs (110) are compared. Notice that there are again two acceptor levels predicted by the theory for surface antisite defects, in accordance with the experimental observation of two distinct Fermi-level pinning positions for the single semiconductor *n*-GaAs.[9–11]

In Fig. 3, we show the comparison of experimental and theoretical Schottky barrier heights, with the barrier (for Au contacts) assigned to the surface antisite defect, cation-on-anion-site (e.g., Ga_{As}). The level of quantitative agreement is fortuitously good, in view of the fact that there are uncertainties in the theory of several tenths of an eV; i.e., we do not claim that the theory can predict Schottky barrier heights to within 0.1 eV! However, we regard Fig. 3 as dramatic evidence that the defect model does provide a very satisfactory explanation of Schottky barrier heights as well as Fermi-level pinning.

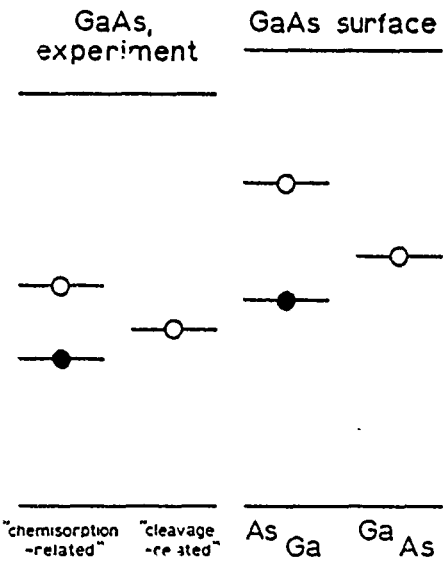


FIG. 2. Experimental Fermi-level pinning positions for *n*- and *p*-type GaAs, (Refs. 2–4 and 8–11) compared with theoretical predictions for surface antisite defects (As_{Ga} and Ga_{As}). Solid circles refer to *p*-type GaAs (experiment), or to donor levels which will produce Fermi-level pinning on *p*-GaAs (theory); open circles refer to *n*-GaAs, or to acceptor levels which will produce pinning on *n*-GaAs.

As the above examples make clear, the theory indicates that most observations of Schottky barriers and Fermi-level pinning for III–V semiconductors are explained by surface antisite defects (with vacancies also involved in a few cases). The antisite defects' deep levels that are responsible for Fermi-level pinning invariably have dangling-bond character. For Group IV semiconductors, however, there are no antisite defects. In this case, an even simpler type of the defect suffices to explain the observations—namely, dangling bonds. These dangling bonds must have a neighboring void, vacancy, or disordered region into which they dangle; otherwise, their energy levels would be severely altered by the metal of the metal/semiconductor contact. Therefore, we refer to them as "sheltered,"[38] and recognize that a sheltered antisite defect is essentially a (vacancy, antisite) pair.

As described in some detail in Refs. 40–43, dangling bonds provide a very good explanation of the experimental measurements for Si, Ge, diamond, and amorphous Si interfaced with various metals.[45–51] For example, on the right-hand side of Fig. 4 we compare the theoretical dangling bond energy with the experimental Fermi-level position for Group IV semiconductors (inferred from the Schottky barrier height ϕ_B for Au contacts, as described in the figure caption). It can be seen that there is good agreement between theory and experiment, with respect to both chemical trends and the position of the energy within the band gap. This same level of quantitative and qualitative agreement is found for various other systems—notably, for the technologically important Si/transition-metal-silicide contacts.[40]

In Fig. 4, we also show the cation dangling-bond energies for III–V semiconductors, together with the experimental

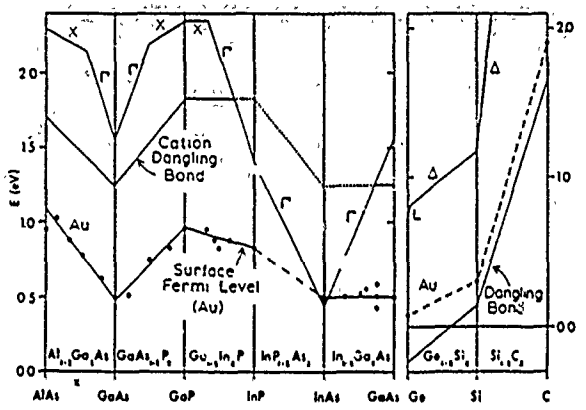


FIG. 4. Surface Fermi energy for *n*-type semiconductors, inferred from Schottky barrier measurements (using $\phi_B = E_c^S - E_F^S$, where ϕ_B is the barrier height, and E_c^S and E_F^S are, respectively, the conduction band minimum and the Fermi energy at the semiconductor surface), compared with the surface Fermi energy predicted for dangling bond defects. The sources for the data are cited in Refs. 39 (III–V semiconductors) and Ref. 42 (Group IV semiconductors). Both the chemical trends and the positions within the band gap are predicted correctly for the Group IV semiconductors. For the III–V semiconductors, the chemical trends are predicted correctly, but the positions within the band gap are too high by 3/4 eV. This indicates that the relevant dangling bonds are associated with antisite defects for the III–V's, whereas simple dangling bonds explain the data for the Group IV's.

surface Fermi-level positions (inferred from the Schottky barrier data displayed in Fig. 3). Since the cation dangling bond state is empty (an acceptor), it would provide Fermi-level pinning on *n*-type III–V semiconductors if dangling bonds were responsible for Fermi-level pinning on these materials. (The anion dangling-bond state is filled—a donor—so it would provide pinning on *p*-type III–V's.)

Two interesting facts are evident on the left-hand side of Fig. 4: (1) The cation dangling bond does a remarkably good job of reproducing the chemical trends of the measurements. (2) However, it does a rather poor job of reproducing the positions of the measured energies within the gap. In fact, the cation dangling-bond energy is about 3/4 eV too high to explain barrier heights and Fermi-level pinning on *n*-type III–V semiconductors.

Although we do not display the results here, we have also calculated the anion dangling-bond energies.[52] These are found to be too *low* by about 3/4 eV to explain barrier heights and Fermi-level pinning on *p*-type semiconductors.

Although *intrinsic dangling bonds* do not appear to produce deep levels in agreement with the measurements on III–V semiconductors, it is important to emphasize that the levels that we invoke to explain Schottky barriers and Fermi-level pinning—for example, the levels of Figs. 1–3—are in most cases due to *antisite dangling bonds*. We note that the antisite dangling-bond energies typically "undermine" the intrinsic dangling-bond energies, in that the antisite acceptor levels lie lower than the intrinsic dangling-bond acceptor levels, and it is the *lowest* acceptor level that tends to produce Fermi-level pinning for an *n*-type semiconductor. [For example, compare the GaAs cation dangling-bond gap deep energy level of Fig. 4 with the lower GaAs antisite dangling-bond level (lower open circle) of Fig. 2.]

In order for a defect to be relevant to Schottky barrier formation, of course, it must be present in appreciable concentration. This means about 1 surface defect per 100 surface atoms, according to our calculations and those of Spicer *et al.*[53] (For bulk defects, the screening of the semiconductor, with dielectric constant $\epsilon \simeq 10$, implies that an order of magnitude more defects are needed—about 1 bulk defect per 10 surface atoms.[54]) However, we do not find it implausible that there may be a rather high concentration of intrinsic and antisite dangling bonds at the rather disordered—and, unfortunately, not yet completely characterized—interfaces between III–V semiconductors and metals. In fact, Mönch *et al.*[11] find that for the "mild" Ge on GaAs system, at low coverage ($\simeq$0.5 monolayer), states (defects) are created at a rate of 0.06 ± 0.04 per deposited Ge atom, which for a coverage of only a single monolayer corresponds to between 1 defect in 10 and 1 defect in 50. If III–V semiconductor/metal interfaces exhibited perfect bonding and perfect order, then the defect model would be inapplicable. However, as evidenced above, we believe that such interfaces are not in fact perfectly ordered as they are grown currently in the laboratory.

III. CONCLUSIONS

The present theory of Schottky barriers and Fermi-level pinning by native defects provides a very satisfactory explanation of the experimental observations for both III–V and Group IV semiconductors. This theory involves two types of defect levels:

(1) Bulk-derived interfacial defect levels, such as the As_{Ga} donor level of Fig. 2. This *s*-like level is derived from the bulk As_{Ga} A_1 donor level,[52] and is merely shifted in energy at the surface.[55]

(2) Dangling bond levels. The observations for Group IV semiconductors—Si, Ge, diamond, and amorphous Si—are explained by intrinsic dangling bonds.[40–43] The observations for III–V semiconductors—GaAs, InP, etc., and their alloys—are explained in most cases by dangling bonds associated with antisite defects, such as As on the Ga site and Ga on the As site. In some cases, such as reactive metals on *n*-InP, surface vacancies—which involve intrinsic dangling bonds—appear to be involved.

Our theory of Schottky barriers is thus primarily a theory of Fermi-level pinning by dangling bonds.

ACKNOWLEDGMENT

We thank the U.S. Office of Naval Research for their support, which made this work possible (Contract Nos. N00014-82-K-0447 and N00014-84-K-0352).

[1] J. Bardeen, Phys. Rev. 71, 717 (1947).
[2] W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. 16, 1422 (1979) and references therein.
[3] W. E. Spicer, I. Lindau, P. R. Skeath, and C. Y. Su, J. Vac. Sci. Technol. 17, 1019 (1980) and references therein.
[4] W. E. Spicer, I. Lindau, P. R. Skeath, C. Y. Su, and P. W. Chye, Phys. Rev. Lett. 44, 520 (1980).

[5]R. H. Williams, V. Montgomery, and R. R. Varma, J. Phys. C 11, L735 (1978).
[6]R. H. Williams and M. H. Patterson, Appl. Phys. Lett. 40, 484 (1982).
[7]R. H. Williams, Surf. Sci. 132, 122 (1983).
[8]W. Mönch and H. J. Clemens, J. Vac. Sci. Technol. 16, 1238 (1979).
[9]H. Gant and W. Mönch, Appl. Surf. Sci. 11/12, 332 (1982).
[10]W. Mönch and H. Gant, Phys. Rev. Lett. 48, 512 (1982).
[11]W. Mönch, Surf. Sci. 132, 92 (1983). This paper, and Refs. 7 and 14, also appear in *Surfaces and Interfaces: Physics and Electronics*, edited by R. S. Bauer (North-Holland, Amsterdam, 1983); W. Mönch, in *Chemistry and Physics of Solid Surfaces V*, edited by R. Vanselow and R. Howe (Springer, New York, 1984), p. 501.
[12]H. H. Wieder, Inst. Phys. Conf. Ser. 50, 234 (1980).
[13]H. H. Wieder, Appl. Phys. Lett. 38, 170 (1981).
[14]H. H. Wieder, Surf. Sci. 132, 390 (1983).
[15]L. J. Brillson, Surf. Sci. Rep. 2, 123 (1982) and references therein.
[16]A variation is the effective work function model, or cluster model: J. M. Woodall and J. L. Freeouf, J. Vac. Sci. Technol. B 2, 510 (1984). Since the present paper emphasizes models for which microscopic quantum-mechanical calculations have been performed, we do not discuss this model or the metallic cluster model of R. Ludeke, T.-C. Chiang, and T. Miller, J. Vac. Sci. Technol. B 1, 581 (1983).
[17]J. van Laar and J. J. Scheer, Surf. Sci. 8, 342 (1967); J. van Laar and A. Huijser, J. Vac. Sci. Technol. 13, 769 (1976).
[18]W. Gudat and D. E. Eastman, J. Vac. Sci. Technol. 13, 831 (1976).
[19]W. E. Spicer, I. Landau, P. E. Gregory, C. M. Garner, P. Pianetta, and P. Chye, J. Vac. Sci. Technol. 13, 780 (1976).
[20]A. Huijser and J. van Laar, Surf. Sci. 52, 202 (1975); A Huijser, J. van Laar, and T. L. van Rooy, Surf. Sci. 62, 472 (1977); G. M. Guichar, C. A. Sebenne, and C. D Thualt, J. Vac. Sci. Technol. 16, 1212 (1979).
[21]A. D. Katnani, G. Margaritondo, R. E. Allen, and J. D. Dow, Solid State Commun. 44, 1231 (1982).
[22]V. Heine, Phys. Rev. A 138, 1689 (1965).
[23]J. C. Inkson, J. Phys. C 6, 1350 (1973).
[24]S. G. Louie and M. L. Cohen, Phys. Rev. Lett. 35, 866 (1975).
[25]E. Louis, F. Yndurain, and F. Flores, Phys. Rev. B 13, 4408 (1976).
[26]S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B 15, 2154 (1977).
[27]C. Tejedor, F. Flores, and E. Louis, J. Phys. C 10, 2163 (1977).
[28]F. Guinea, J. Sanchez-Dehesa, and F. Flores, Phys. Rev. B 16, 6499 (1983).
[29]J. Tersoff, Phys. Rev. Lett. 52, 465 (1984).
[30]J. Sanchez-Dehesa and F. Flores, Solid State Commun. 50, 29 (1984); J. Sanchez-Dehesa, F. Flores, and J. D. Dow (to be published) and references therein.
[31]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), p. 428.
[32]The first calculations for surface vacancies were performed by M. S. Daw and D. L. Smith, Phys. Rev. B 20, 5150 (1979).
[33]R. E. Allen and J. D. Dow, Phys. Rev. B 25, 1423 (1982).
[34]R. E. Allen and J. D. Dow, J. Vac. Sci. Technol. 19, 383 (1981).
[35]R. E. Allen and J. D. Dow, Appl. Surf. Sci. 11/12, 362 (1982).
[36]R. E. Allen, H. P. Hjalmarson, and J. D. Dow, Solid State Commun. 41, 419 (1982).
[37]J. D. Dow and R. E. Allen, J. Vac. Technol. 22, 659 (1982).
[38]R. E. Allen, R. P. Beres, and J. Dow, J. Vac. Sci. Technol. B 1, 401 (1983).
[39]R. E. Allen, T. J. Humphreys, J. D. Dow, and O. F. Sankey, J. Vac. Sci. Technol. B 2, 449 (1984).
[40]O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1983).
[41]O. F. Sankey, R. E. Allen, and J. D. Dow, J. Ultramicrosc. 14, 127 (1984).
[42]O. F. Sankey, R. E. Allen, and J. D. Dow, J. Vac. Sci. Technol. B 2, 491 (1984).
[43]O. F. Sankey, R. E. Allen, and J. D. Dow, in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, edited by D. J. Chadi and W. A. Harrison (Springer, New York, 1985).
[44]J. D. Dow, R. E. Allen, and O. F. Sankey in *Chemistry and Physics of Solid Surfaces V*, edited by R. Vanselow and R. Howe (Springer, New York, 1984), p. 483.
[45]C. A. Mead and W. G. Spitzer, Phys. Rev. A 134, 713 (1964).
[46]See, for example, *Thin Films and Interfaces*, edited by P. S. Ho and K. N. Tu (North-Holland, New York, 1982). Reprinted from the Thin Solid Films 93, (1983).
[47]P. S. Ho and G. W. Rubloff, Thin Solid Films 89, 433 (1982).
[48]E. H. Rhoderick, *Metal-Semiconductor Contacts* (Clarendon, Oxford, 1978).
[49]J. M. Andrews and J. C. Phillips, Phys. Rev. Lett. 35, 56 (1975).
[50]G. Ottaviani, K. N. Tu, and J. W. Mayer, Phys. Rev. B 24, 3354 (1981).
[51]D. Cherns, G. R. Anstis, J. L. Hutchinson, and J. C. H. Spence, Philos. Mag. A 46, 849 (1982).
[52]Dangling-bond energies for III–V semiconductors were determined in exactly the same way as the Si dangling-bond energy was determined in Ref. 40—by fitting to calculated A_1 and T_2 vacancy levels. For the results shown in Fig. 4, we used the vacancy levels predicted by the sp^3s^* model of P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983). The results of Figs. 1–3 were calculated with this model and the impurity potential model of H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).
[53]W. E. Spicer, S. Pan, D. Mo, N. Newman, P. Mahowald, T. Kendelewicz, and S. Eglash, J. Vac. Sci. Technol. B 2, 476 (1984).
[54]A. Zur, T. C. McGill, and D. L. Smith, J. Vac. Sci. Technol. B 1, 608 (1983).
[55]It is also possible that bulk defects are involved in some cases of Fermi-level pinning; for example, the bulk complex (Ga_{Sb},V_{Ga}), (V = vacancy) may explain the shallow pinning level observed for GaSb (Refs. 2 and 36).

Predicted energy band gaps of $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ metastable, substitutional, crystalline alloys

David W. Jenkins, Kathie E. Newman, and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 21 January 1985)

Predictions of the energy band gaps as functions of alloy composition are given for the Greene alloys, which are metastable, crystalline, substitutional alloys of III-V compounds and group-IV elemental materials. All possible combinations of these alloys involving Al, Ga, In, P, As, Sb, Si, Ge, and Sn are considered. The Γ and L conduction-band minima, relative to the valence-band maxima, exhibit characteristic V-shaped bowing and kinks as functions of composition x; the band edges at point X bifurcate at critical compositions corresponding to the order-disorder transition of Newman *et al.* The V-shaped bowing due to the transition offers the possibility of band gaps significantly smaller than expected on the basis of the conventional virtual-crystal approximation. Alloys with modest lattice mismatches that are predicted to have especially interesting band gaps include $(InP)_{1-x}Ge_{2x}$, $(AlSb)_{1-x}Sn_{2x}$, $(GaSb)_{1-x}Sn_{2x}$, and $(InAs)_{1-x}Sn_{2x}$, which are alloys with potentially small band gaps, and $(AlAs)_{1-x}Ge_{2x}$ and $(GaAs)_{1-x}Si_{2x}$, which are alloys with larger gaps and several interesting band-edge crossings as functions of composition.

I. INTRODUCTION

Recently, Greene and co-workers have fabricated a new class of semiconducting $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ alloys for a wide range of compositions.[1–4] The III-V compounds and group-IV elemental materials are normally immiscible at equilibrium,[5] but can be forced to mix by ion bombardment during growth. The resulting material, in the case of $(GaAs)_{1-x}Ge_{2x}$ or $(GaSb)_{1-x}Ge_{2x}$, is a metastable, crystalline, substitutional alloy with a lifetime at room temperature of order 10^{29} years.[6] The fundamental energy band gap of $(GaAs)_{1-x}Ge_{2x}$ has been determined from optical-absorption measurements and shows a nonparabolic V-shaped bowing as a function of alloy composition x (Ref. 7). A V-shaped band gap cannot be explained using the conventional virtual-crystal approximation, which gives approximately parabolic bowing. This V-shaped bowing is explained, however, with a zinc-blende-to-diamond, order-disorder phase transition.[8]

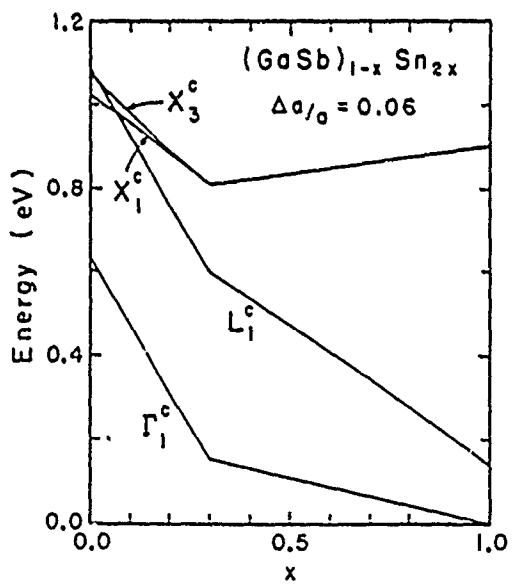


FIG. 1. Predicted band gaps at points Γ, L, and X versus alloy composition for $(GaSb)_{1-x}Sn_{2x}$. Kinks are seen in the Γ and L levels and the level at point X bifurcates at the assumed critical composition of Newman's zinc-blende-to-diamond phase transition, $x_c \simeq 0.3$. The gap is direct for all compositions, ranges from $\simeq 0.6$ to zero and decreases slowly as a function of composition from 0.15 eV to zero for compositions greater than the critical composition.

A theory for this transition has been developed by Newman *et al.*[8–10] and applied to $(GaAs)_{1-x}Ge_{2x}$. As seen in Fig. 1, where the theory is evaluated for the conduction-band minima near points Γ, L, and X for $(GaSb)_{1-x}Sn_{2x}$, the fundamental band gap exhibits a V-shaped bowing as a function of composition, with a kink at the critical composition x_c. This theory also gives smaller gaps than those of the conventional virtual-crystal approximation.

In this paper we apply this theory to the entire class of $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ alloys involving all possible combinations of Al, Ga, In, P, As, Sb, Si, Ge, and Sn, and we predict the energy band edges for these new metastable materials as functions of alloy composition x. We also establish general rules for understanding the chemical trends in the band gaps and for choosing a metastable $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ alloy with a desired energy band gap.

II. THEORY

The central idea of the present work is that all of the $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ metastable alloys should exhibit an order-disorder transition from an ordered zinc-blende structure (in which cations "know" which sites are supposed to be cation sites) to the disordered diamond structure in which there is no distinction between anion and cation sites. The critical composition x_c at which this tran-

 ©1985 The American Physical Society

sition occurs depends on the growth conditions of the alloy.

In developing a theory of the electronic structures of these alloys, we must remember that very little is presently known about these new and interesting materials. Many of the metastable alloys have not yet been grown; in most cases, satisfactory growth conditions are not yet known; and it is not yet definitely known if any of the Greene alloys other than $(GaSb)_{1-x}Ge_{2x}$ exhibits the order-disorder transition [which should be detected in x-ray diffraction as the disappearance of the (200) zinc-blende spot as x approaches x_c from below].[11] These facts are important in defining the nature of the theory that is appropriate at this time; it should be global and simple, rather than detailed and excessively quantitative. With this in mind, we assume both that all of the Greene alloys exhibit the Newman *et al.* transition, and that there exist growth conditions that will result in a critical composition $x_c=0.3$, the value appropriate for the two alloys grown to date by Greene and co-workers: $(GaAs)_{1-x}Ge_{2x}$ and $(GaSb)_{1-x}Ge_{2x}$ (x_c is probably experimentally adjustable).[12] We then predict the band structures (as functions of alloy composition x) of the remaining $(A^{III}B^{V})X^{IV}_{2x}$ metastable alloys with the intent of determining which alloys are likely to exhibit interesting and useful electronic structures—thereby targeting specific alloys for priority growth. Thus, we present these calculations in order to predict which materials are most likely to be interesting, rather than pretending to specify the band structures with any precision.

A. Order-disorder transition

The order-disorder transition involves a change of symmetry from the zinc-blende structure to the diamond-crystal structure. In this transition, the distinction between anion and cation sites is lost. The relevant order parameter is:[9]

$$M(x)=\langle P_{III}\rangle_{cation}-\langle P_{III}\rangle_{anion}, \tag{1}$$

where we imagine a zinc-blende lattice with sites labeled nominally "cation" and "anion," and $\langle P_{III}\rangle_{cation}$ is the average over all the lattice sites of the probability that a column-III atom occupies a nominal cation site. Thus $M(x)$ is proportional to the average electric dipole moment per unit cell. The order parameter depends on the growth conditions (e.g., substrate temperature, ion-bombardment energy) as well as on the composition x. For a completely ordered zinc-blende alloy, in which all column-III (column-V) atoms occupy nominal cation (anion) sites, we have $M=1-x$. If all the cations are on anion sites and the anions are on cation sites, we have merely mislabeled the nominal lattice and the order parameter is $x-1$. For the metastable ordered phase ($x<x_c=0.3$), we have $0<|M(x)|<1-x$. For the disordered diamond phase ($x>x_c$), we have $M=0$.

The theoretical problem posed by the Greene alloys is that of predicting the electronic structure of metastable alloys which are described by the order parameter $M(x)$. Thus, we must first execute a nonequilibrium phase-transition [i.e., of $M(x)$] and then calculate the changes of the electronic structure as the alloys (with different composition x) undergo the order-disorder transition. Newman showed that this formidable problem could be solved by breaking it into four connected parts: (i) an equilibrium phase-transition theory of the order parameter $M(x)$, based on a three-component "spin"-Hamiltonian model similar to the Blume, Emery, Griffiths model[13] of He^3-He^4 solutions. [Spin-up, spin-down, or zero at a site in $(GaAs)_{1-x}Ge_{2x}$ signifies occupation of that site by Ga, As, or Ge, respectively.] (ii) Introduction of the nonequilibrium character of the alloys by eliminating those equilibrium phases that cannot be reached due to growth conditions (e.g., phase separation, which occurs at equilibrium, is prevented because characteristic growth times are small in comparison with the time required for the phases to diffuse apart); (iii) mutual elimination of two unknown parameters of the spin-Hamiltonian model, i.e., a spin-coupling constant J and an effective growth temperature T, in favor of one empirical parameter, the critical composition x_c;[14] and (iv) evaluation of the electronic structure using a modified virtual-crystal approximation and a tight-binding model[15] whose matrix elements depend parametrically on the order parameter $M(x;x_c)$. Thus, in the Newman approach there are two Hamiltonians: (i) a spin-Hamiltonian for treating the order-disorder transition and for calculating the order parameter $M(x;x_c)$ and (ii) an empirical tight-binding Hamiltonian—that depends parametrically on $M(x;x_c)$—for calculating the electronic structure.

B. Spin-Hamiltonian model

Newman *et al.* have shown that a III-V compound semiconductor such as GaAs can be modeled in a spin-Hamiltonian language as an "antiferromagnet" where spin-up or spin-down on a site represents occupation by a group-III atom or a group-V atom, respectively. Thus GaAs, with alternating Ga and As atoms, in this language, is an "antiferromagnet." The "magnetization" is proportional to the net electric dipole moment per unit cell, Eq. (1), and for zero-temperature GaAs at equilibrium, equals unity. In metastable $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ alloys, such as $(GaAs)_{1-x}Ge_{2x}$, occupation of a site by a column-IV atom such as Ge is represented by "spin" zero. If the Ge were to occupy both anion and cation sites without disturbing the occupation of these sites by Ga and As, then the order parameter would be $M(x)=1-x$. However, M is not $1-x$ because Ge (spin zero) dilutes the "magnetization" $M(x;x_c)$ of this "antiferromagnet," by removing nonzero "spins" at various sites, until there is insufficient "spin-spin" interaction for an average site to "know" it should have spin-up or spin-down. With a sufficient concentration x of dilutants (that depends on temperature), the "magnetization" vanishes, and the system undergoes a phase transition, from an "antiferromagnetic" zinc-blende state with $M\neq 0$ to an "unmagnetized" phase ($M=0$). That is, as Ge dilutes GaAs, an average cation site is no longer fully surrounded by As atoms and no longer feels electronically compelled to be occupied by a Ga atom rather than an As atom. The average electric dipole moment $M(x)$ of the ordered zinc-blende phase decreases and the system undergoes a transition from an ordered zinc-blende phase in which Ga atoms preferentially

occupy nominal cation sites to a disordered ($M=0$) diamond phase in which there is no distinction between anion and cation sites. Newman constructed a spin-Hamiltonian model of this order-disorder transition. The important physical parameter of this Hamiltonian is a nearest-neighbor spin coupling (which is related to energies of interaction of the pairs of atoms V-V, III-III, and III-V). The Hamiltonian, when treated in a mean-field approximation, yields the following equation for the order parameter $M(x;x_c)$:

$$\tanh[(M/(1-x_c)]=[M/(1-x)] \quad . \tag{2}$$

where x_c is the critical composition of the order-disorder transition.

C. Tight-binding Hamiltonian

The electronic structure calculations are based on an empirical, ten-band, second-nearest-neighbor, tight-binding theory, which employs an sp^3s^* basis at each site of the zinc-blende lattice. The on-site and nearest-neighbor matrix elements of this model have been obtained previously by Vogl *et al.*,[15] who fit the known band structures of many III-V compounds and group-IV semiconductors. The Vogl matrix elements are augmented by one or two second-neighbor parameters[16] (see Table I) in order to obtain a better fit to the band structures of these semiconductors at the L point of the Brillouin zone. (The Vogl model was designed to fit the conduction-band structures well near points Γ and X.) The on-site matrix elements for these many semiconductors exhibit manifest chemical trends that depend only on the atomic energies of the atom on the site, to a good approximation. The off-diagonal nearest-neighbor matrix elements are inversely proportional to the square of the bond length d, according to the rule of Harrison *et al.*[17] For our purposes the important physical parameters of the tight-binding Hamiltonian are the on-site energies of the column-III, -IV, and -V atoms, which we shall interpolate using a generalized virtual-crystal approximation.[9] The on-site matrix elements are interpolated according to Eq. (3), as are Vd^2, where V is the off-diagonal matrix elements and d is the bond length of the alloy predicted by Vegard's law:[18] $d(x)=(1-x)d_{\mathrm{III-V}}+xd_{\mathrm{IV}}$.

TABLE I. Second-neighbor parameters. Note here that $\epsilon(p_xa,p_ya')=\epsilon(p_xc,p_yc')$ and $\epsilon(sa,p_xa)=\epsilon(p_xc,sc')$. See Ref. 16 for details.

Semiconductor	$\epsilon(sa,p_ya')$	$\epsilon(p_xa,p_ya')$
AlP	1.990	0.000
AlAs	1.830	−0.876
AlSb	0.101	0.000
GaP	0.641	0.000
GaAs	0.464	0.000
GaSb	0.688	0.000
InP	0.368	0.000
InAs	0.187	0.000
InSb	0.107	0.000
Si	0.000	0.146
Ge	0.157	0.000
Sn	0.000	0.056

We expect these $(A^{\mathrm{III}}B^{\mathrm{V}})_{1-x}X^{\mathrm{IV}}_{2x}$ alloys to satisfy adequately the Onodera-Toyozawa[19] criterion for an "amalgamated" electronic spectrum, since the variations in on-site diagonal matrix elements are small in comparison with nearest-neighbor transfer matrix elements.[20] Therefore, we expect them to have relatively well-defined band structures which can be described (in a first approximation) by a mean-field theory of the virtual-crystal type. They cannot be treated with the ordinary virtual-crystal approximation, however, because (in the disordered "diamond" phase, in particular) they contain many antisite atoms (e.g., a column-III atom on a nominal anion site)—and the usual virtual-crystal approximation does not allow for antisite atoms. We circumvent this problem by using the generalized virtual-crystal approximation,[9] which has virtual anions and cations such that the virtual cation is (schematically):

$$[(1-x+M)/2]A^{\mathrm{III}}+[(1-x-M)/2]B^{\mathrm{V}}+xX^{\mathrm{IV}} \quad . \tag{3}$$

Here, A^{III}, X^{IV}, and B^{V} represent the column-III, -IV, and -V atoms, and $M(x;x_c)$ is the order parameter (1) of the order-disorder transition, obtained by solving Eq. (2).

III. RESULTS

The energies of the band edges (relative to the valence-band maximum, which is defined to be the zero of energy) are given in Fig. 1 for $(GaSb)_{1-x}Sn_{2x}$. Corresponding results for all possible $(A^{\mathrm{III}}B^{\mathrm{V}})_{1-x}X^{\mathrm{IV}}_{2x}$ alloys are given in Figs. 2–4. The Γ conduction-band minimum occurs at $\mathbf{k}=(0,0,0)$ in the band structure. The edges labeled Δ and Λ refer to the conduction minima near the $(1,0,0)$ and $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ points, respectively (i.e., near points X and L).[21] For $\mathbf{k}$ at the X point of the Brillouin zone, the conduction-band edge actually bifurcates as a function of alloy composition at the critical composition x_c, producing both an X_1 and an X_3 minimum in the zinc-blende (ordered) phase for $x<x_c$, but only one minimum for $x>x_c$ in the diamond (disordered) phase. This bifurcation is reflected in the dependence of the minima along the Δ line as functions of composition x (see Fig. 4), because these minima lie at wave vectors near point X. The relative minimum at point Γ, when plotted as a function of composition x, exhibits a kink at x_c, as does the band edge at the L point. The minimum in the Λ direction reflects the kinked behavior of the nearby L point.

In addition to the dependences on alloy composition x, there are discernible trends depending on the positions of the atoms in the Periodic Table. To facilitate quantification of these trends, we define an effective average atomic number:

$$\langle Z\rangle=xZ_{\mathrm{IV}}+(1-x)(Z_{\mathrm{III}}+Z_{\mathrm{V}})/2 \, , \tag{4}$$

where, for example, Z_{III} is the atomic number of the column-III atom. Figure 5 shows that the Γ, Δ, and Λ band edges tend to decrease in energy with increasing

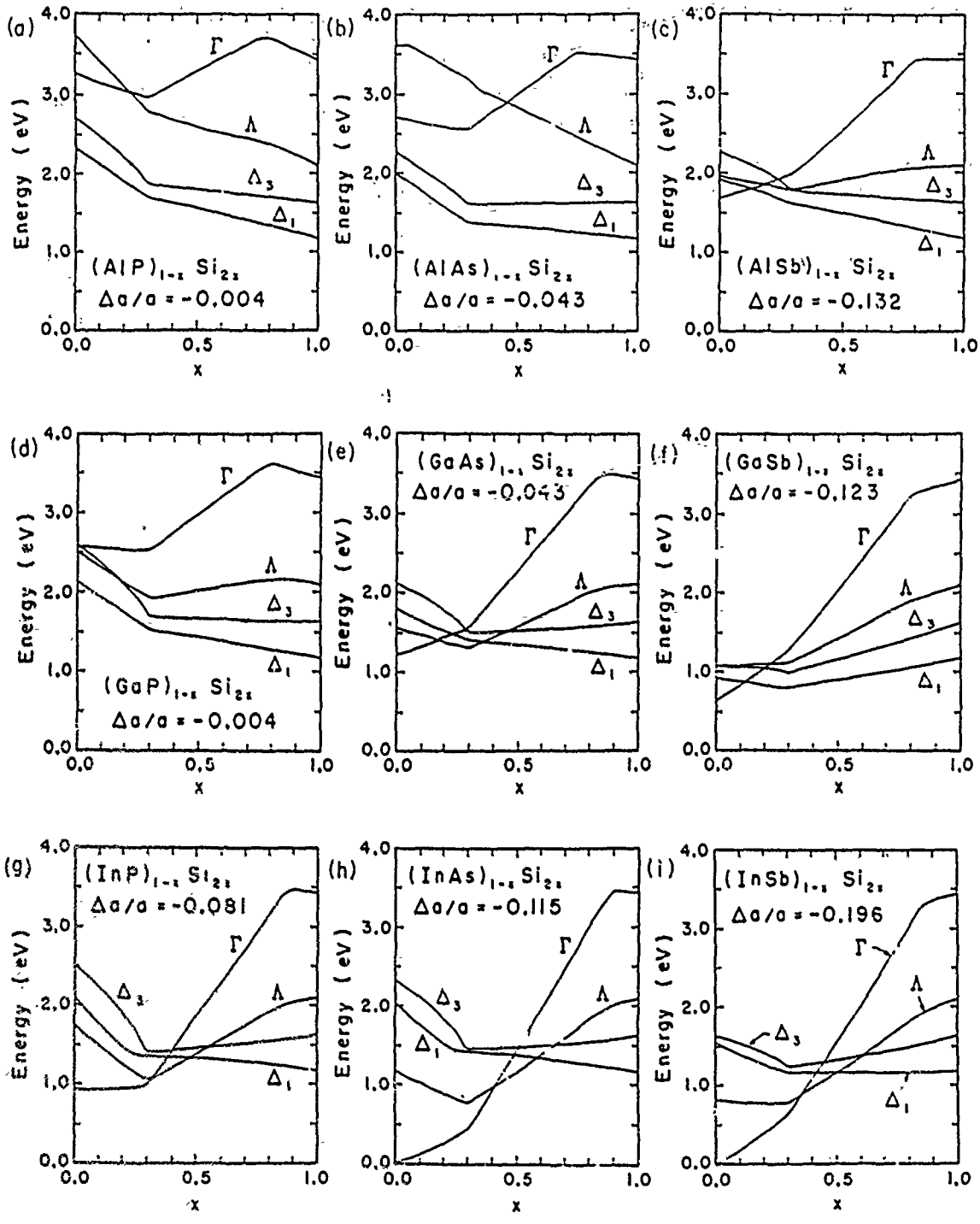


FIG. 2. Predicted band gaps of $(A^{III}B^{V})_{1-x}Si_{2x}$ alloys versus x, for the following III-V compounds: AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, and InSb. Lattice mismatches, defined by Eq. (5), are shown. Kinks can be seen in the Γ, Λ, and Δ levels at the assumed critical composition $x_c=0.3$. The Δ minimum generally lies some distance from the X point in our tight-binding model, so the strict bifurcation at the X point is not clearly visible. The kinks near $x=1$ are due to a crossing of the Γ_c^{15} and Γ_c^{1} levels.

$\langle Z \rangle$, with Γ decreasing most rapidly and Δ decreasing least rapidly with $\langle Z \rangle$. This trend can be exploited for example, to find metastable alloys with small fundamental band gaps for possible applications in infrared photography. The smaller gaps are associated with large average atomic numbers. Hence $(GaSb)_{1-x}Sn_{2x}$, with average atomic numbers ranging from 36.5 to 50, should be an interesting small-band-gap material, provided its electronic transport properties can be made suitable for device applications.

Predicted band gaps of the metastable zinc-blende-diamond Greene alloys fabricated from Al, Ga, In, P, As, Sb, Si, Ge, and Sn are shown in Figs. 2–4. General trends follow those of the prototypical alloy

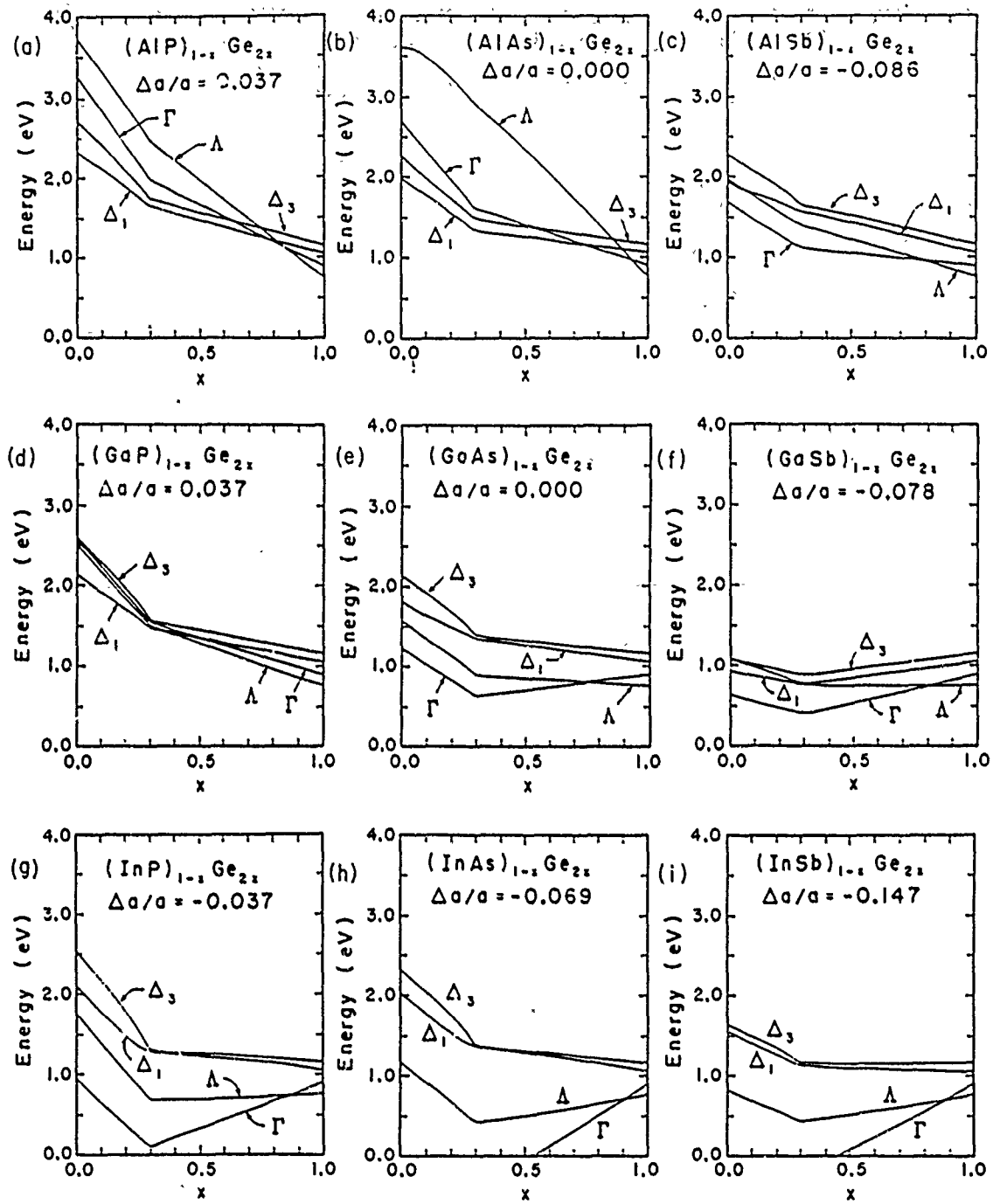


FIG. 3. Predicted band gaps of $A^{III}B^{V})_{1-x}Ge_{2x}$ alloy versus x, for the following III-V compounds: AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, and InSb. Lattice mismatches, defined by Eq. (5), are shown. Kinks can be seen in all levels, at the assumed critical composition $x_c=0.3$. For some alloys, notably $(InP)_{1-x}Ge_{2x}$ for $x<0.4$ and $(InAs)_{1-x}Ge_{2x}$ for $x<0.5$, the Δ minimum occurs at the X point in our tight-binding model and the strict bifurcation at point X is clearly visible.

$(GaSb)_{1-x}Sn_{2x}$, shown in Fig. 1. All alloy band gaps exhibit kinks at x_c as a function of composition. There is always at least one kink in the minimum conduction-band edge at $x=x_c$, due to the phase transition. This kink is not associated with a crossing of the band edges, although these types of effects can also be seen at other compositions. For example, in $(InP)_{1-x}Ge_{2x}$ (Ref. 22) at $x=0.85$, the conduction band at Γ crosses with Λ and the alloy goes from being a direct-gap semiconductor to one with an indirect gap.

The alloys with the smallest lattice mismatches

$$\Delta a/a=(a_{IV}-a_{III\text{-}V})/a_{IV} \tag{5}$$

are especially interesting. We focus primarily on alloys with $\Delta a/a<0.07$. Values of $\Delta a/a$ are given in each figure.

Since the details of the band gaps for these alloys depend on the constituents, we summarize details below fig-

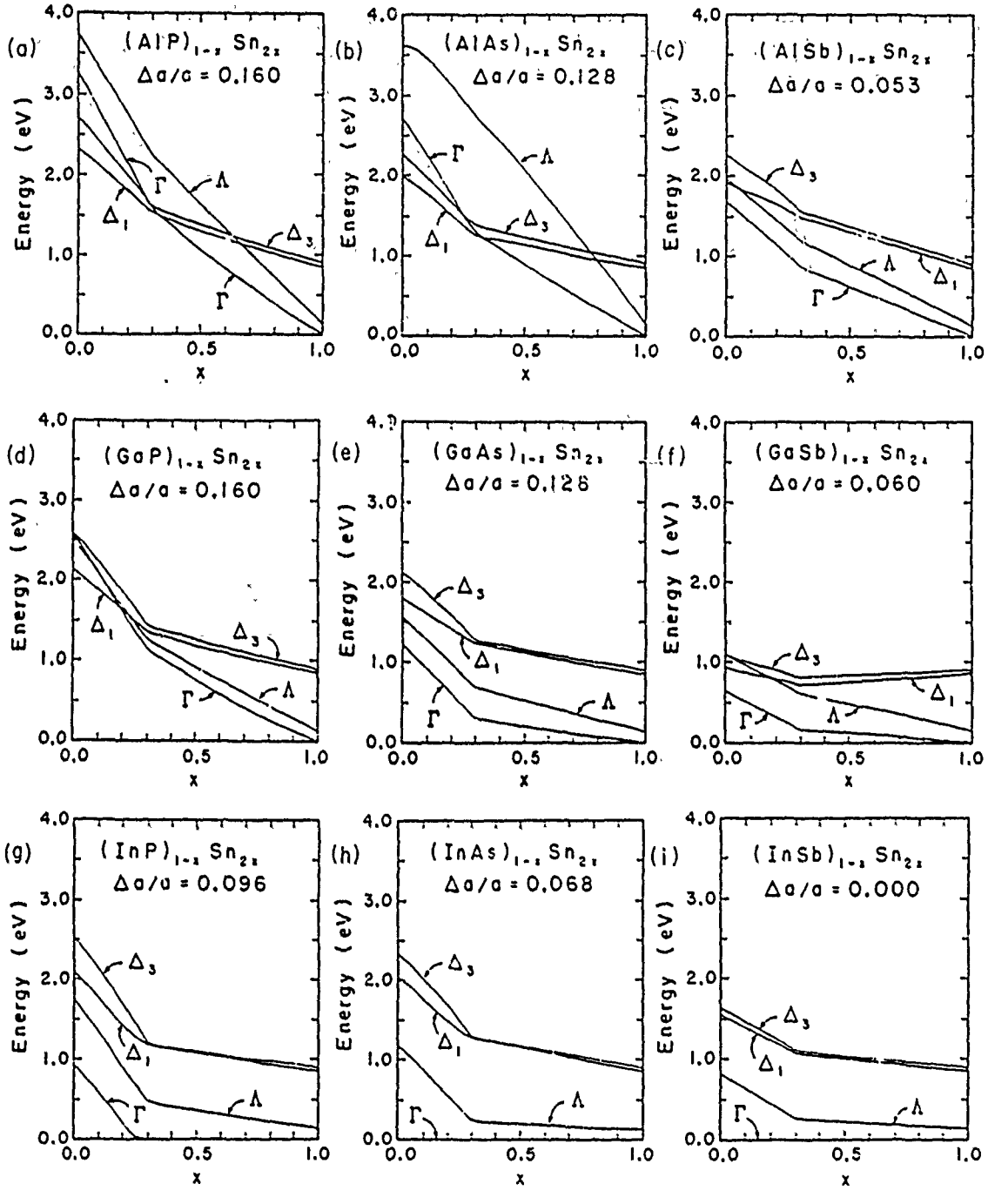


FIG. 4. Predicted band gaps of $(A^{III}B^{V})_{1-x}Sn_{2x}$ alloys versus x, for the following III-V compounds: AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, and InSb. Lattice mismatches, defined by Eq. (5), are shown. Kinks can be seen in all levels, at the critical composition $x_c = 0.3$. For some alloys, notably $(InP)_{1-x}Sn_{2x}$ for $x < 0.6$ and $(InAs)_{1-x}Sn_{2x}$ for $x < 0.5$, the Δ minimum occurs at the X point in our tight-binding model and the strict bifurcation at point X is clearly visible.

ure by figure. Figure 2 displays predicted band edges for zinc-blende materials combined in metastable alloys with Si. Those with the smallest lattice mismatches are $(AlP)_{1-x}Si_{2x}$ ($\Delta a/a = -0.004$), $(AlAs)_{1-x}Si_{2x}$ (-0.043), $(GaP)_{1-x}Si_{2x}$ (-0.004), and $(GaAs)_{1-x}Si_{2x}$ (-0.043).[23] Thus, of this class of well-lattice-matched alloys, one is restricted to materials with $\langle Z \rangle \leq 23$. The fundamental band gaps of these alloys vary from 1.17 eV for Si to 2.5 eV for ordinary AlP.[23] These gaps tend to have only one kink, at the critical composition $x = x_c$, because the fundamental gap, like that of Si, is along the Δ_1 line for all x, and does not cross Γ or Λ [the exception being $(GaAs)_{1-x}Si_{2x}$ for which we find crossings from Γ to Λ to Δ as a function of increasing composition]. The kink in Γ for $x \simeq 0.8$ is due to mixing of this level[23] and a Γ_c^{15} level not displayed (Si has $\Gamma_c^{25'} < \Gamma_c^{2}$). In contrast to the small-lattice-mismatched materials, the heavily strained alloys (see the last row of Fig. 2), all show multiple band-

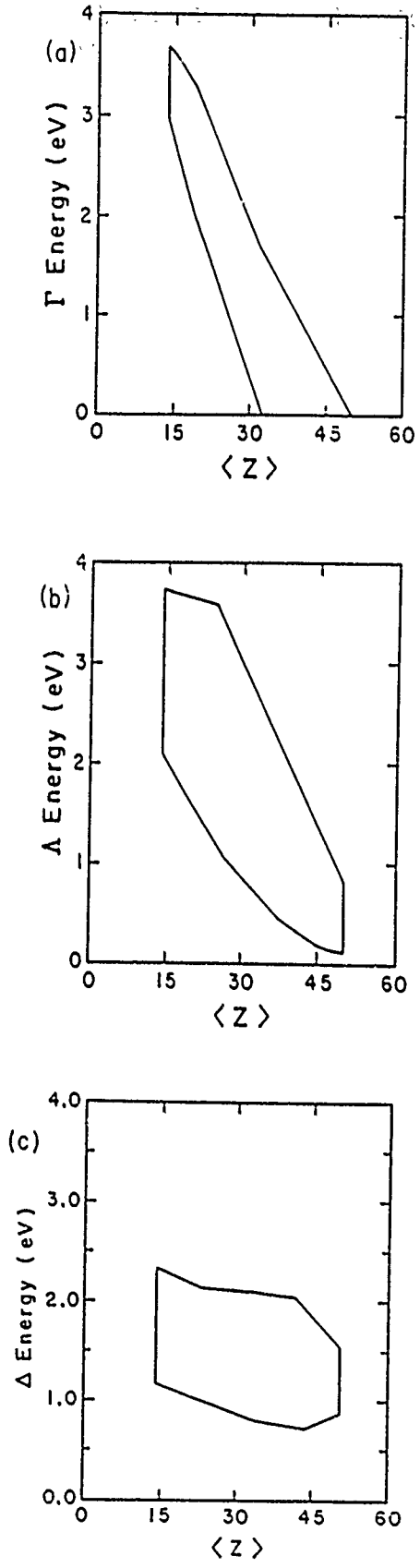


FIG. 5. Trends of the (a) Γ, (b) Λ, and (c) Δ band edges versus average atomic number ⟨Z⟩. The relevant energies for the $(A^{III}B^{V})_{1-x}X^{IV}_{2x}$ alloys in question lie within the boxes of the figures. Hence those at Γ and Λ, in particular, tend to decrease with increasing ⟨Z⟩.

edge crossings from Γ to Λ to Δ as a function of increasing composition x.

Figure 3 gives band edges for zinc-blende materials in metastable mixtures with Ge. Those with the smallest lattice mismatches are $(AlP)_{1-x}Ge_{2x}$ ($\Delta a/a = -0.037$), $(AlAs)_{1-x}Ge_{2x}$ (0.0), $(GaP)_{1-x}Ge_{2x}$ (−0.037), $(GaAs)_{1-x}Ge_{2x}$ (0.0), and $(InP)_{1-x}Ge_{2x}$ (−0.037). In this class of alloys we are restricted to well-lattice-matched materials with $\langle Z\rangle \leq 32$. The band gaps of these alloys vary from 0.1 eV for $(InP)_{1-x}Ge_{2x}$ at $x = 0.5$ to 2.5 for ordinary AlP. The band gaps of these alloys have crossings from Δ to Γ to Λ for $(AlP)_{1-x}Ge_{2x}$ and $(AlAs)_{1-x}Ge_{2x}$ and from Γ to Λ for the others. Of the remaining alloys with larger mismatches, some, such as $(InAs)_{1-x}Ge_{2x}$ and $(InSb)_{1-x}Ge_{2x}$, have zero gap for some compositions x but, because the mismatch is larger, they may be difficult to grow.

Figure 4 presents our predictions for metastable alloys resulting from mixing zinc-blende materials with Sn. Those with the smallest lattice mismatches are $(AlSb)_{1-x}Sn_{2x}$ ($\Delta a/a = 0.053$), $(InAs)_{1-x}Sn_{2x}$ (0.068), $(GaSb)_{1-x}Sn_{2x}$ (0.060), and $(InSb)_{1-x}Sn_{2x}$ (0.0). In this class of alloys, lattice matching restricts us to materials with $32 \leq \langle Z\rangle \leq 50$. These are especially interesting materials because Sn has a zero band gap. The band gaps are predicted to be zero for the metastable alloys $(InAs)_{1-x}Sn_{2x}$ and $(InSb)_{1-x}Sn_{2x}$ for all compositions (despite the fact that the equilibrium compounds InAs and InSb have nonzero gaps[24]). All of the Sn-based metastable alloys (with small lattice mismatches) mentioned above are either direct-gap or zero-gap materials. $(GaSb)_{1-x}Sn_{2x}$ is particularly interesting, because the predicted gap varies from 0.15 eV to zero over a large range in composition, from 0.3 to 1.0. Hence, the gap is small and may not be too sensitive to fluctuations in local environment. This, along with $(InP)_{1-x}Ge_{2x}$, may be an especially good candidate for an infrared detector.[22] The remaining alloys, while covering a large range in gap size, from 2.5 eV for ordinary AlP to zero for Sn, all have large lattice mismatches, $\Delta a/a > 0.096$, and good-quality, long-lived, metastable samples of these materials may be difficult to grow.

IV. CONCLUSIONS

We have presented predictions of the energy band gaps versus alloy composition x for the Greene alloys: metastable, crystalline, substitutional alloys of III-V compounds and group-IV elemental materials. The band gaps at points Γ and L exhibit kinks and the X points bifurcate as functions of composition x, at a critical value x_c corresponding to the order-disorder transition of Newman *et al.* The V-shaped bowing offers the possibility of band gaps significantly smaller than expected on the basis of the conventional virtual-crystal approximation. Alloys with modest lattice mismatches that are predicted to have small band gaps include $(InP)_{1-x}Ge_{2x}$, $(AlSb)_{1-x}Sn_{2x}$, $(GaSb)_{1-x}Sn_{2x}$, and $(InAs)_{1-x}Sn_{2x}$. Larger band-gap alloys with several potentially interesting level crossings in the band gap include $(AlAs)_{1-x}Ge_{2x}$ and $(GaAs)_{1-x}Si_{2x}$.

ACKNOWLEDGMENTS

We gratefully acknowledge the generous support of the U.S. Army Research Office (Contract No. DAAG29-83-K-0122) and the U.S. Office of Naval Research (Contract Nos. N00014-77-C-0537 and N00014-84-K-0352). We thank J. E. Greene and S. Barnett for stimulating conversations about this work and their alloys.

[1]J. E. Greene, J. Vac. Sci. Technol. B 1, 229 (1983).

[2]See also, Z. I. Alferov, R. S. Vantanyan, V. I. Korol'kov, I. I. Mokan, V. P. Ulin, B. S. Yavich, and A. A. Yakovenko, Fiz. Tekh. Poluprovodn. 16, 887 (1982) [Sov. Phys. Semicond. 16, 567 (1982)].

[3]A. J. Noreika and M. H. Francombe, J. Appl. Phys. 45, 3690 (1974).

[4]P. Duwez, R. H. Williams, and W. Klement, Jr., J. Appl. Phys. 31, 1500 (1960) were, to our knowledge, the first to grow some of these alloys for selected compositions.

[5]V. M. Glazov and V. S. Zemskov, *Physiochemical Principles of Semiconductor Doping* (IPST, Jerusalem, Israel, 1978). See also S. I. Shah, K. C. Cadien, and J. E. Greene, J. Electron. Mater. 11, 53 (1980).

[6]K. C. Cadien, A. H. Eltoukhy, and J. E. Greene, Appl. Phys. Lett. 38, 773 (1981); Vacuum 31, 253 (1981).

[7]S. A. Barnett, M. A. Ray, A. Lastras, B. Kramer, J. E. Greene, P. M. Raccah, and L. L. Ables, Electron. Lett. 18, 891 (1982). Typographical errors in the figures have been corrected in Ref. 8.

[8]K. E. Newman, A. Lastras-Martinez, B. Kramer, S. A. Barnett, M. A. Ray, J. D. Dow, J. E. Greene, and P. M. Raccah, Phys. Rev. Lett. 50, 1466 (1983).

[9]K. E. Newman and J. D. Dow, Phys. Rev. B 27, 7495 (1983).

[10]Equation (A17) of Ref. 9 is (obviously) misprinted and should read

$$4KzQ/(k_bT)=\ln[(1-x-Q)^2-M^2]/[(1-x+Q)^2-M^2)]$$
$$+2\ln[(x-Q)/(x+Q)]\ .$$

[11]S. A. Barnett presented preliminary x-ray diffraction data for $(GaSb)_{1-x}Ge_{2x}$ indicating that this does indeed happen: Bull. Am. Phys. Soc. 29, 203 (1984). See also, S. A. Barnett, B. Kramer, L. T. Romano, S. I. Shah, M. A. Ray, S. Fang, and J. E. Greene, *Layered Structure, Epitaxy, and Interfaces*, edited by J. M. Gibson and L. R. Dawson (North-Holland, Amsterdam, 1984).

[12]The critical composition x_c is bounded above by x_p, the percolation composition. An alternative theoretical description of $(GaAs)_{1-x}Ge_x$ in terms of percolation, was given by M. I. D'yakonov and M. E. Raikn, Fiz. Tekh. Poluprovodn. 16, 890 (1982) [Sov. Phys. Semicond. 16, 570 (1982)]. However, percolation would produce a critical composition x_c near 0.6 not 0.3, as observed for $(GaAs)_{1-x}Ge_{2x}$. Preliminary x-ray diffraction studies (Ref. 11) favor the order-disorder model.

[13]M. Blume, V. J. Emery, and R. B. Griffiths, Phys. Rev. A 4, 1071 (1971).

[14]Generally, neither J nor T is known, but both can be simultaneously eliminated from the equation for $M(x;x_c)$ if the phase-transition composition x_c is known. See Ref. 9.

[15]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 353 (1983). Some parameters have been refit more accuarately to existing data or fit to more recent data in the case of GaSb [T.-C. Chiang and D. E. Eastman, Phys. Rev. B 22, 2940 (1980)]. These are $V(sa,pc)=V(sc,pa)=4.9617$, $V(s^*a,pc)=V(pa,s^*c)=4.5434$ for Ge; $V(sa,pc)=V(sc,pa)=4.2288$, $V(s^*a,pc)=V(pa,s^*c)=3.9665$ for Sn; $V(sa,pc)=4.2285$, $V(sc,pa)=5.2671$, $E(s^*,a)=8.5014$, $V(s^*a,pc)=4.7325$, $V(pa,s^*c)=4.2547$ for GaAs; $E(s,a)=-7.1256$, $E(p,a)=0.6718$, $E(s,c)=-3.7042$, $E(p,c)=2.7312$, $V(s,s)=-5.9554$, $V(x,x)=1.3546$, $V(x,y)=4.4438$, $V(sa,pc)=5.1093$, $V(sc,pa)=4.4708$, $V(s^*a,pc)=5.1609$, and $V(pa,s^*c)=4.1199$ for GaSb.

[16]We have modified the nearest-neighbor model to include second-neighbor interactions, as discussed for Ge and Si by K. E. Newman and J. D. Dow, Phys. Rev. B 30, 1929 (1984).

[17]W. A. Harrison and S. Ciraci, Phys. Rev. B 10, 1516 (1974); W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

[18]L. Vegard, Z. Phys. 5, 17 (1921).

[19]Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968).

[20]H. Holloway and L. C. Davis, Phys. Rev. Lett. 53, 830 (1984), have recently suggested that the satisfaction of the Onodera-Toyozawa criterion may not be sufficient to guarantee the validity of a virtual-crystal approximation. We disagree with many of their statements about Ref. 9. The central theoretical arguments of their paper are contradicted by recent experimental results (Ref. 11).

[21]The relative minimum of the conduction band is either at point Γ, or near (or at) points X or L, in the Δ or Λ direction.

[22]D. W. Jenkins, K. E. Newman, and J. D. Dow, J. Appl. Phys. 55, 3871 (1984).

[23]We are using a notation Γ for s-like states Γ^1, as opposed to p-like states Γ^{15}_c (c denoting the conduction band and v denoting valence band).

[24]Recall that because of ion bombardment during growth, metastable AlP, for example, has more antisite defects and thus should be different from stable equilibrium AlP. This model gives a smaller band gap for metastable II-V compound semiconductors than for the stable III-V compounds. The "magnetization" $M(x;x_c)$ is not unity for $x=0$ in this theory, i.e., the theory predicts a significant concentration of antisite defects.

Reprint from

Proceedings of the 17th
International Conference
on the Physics of Semiconductors

Edited by J.D. Chadi and W.A. Harrison

© 1985 by Springer-Verlag New York, Inc.
Printed in the United States of America.



Springer-Verlag
New York Berlin Heidelberg Tokyo

THEORY OF SCHOTTKY BARRIERS FOR III-V AND GROUP-IV SEMICONDUCTORS

Otto F. Sankey
Department of Physics, Arizona State University
Tempe, Arizona 85287
and
Roland E. Allen
Department of Physics, Texas A&M University
College Station, Texas 77843
and
John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556

A theory of Schottky barriers for III-V and Group-IV semiconductors must explain numerous experimental observations, two of which are the following: (1) The barrier height $\phi_B = E_c^s - E_F^b$ often attains nearly its final value with the deposition of only a fraction of a monolayer of metal on the semiconductor surface [1,2]. (Here E_c^s and E_F^b are respectively the energies of the conduction band edge at the surface and the Fermi level in the bulk.) This finding implies that the mechanism responsible for Schottky barrier formation requires only submonolayer coverage rather than a complete semiconductor/metal interface. It thus appears to rule out metal-induced gap states [3-6] as the dominant mechanism for most systems. (2) In many cases, more than one value of the Schottky barrier height is observed for a single semiconductor. In InP, for example, there is a "switching" of barrier heights from about 0.5 eV to nearly zero as the reactivity of the metal treatment of the surface is varied [7]. A theory that yields only a single barrier height for all metals and all surface treatments is thus unable to explain these Schottky barriers.

Here we describe a theory of Schottky barriers for III-V and Group-IV semiconductors that is in agreement with the above observations and many other features of the experimental measurements [1,2,7-14]. The theory is based on Fermi-level pinning [15] by levels associated with defects at the semiconductor/metal interface [1,8-11]. We believe that this theory is applicable to most observed barriers, but recognize, of course, that other mechanisms can be important in some cases. For example, the original Schottky mechanism of charge transfer between a metal and a semiconductor without defect states in the fundamental band gap appears to dominate for nonreactive metals on GaSe [8].

For III-V semiconductors, the states responsible for pinning the Fermi-level are associated with native defects, either antisite defects [16-18] or vacancies [17-19] at the semiconductor surface. For Group-IV semiconductors, the pinning

defects are interfacial dangling bonds [20,21] -- e.g., Si dangling bonds "sheltered" from the metal by interfacial vacancies.

In Fig. 1, results are shown for the (110) surface of InP [17]. (Our calculations employ methods introduced by Vogl et al. [22] and Hjalmarson et al. [23], together with the analytic Green's function technique [24], using the surface relaxation characteristic of III-V semiconductors [25,26]. It can be shown [20] that a defect at a free surface has virtually the same deep energy levels as a sheltered defect at a semiconductor/metal contact.) At the surface, both antisite defects -- In_P and P_{In} -- are predicted to produce deep acceptor levels at about 0.5 eV below the conduction band edge. The experimental Fermi-level pinning position for non-reactive metals on n-type InP [7,14], also shown in Fig. 1, can be explained by either of these levels. In addition, a surface phosphorous vacancy V_P is predicted to yield a shallow donor level, which appears to explain the other Fermi-level pinning position just beneath the conduction band edge observed for reactive metals on n-InP [7] and for treatments of the surface with O or Cl [8]. A shallow level is also predicted for the surface impurities S_P and Sn_{In}, and this may explain the experimental results for S and Sn treatments of the InP surface [8]. Although not shown in Fig. 1, the predictions for p-InP are also in agreement with experiment [17].

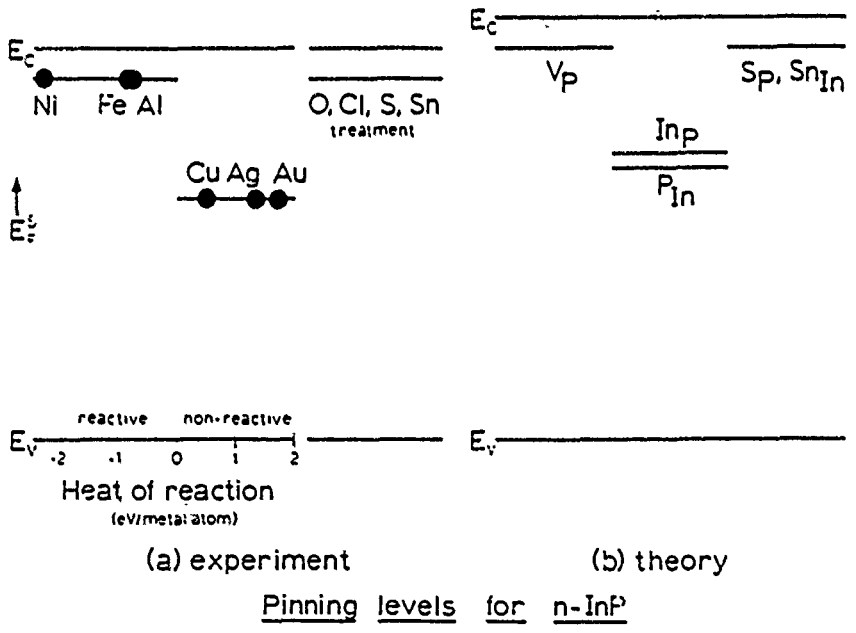


Fig. 1. Theoretical and experimental Fermi-level pinning positions for the (110) surface of InP (after Ref. 17).

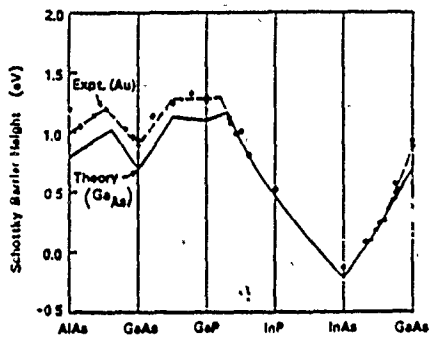


Fig. 2. Theoretical and experimental Schottky barrier heights for various III-V semiconductors and their alloys. The results for $Ga_{1-x}Al_xAs$ and $In_{1-x}Ga_xAs$ are also given in [16].

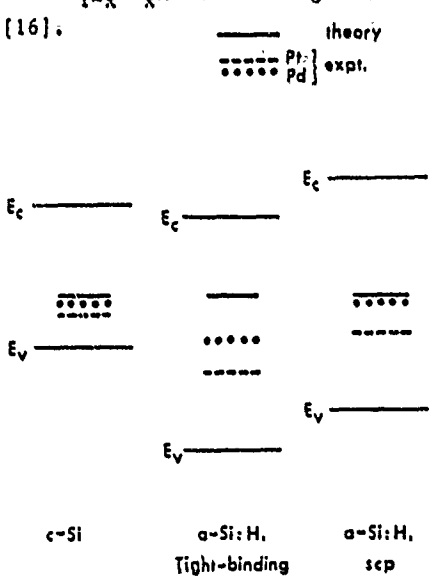


Fig. 3. Calculated Fermi-level pinning positions for c-Si and a-Si:H. The experimental positions are inferred from the measurements of [29] in conjunction with the band gap of a-Si:H of 1.8 eV and a theoretical value of E_c obtained from [28] for a-Si:H with a band gap 0.7 eV larger than that of c-Si (scp) or a virtual-zincblende approximation to the theory in [27] (tight-binding). Data for Pt and Pd contacts are from [29].

In Fig. 2, we show the Schottky barrier heights predicted by the present theory (in the virtual crystal approximation) for several III-V semiconductors and their alloys. These results are for n-type semiconductors, with the defect responsible for Fermi-level pinning taken to be the cation-on-anion site surface antisite defect -- e.g., Ga on the As site [16]. The predictions are compared with measured barrier heights for Au contacts [10]. The agreement between theory and experiment is fortuitously good, in view of the several tenths of an eV uncertainty in the theory. It can be seen, however, that the trends with alloy composition are very well described by the theory.

In Fig. 3, results are shown for both crystalline Si (c-Si) and hydrogenated amorphous Si (a-Si:H). The theory for c-Si, which is described elsewhere [20], provides a quite satisfactory description of numerous experimental observations.

Extensions of the theory to Ge, $Ge_{1-x}Si_x$ alloys, and diamond were also found to be in good agreement with experimental data [21]. Here we present new results for a-Si:H. Details of the theory will be presented in a longer article, but the essential idea was to extend the c-Si prediction for the Si dangling-bond energy to a-Si:H by employing the tight-binding theory of Papaconstantopoulous and Economou [27,28] for this material. For example, we treated hydrogenated amorphous Si as a virtual zincblende ternary alloys, $Si(Si_{1-x}H_{4x})$, deducing tight-binding parameters for Si and $Si_{1-x}H_{4x}$ from the parameters of Ref. [27]. These parameters were then used to compute the band edges of a-Si:H relative to those of c-Si and the deep level of a Si

bond in a-Si:H surrounded by a local environment similar to that of an interface with $NiSi_2$. Our results are given in Fig. 3, and agree with the data.

In conclusion, the present theory provides a satisfactory description of Schottky barriers for a number of III-V semiconductors and their alloys (both n- and p-type), and for Si, Ge, $Ge_{1-x}Si_x$ alloys, diamond and a-Si:H.

ACKNOWLEDGMENT: We thank the Office of Naval Research for their generous support of this work (Contract Nos. N00014-82-K-0447 and N00014-84-K-0352).

REFERENCES

[1] W. E. Spicer, I. Lindau, P. R. Skeath, C. Y. Su, and P. W. Chye, Phys. Rev. Lett. 44, 420 (1980).

[2] P. S. Ho and G. W. Rubloff, Thin Solid Films 89, 433 (1982). G. W. Rubloff, Surf. Sci. 132, 268 (1983).

[3] V. Heine, Phys. Rev. 138, A1689 (1965).

[4] S. G. Louie and M. L. Cohen, Phys. Rev. B13, 2461 (1976).

[5] J. Sanchez-Dehesa and F. Flores, Solid State Commun. 50, 29 (1984).

[6] J. Tersoff, Phys. Rev. Lett. 52, 465 (1984).

[7] R. H. Williams, V. Montgomery, and R. R. Varma, J. Phys. C11, 2735 (1978).

[8] R. H. Williams, Surf. Sci. 132, 122 (1983) and references therein.

[9] W. Monch and H. Gant, Phys. Rev. Lett. 48, 512 (1982).

[10] W. Monch, Surf. Sci. 132, 92 (1983) and references therein.

[11] H. H. Weider, Surf. Sci. 132, 390 (1983).

[12] L. J. Brillson, Surf. Sci. Rep. 2, 123 (1982).

[13] G. Ottaviani, K. N. Tu, and J. W. Meyer, Phys. Rev. B24, 3344 (1981).

[14] C. A. Mead and W. G. Spitzer, Phys. Rev. A134, 713 (1964).

[15] J. Bardeen, Phys. Rev. 71, 717 (1947).

[16] R. E. Allen and J. D. Dow, Phys. Rev. B25, 1423 (1982).

[17] J. D. Dow and R. E. Allen, J. Vac. Sci. Tech. 20, 659 (1982).

[18] R. E. Allen, R. P. Beres, and J. D. Dow, J. Vac. Sci. Tech. B1, 401 (1983).

[19] M. S. Daw and D. L. Smith, Phys. Rev. B20, 5150 (1979).

[20] O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1984).

[21] O. F. Sankey, R. E. Allen, and J. D. Dow, J. Vac. Sci. Tech. (in press).

[22] P. Vogl, H. P. Hjalmarson, J. D. Dow, J. Phys. Chem. Solids 44, 365(1983).

[23] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).

[24] R. E. Allen, Phys. Rev. B20, 1454 (1979) and to be published.

[25] A. Kahn, E. So, P. Mark, and C. B. Duke, J. Vac. Sci. Tech. 15, 580 (1978).

[26] S. Y. Tong, A. R. Lubinsky, B. J. Mrstik, and M. A. Van Hove, Phys. Rev. B17, 3303 (1978).

[27] D. A. Papaconstantopoulous and E. N. Economou, Phys. Rev. B24, 7233 (1981).

[28] See also, D. P. Divincenzo, J. Bernholc, and M. H. Brodsky, Phys. Rev. B28, 3246 (1983) for calculations of the band edges of a-Si:H.

[29] R. J. Nemanich, M. J. Thompson, W. B. Jackson, C. C. Tsai, and B. L. Stafford, J. Vac. Sci. Tech. B1, 519 (1984).

Applications of Surface Science 22/23 (1985) 937–947
North-Holland, Amsterdam

INTERFACIAL DEEP LEVELS RESPONSIBLE FOR SCHOTTKY BARRIER FORMATION AT SEMICONDUCTOR/METAL CONTACTS

John D. DOW

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA

Otto F. SANKEY

Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

and

Roland E. ALLEN

Department of Physics, Texas A&M University, College Station, Texas 77843, USA

Received 27 August 1984; accepted for publication 31 October 1984

The following facts, and many others, concerning III–V (e.g., GaAs, InP) Schottky barriers can be understood in terms of Fermi-level pinning by interfacial antisite defects (sheltered by vacancies) at semiconductor/metal contacts: (i) the barrier heights are almost independent of the metal in the contact; (ii) the surface Fermi levels can be pinned at sub-monolayer coverages and the pinning energies are almost unaffected by changes of stoichiometry or crystal structure; (iii) the Schottky barrier height for n-InP with Cu, Ag, or Au is ≈ 0.5 eV, but changes to ≈ 0.1 eV when reactive metal contacts (Fe, Ni, or Al) are employed because the antisite defects are dominated by P vacancies; and (iv) the dependence on alloy composition for alloys of AlAs, GaAs, GaP, InAs, and GaAs is extremely complex – owing to the dependence of the binding energy for the cation-on-anion-site deep level on alloy composition. Fermi-level pinning by Si dangling bonds at Si/transition-metal silicide interfaces accounts for the following facts: (i) the barrier heights are independent of the transition-metal, to within ≈ 0.3 eV; (ii) on the 0.1 eV scale there are chemical trends in barrier heights for n-Si, with the heights decreasing in the order Pt, Pd, and Ni; (iii) barriers form at low metallic coverage, (iv) barrier heights are independent of silicide crystal structure or stoichiometry to ± 0.1 eV; and (v) the barrier heights for n-Si and p-Si add up to approximately the energy of the band gap.

1. Introduction

When a metal is deposited on a semiconductor surface, a potential barrier to electron motion is formed, which prevents the flow of electrons between the metal and the semiconductor. The physics governing the formation of this Schottky barrier is controversial even today. Here we present theoreti-

0378-5963/85/$03.30 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

cal calculations which support the notion that Schottky barriers are commonly (but not exclusively [1]) formed as a result of "Fermi-level pinning" by deep trap states associated with defects at the semiconductor surface.

The basic idea of Fermi-level pinning was enunciated by Bardeen [2] in 1947, and is most easily described for the limit of a degenerate n-type semiconductor in contact with a metal. The Fermi energies of the semiconductor, the semiconductor surface, and the metal must align in electronic equilibrium. At zero temperature, the Fermi level of the semiconductor lies almost at the conduction band edge (more precisely, at the donor level), and lines up with the Fermi level of the metal. The Fermi energy of the semiconductor's surface, however, can lie deep in the fundamental band gap if there are deep impurity levels in the gap. In this case of sufficient concentration of deep levels in the gap, the deep levels determine and "pin" the Fermi energy of the surface, which does not align with the bulk semiconductor's Fermi energy if the valence band maxima of the bulk and the surface are assumed to be at the same energy. Hence, the semiconductor and its surface are not in electronic equilibrium when the valence band maxima align. As a result, carriers must diffuse in order to bring the surface into electronic equilibrium with the bulk semiconductor and the metal: a surface dipole must build up, and the bands must bend near the surface to align the Fermi energies of the bulk and the surface. This results in a Schottky barrier (see fig. 1). Bardeen, in his Fermi-level pinning paper, left open the possibility that the deep levels responsible for the pinning might be either intrinsic (e.g., surface states) or extrinsic. Spicer and co-workers [3] have championed the idea that *native defects* produced during the formation of the semiconductor/metal contact pin the Fermi energy.

In this Fermi-level pinning model, one can estimate the Schottky barrier height for an n-type semiconductor by first determining the defect responsible for the pinning and then calculating the difference in energy between

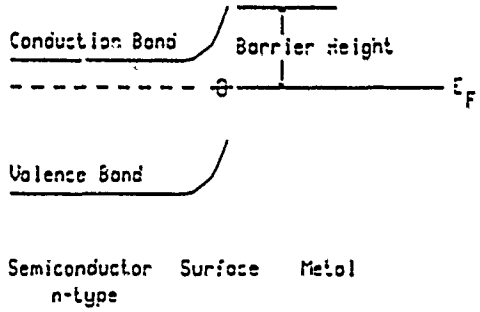


Fig. 1. Schematic illustration of the Bardeen model of Fermi-level pinning by a semiconductor-surface defect deep level (denoted by an open circle with a bar through it).

the neutral defect's lowest unfilled deep level in the gap and the conduction band edge. (For p-type material, the highest filled level in the gap pins the Fermi level.) Hence the problem of calculating the Schottky barrier height is reduced to the equivalent problem of computing the binding energy of a deep level. (See fig. 1.)

2. Deep levels in the bulk

To understand the physics of deep levels at metal/semiconductor contracts, one must first comprehend the basic physics of deep impurities in the bulk of a semiconductor. An impurity level, by current definition [4], is "deep" if that level originates from the central-cell defect potential of the impurity (as opposed to originating from the long-ranged Coulombic tail of the defect potential, as for "shallow" levels). In covalently bonded semiconductors, sp^3-bonded substitutional defects have typically *four* deep levels near the fundamental band gap and an infinite number of shallow levels. The infinite number of shallow levels is associated with the fact that the Coulomb potential has an infinite number of bound states, and the four deep levels are due to there being one s-like and three p-like orbitals for an sp^3-bonded defect. In the bulk of a tetrahedral semiconductor, the three p-states are degenerate, forming a T_2-symmetric deep level, and the s-state gives rise to an A_1 level.

The four deep levels need not all lie within the fundamental band gap, however. In fact, it is rare that all four do. Indeed, a "shallow-impurity" is one for which all of its deep levels lie *outside* the fundamental band gap (fig. 2). A "deep impurity" is an impurity that produces at least one deep level in the band gap. The issue of whether a deep level lies within the gap or not is a quantitative one: if the host bands are broad enough and the fundamental band gap is narrow enough, then the bands are likely to cover up all of the deep levels, making them resonant with the host bands. Hence, narrow-gap semiconductors tend to have relatively fewer "deep impurity" centers (with levels in the gap) than large band-gap materials.

The basic physics of deep levels is illustrated schematically in fig. 3 for the case of an N impurity replacing P in the bulk of GaP. For simplicity we consider only the A_1 or s-like deep state of the defect. First consider atomic Ga and P, which, when combined into a molecule, form bonding and antibonding levels. The bonding–antibonding splitting is of order $v^2/(\epsilon_{Ga} - \epsilon_P)$, where v is the nearest-neighbor transfer matrix element and $\epsilon_{Ga} - \epsilon_P$ is the energy denominator resulting from perturbation about the extreme tight-binding limit [5,6]. The bonding and antibonding states of the molecule are the parents of the conduction and the valence bands of the solid, respectively. If now one P atom is replaced by an N impurity atom, the N

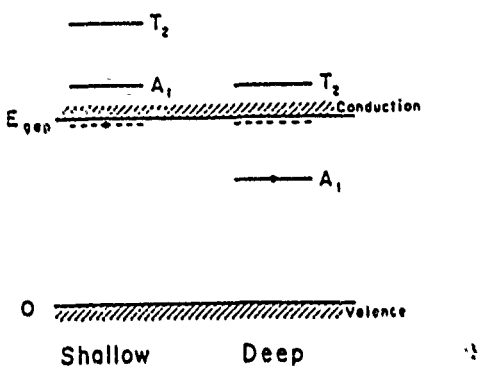


Fig. 2. Schematic illustration of the difference between "shallow" and "deep" sp^3-bonded substitutional (donor) impurities, after ref. [21]. The shallow energy levels in the band gap are dashed. The deep levels of A_1 (s-like) and T_2 (p-like) symmetry are denoted by heavy lines. In the case of a "shallow impurity" the deep levels are resonances and lie outside the fundamental band gap; for a "deep impurity" at least one deep level lies within the gap. The lowest level is occupied by an extra electron (dark circles) if the impurity has a valence one greater than the host atom it replaces (e.g., S or O on a P site in GaP).

will try to hybridize with its neighbors. However, the atomic energy of the N is ≈ 7 eV lower than the corresponding energy of the P atom it replaces (i.e., the defect potential in the central-cell is $V \approx -7$ eV). As a result, the energy denominator is ≈ 7 eV larger for N than for P, and (since v is almost the same for P and N [7]), the bonding–antibonding splitting is smaller – and the deep level lies within the band gap. For a slightly less negative value of V (i.e., a slightly more electropositive defect than N, such as S), however, the deep level is resonant with the conduction band – so that at most "shallow" states bound by the long-ranged Coulombic, $-Ze^2/\epsilon r$, part of the defect potential (neglected here) would lie in the gap. (Here Z is the impurity-host valence difference and is zero and unity for N and S, respectively, replacing a P atom in GaP.)

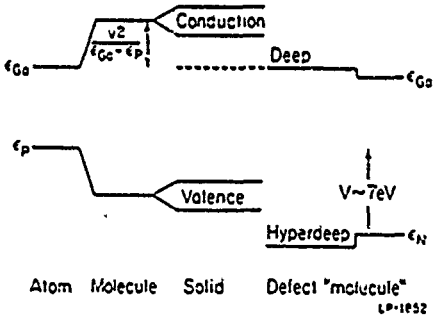


Fig. 3. Schematic illustration of the qualitative physics of deep levels, as discussed in ref. [4].

3. Deep levels of surface defects

The same basic physics holds for a defect at a surface. The reduction of the tetrahedral symmetry by the surface causes the A_1 bulk levels to shift and the T_2 deep levels to split into three orbitally non-degenerate levels. For a defect at a free surface the splittings of the T_2 level are of order 1 eV. Therefore, the free-*surface defect levels lie at substantially different energies from the bulk defect levels.*

A central question is whether the pinning defects for Schottky barrier formation are at the semiconductor's surface or in the semiconductor's bulk. There is no definitive experimental answer to this question at the present, but we believe that the defects are *at or very near* the semiconductor/metal interface *in environments that are quite similar to the environment of a defect at a free surface.* By this we mean the pinning defects are each adjacent to a vacancy or a void (or a highly electropositive atom) that "shelters" it electronically from its more distant neighbors: a defect–vacancy pair in the semiconductor's bulk has essentially the same energy levels as a defect at a surface [8] – because deep-level wavefunctions are rather localized to the shell of first-neighbors of the defect, and the main difference between a defect–vacancy pair and the same defect at a surface is that, at the surface (which can be thought of as a sheet of vacancies), some second- and more-distant neighbors are vacancies rather than atoms. (Second-neighbor effects on a deep level are rarely major.)

Our reasons for adopting this viewpoint that the pinning defects are near the semiconductor/metal interface and "sheltered" in free-surface-like environments are: (i) Fermi-level pinning can occur at sub-monolayer metallic coverages, a fact that is difficult to explain unless the relevant defects are at or near the interface; (ii) the simple bulk point defects, such as vacancies and antisites, unquestionably give qualitatively as well as quantitatively incorrect predictions for the observed behavior of Fermi-level pinning and Schottky barrier heights (e.g., the bulk antisite As_{Ga} – As on a Ga site – in GaAs cannot explain the Fermi-level pinning for n-GaAs because it produces only an occupied deep donor level in the gap, whereas an unoccupied acceptor is required to achieve Fermi-level pinning in the gap for n-type material – i.e., the next available level for an electron is the deep level, rather than the conduction band edge). In contrast, As_{Ga} at the surface produces *two* deep levels in the gap: a deep donor and a deep acceptor; (iii) without the concept of sheltering, the defect theory would be in conflict with the experimental fact that, for GaAs and some other semiconductors, the deposition of different (non-reactive) metals in a semiconductor/metal contact most often leads to the same Schottky barrier height (if the defect were in direct contact with the metal, its energy levels would be significantly altered by changing the metal). Therefore, the pinning defect must be

adjacent either to a vacancy or to a very electropositive atom (recall that electronically a vacancy is an infinitely electropositive "atom" [9]).

Spicer and co-workers [3], Wieder et al. [10], Mönch et al. [11], and Williams et al. [12] have presented numerous data which indicate that the surfaces of III–V semiconductors have Fermi levels determined or "pinned" by the deep impurity states of *native defects*. The exact mechanisms by which these defects are created are not presently understood, but it is believed that they are normally generated during the formation of the surface (e.g., by cleavage) or during the deposition of a metal contact. Indeed, the precise nature of the native defects is not presently known, and one purpose of this work is to provide a theoretical framework for identifying the "pinning defects". We shall enumerate the possible native defects, argue that the pinning levels of many complex defects are virtually identical to the pinning levels of a few simple ones, show that some simple defects can explain the observed chemical trends in Schottky barrier data for III–V semiconductors while others cannot, and propose a relatively simple and specific picture of the pinning defects.

The possible native defects are anion and cation vacancies, both types of antisite defects, anion and cation interstitials, and combinations of these. It can be shown, however, that the combination defects normally have spectra similar to the sum of their constituents' spectra [8] – and so we consider only the isolated defects. We also eliminate interstitials from consideration, because (i) interstitials are known to be very sensitive to the local environment [13] (whereas Fermi-level pinning defects are not), and (ii) in the bulk, the Group-III and Group-V atoms have been observed either on their own sites or on the antisite, but (to our knowledge) not at interstitial positions. Defects associated with the metal atoms originating from the metal of the contact are not considered because (i) for some semiconductors at least, the Schottky barrier heights are relatively independent of the metal, and (ii) for most of the semiconductors of interest, the metal atoms themselves do not produce the required deep levels in the fundamental band gap.

Thus we are left with an apparently simple problem: compute the deep levels of the vacancies and the antisite defects, and determine if these levels explain the observations. In making these calculations, however, we must recognize that this or any theory has uncertainties of order ± 0.5 eV (part of which is due to the neglect of lattice relaxation around the defect). Therefore, we do not simply compare the theory with data, but instead we (i) eliminate as many as possible of the Fermi-level pinning assignments because the theory and the data disagree by > 0.5 eV, and (ii) make our final assignments on the basis of the observed chemical trends in the Fermi-level pinning positions from one semiconductor to another.

The calculations employ an empirical tight-binding Hamiltonian [6] for the host semiconductor. Since the parameters of this Hamiltonian exhibit

chemical trends, the defect potential's matrix elements can be estimated from the trends [4]. In the localized-orbital sp^3s^* tight-binding basis, the defect potential V is diagonal (provided we assume that the lattice does not relax around the defect) with elements proportional to the differences between the atomic energies of the defect and the host atom it replaces. The deep levels of the defect are obtained by solving the secular equation

$$\det[1-(E-H_0)^{-1}V]=0\ ,$$

where H_0 is the host Hamiltonian and E is the deep level. For the calculations reported here, H_0 describes a relaxed (110) III–V surface with the 27° rigid rotation of the anions out of the surface plane, and V is a matrix simulating the central-cell potential of a defect at the surface. However, a far simpler model involving defect–vacancy pairs in the bulk or at an interface would give similar results [14]: in the simpler case H_0 would represent the sheltering vacancy and the cluster of atoms at the defect site (before the defect is introduced) and at surrounding first- and possibly more-distant-neighbor sites. The details of solving the secular equation either for a defect at the free surface [15] or for a simplified cluster model [14] have been described elsewhere.

4. Results for III–V semiconductors

The results of our calculations of the Schottky barrier heights (i.e., the binding energies of the lowest incompletely occupied one-electron level of the neutral impurity with respect to the conduction band edge) are given in fig. 4, where we have assumed that the defect responsible for Fermi-level pinning is the cation-on-anion-site antisite defect at the surface. The agreement between theory and data is strikingly excellent, and strongly supports the hypothesis that this antisite defect is responsible for the observed Schottky barrier formation. (The two vacancies and the other antisite defect fail to reproduce all the observed trends.)

This success does not mean that all Schottky barrier formation in III–V semiconductors is attributable to Fermi-level pinning by cation-on-anion-site defects. Although an antisite defect can be formed with less free energy that a vacancy [16], we believe Fermi-level pinning by vacancies has been observed for InP contacts with reactive metals [17]. Indeed, the apparent dependence of Schottky barrier height on chemical reactivity [12,17,18] can be explained in terms of chemical reactions changing the dominant defect from an antisite to a vacancy. The reactive metals combine with P making stable compounds, leave P vacancies (V_P). In InP these vacancies are predicted to yield shallow donor levels in the fundamental gap near the conduction band edge; these levels pin the Fermi energy and yield a small

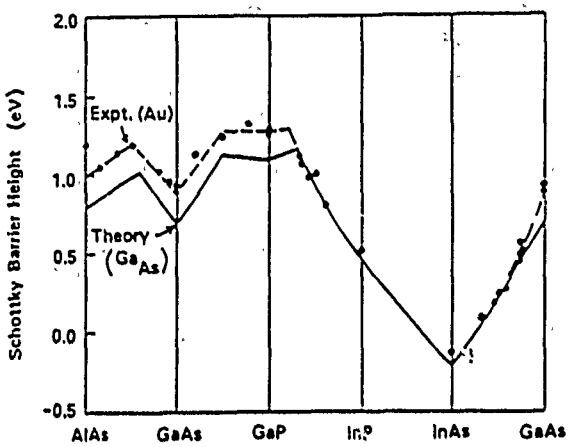


Fig. 4. Experimental (dashed) and theoretical (solid) dependence of the Schottky barrier heights of III–V semiconductor alloys with Au contacts versus alloy composition, after ref. [22]. The theory assumes Fermi-level pinning by a surface cation-on-anion-site defect.

(≈ 0.1 eV) Schottky barrier height. Thus, when reactive metals (viz., Fe, Ni, or Al) are deposited on InP, the dominant Fermi-level pinning defects appear to be P vacancies; but when non-reactive metals are deposited (viz., Cu, Ag, or Au), the antisite defect levels appear to dominate, and the barrier height is approximately 0.5 eV. Thus the theory, supplemented by the hypothesis that the reactive metals produce P vacancies, can account for the InP data.

The ability of the theory to provide a natural explanation of the dependence of Schottky barrier height on chemical reactivity is especially important, because it offers a resolution of a major controversy between the viewpoints (i) that Schottky barrier formation is due to Fermi-level pinning by defects (championed by Spicer and co-workers [3]), and (ii) that Schottky barrier formation depends critically on chemical reactivity (advocated by Brillson and associates [18]). Our own viewpoint is that both sides of the controversy are essentially correct, and that different chemical reactions produce different dominant defects and Fermi-level pinning positions.

Presently, it is not known if reactive metals do indeed produce a sufficient number of interfacial P vacancies in InP. Indeed, studies of P diffusion indicate that the diffusion rate is greater for non-reactive metals and that P concentrates at the reactive-metal/InP interface [19]!

There are many other experimental facts concerning III–V semiconductors, most of them of a detailed nature, that the Fermi-level pinning theory can explain. But rather than focus on those details in our limited space, we instead turn our attention to Si (which, being homopolar, has no antisite defects) and the question of whether Si's Schottky barriers are similar to those of the III–V's.

5. Si/transition-metal silicide Schottky barriers

The best-studied Si Schottky barriers are those with transition-metal silicides. The silicides themselves are metals created by the reaction of the transition-metals with the Si. For these systems we believe the pinning defects are Si dangling bonds.

The local defect we think is responsible for the Fermi-level pinning is similar to a bulk-Si vacancy whose four nearest-neighbor atoms (instead of all being Si) are three transition-metal atoms and one Si atoms (the one whose bond dangles into the sheltering vacancy due to a missing Si bridge atom at the Si/transition-metal silicide interface) [20]. If the transition-metal atoms were instead Si atoms, locally this defect would be a bulk-Si vacancy – with an A_1 symmetric deep level resonant with the valence band and a T_2 level in the Si band gap. Hence, to determine the physics of the Fermi-level pinning at the Si/transition-metal silicide interface, we need only understand how the bulk-Si vacancy's deep levels change as three of the vacancy's neighbors change from Si into transition-metal atoms. The change of the three neighbors from Si into transition-metal atoms can be simulated by increasing the sp^3 hybrid energies of the atoms on the transition-metal sites (for the hybrids oriented toward the vacancy) from ϵ_h (for Si) to $\epsilon_h + V$ (for transition-metal atoms), with V of order 5 eV. That is, relative to Si, the transition-metal atoms are very electropositive (electronically like vacancies). The large positive repulsive potential V on the transition-metal sites merely pushes the Si dangling bond away from the silicide and into the Si. In the process, it drives the energy of the T_2 deep level for the bulk-Si vacancy out of the fundamental band gap and into the conduction band, and brings the A_1 level up into the gap. (For $V \to +\infty$, the A_1-derived level approaches the hybrid energy ϵ_h asymptotically from below.) This level, for the neutral defect, is singly occupied by one electron, and therefore can pin the Fermi energy of either n-Si or p-Si. Hence the barrier heights for n- and p-type material add up to the band gap. Since changes of V of order 1 eV have little effect on a pinning level that asymptotically approaches ϵ_h, the theory explains why different transition-metals have the same barrier heights to within ≈ 0.3 eV, while the differences in barrier heights on the 0.1 eV scale reflect the chemical trends in V (which is proportional to the difference between the atomic energy of a transition-metal and that of Si) giving decreasing barrier heights for Pt, Pd, and Ni silicides. Moreover, since the pinning defect is localized and has properties that depend primarily on the electropositivity of the transition-metal atoms, one can understand why barriers form at low metallic coverages and have heights that are insensitive to stoichiometry or the silicide crystal structure. Thus Fermi-level pinning by dangling bonds can account for the main experimental facts concerning Schottky barrier heights at Si/transition-metal silicide interfaces.

6. Discussion

We have touched on a small subset of the many and varied experimental facts that can be explained by the Fermi-level pinning model of Schottky barrier formation. This model has been successfully applied to Schottky barriers involving Si, $Si_{1-x}Ge_x$, and diamond (with Fermi-level pinning by dangling bonds) as well as to III–V semiconductors (in which antisite defects and dangling bonds pin the Fermi level). It appears to be applicable to any covalent semiconductor which responds to contact formation by spontaneously producing a sufficient number of native defects.

Acknowledgements

We are grateful to the US Office of Naval Research and the Army Research Office for supporting this research (Contract Nos. N00014-84-K-0352 and DAAG29-83-K-0122). We also thank D. Jenkins for his assistance with the figures.

References

[1] Materials like GaSe that are resistant to the formation of surface defects do not exhibit Fermi-level pinning, but obey the Schottky model.
[2] J. Bardeen, Phys. Rev. 71 (1947) 717.
[3] W.E. Spicer, P.W. Chye, P.R. Skeath, C.Y. Su and I. Lindau, J. Vacuum Sci. Technol. 16 (1979) 1422;
W.E. Spicer, I. Lindau, P.R. Skeath and C.Y. Su, J. Vacuum Sci. Technol. 17 (1980) 1019.
[4] H.P. Hjalmarson, P. Vogl, D.J. Wolford and J.D. Dow, Phys. Rev. Letters 44 (1980) 810;
H.P. Hjalmarson, PhD Thesis, University of Illinois (1979) (unpublished);
H.P. Hjalmarson, P. Vogl, D.J. Wolford and J.D. Dow (unpublished).
Many of the ideas of this work are based on earlier work by W.Y. Hsu, J.D. Dow, D.J. Wolford and B.G. Streetman, Phys. Rev. B16 (1977) 1597.
[5] ϵ_c and ϵ_v here are s atomic orbital energies, as given in ref. [6].
[6] P. Vogl, H.P. Hjalmarson and J.D. Dow, J. Phys. Chem. Solids 44 (1983) 365.
[7] W.A. Harrison, Electronic Structure and the Properties of Solids (Freeman, San Francisco, 1980).
[8] O.F. Sankey and J.D. Dow, Phys. Rev. B26 (1982) 3243, and references therein.
[9] M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30 (1969) 2409.
[10] H.H. Wieder, Surface Sci. 132 (1983) 390, and references therein.
[11] W. Mönch, Surface Sci. 132 (1983) 92, and references therein.
[12] R.H. Williams, Surface Sci. 132 (1983) 122, and references therein.
[13] O.F. Sankey and J.D. Dow, Phys. Rev. B27 (1983) 7641.
[14] O.F. Sankey, R.E. Allen and J.D. Dow, to be published.
[15] J.D. Dow, R.E. Allen and O.F. Sankey, in: Chemistry and Physics of Solid Surfaces, Vol. 5, Eds. R. Vanselow and R. Howe (Springer, Berlin, 1984).

[16] J.A. Van Vechten, J. Electrochem. Soc. 122 (1975) 419, 423. The defect formation process is almost certainly an event that occurs far from equilibrium.
[17] J.D. Dow and R.E. Allen, J. Vacuum Sci. Technol. 20 (1982) 659.
[18] L.J. Brillson, Surface Sci. Rept. 2 (1982) 132, and references therein.
[19] Y. Shapira and L.J. Brillson, J. Vacuum Sci. Technol. B1 (1983) 618.
[20] O.F. Sankey, R.E. Allen and J.D. Dow, Solid State Commun. 49 (1984) 1.
[21] J.D. Dow, Lectures (Varenna), to be published.
[22] J.D. Dow and R.E. Allen, to be published.

Applications of Surface Science 22/23 (1985) 948–968
North-Holland, Amsterdam

ADVANCES IN CHARACTERIZING AND CONTROLLING METAL–SEMICONDUCTOR INTERFACES *

L.J. BRILLSON

Xerox Webster Research Center, 800 Phillips Road, W-114, Webster, New York 14580, USA

Received 27 August 1984; accepted for publication 7 November 1984

We have used a variety of novel approaches in characterizing metal–semiconductor interfaces – soft X-ray photoemission spectroscopy with interlayers or markers, surface photovoltage spectroscopy, and cathodoluminescence spectroscopy, coupled with pulsed laser annealing – to reveal systematics between interface chemical and electronic structure. The chemical basis for these interfacial properties suggests new avenues for controlling electronic structure on a microscopic scale.

1. Introduction

With the application of surface science techniques to the study of metal–semiconductor interfaces, considerable progress has been achieved in understanding the interactions which take place at the microscopic junction and their influence on macroscopic electronic properties [1–6]. In particular, it is now generally accepted that the extrinsic electronic states of a metal–semiconductor interface – e.g., those due to some interaction between metal and semiconductor – rather than any intrinsic states present at the semiconductor surface – dominate the Schottky barrier formation. Considerable evidence for these conclusions has been derived from contact potential [7,8], surface photovoltage, low energy electron loss [9,10], UV [12,13], and soft X-ray photoemission spectroscopies [14–21]. With these techniques, research groups around the world have found strong charge transfer and atomic redistribution occurring with the deposition of only a few monolayers or less of deposited metal on clean, ordered semiconductor surfaces. Thus related phenomena such as chemical reactions, diffusion, formation of defects, dipoles, and alloy layers at the metal–semiconductor interface are observed which can account for Schottky barrier formation on an atomic scale. Within the last few years, this body of work has been extended to reveal further

* This work reported here was carried out in collaboration with C.F. Brucker, A. Katnani, M. Kelly, G. Margaritondo, H. Richter, Y. Shapira, M. Slade, and N.G. Stoffel.

0378-5963/85/$03.30 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

Densities of phonon states for $(GaSb)_{1-x}(Ge_2)_x$

Akiko Kobayashi*
Department of Physics and Materials Research Laboratory, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801
and Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

Kathie E. Newman and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 11 March 1985)

A theory of the phonon spectral densities of states is reported for the long-lived, metastable, crystalline alloys $(GaSb)_{1-x}(Ge_2)_x$, assuming both a zinc-blende–diamond phase-transition model of the metastable structure and an on-site substitutional model without a phase transition. Comparison of both models with the data strongly favors the phase-transition model, but not so strongly as to be absolutely compelling evidence for the model. The theory is evaluated with use of the recursion method and a rigid-ion approximation with first- and second-nearest-neighbor force constants. The evolution of the spectra with increasing x is predicted, and is considerably more complicated than that given by either a virtual crystal or a persistence approximation. Principal spectral features are associated with vibrations of various bonds. All of the major anomalies in the Raman data are explained: (i) The alloy dependence of the Ge-like LO Raman linewidth is related to the entropy per site. (ii) The discontinuity as a function of x in the Raman peak position of the Ge-like LO mode is due to the changing importance of vibrations associated with Ge–Ge and Ge–Ga bonds. (iii) The anomalous asymmetries of the GaSb-like and Ge-like LO Raman peaks are attributed to spectral features associated with Sb–Sb bonds and Ga–Ge bonds, respectively. (iv) The LO-TO splitting at $k=0$ is proportional to the order parameter of the phase transition and decreases from the GaSb value to zero as x increases to $x_c \simeq 0.3$. The model leads strong but indirect support to the phase-transition model over the on-site model, because, as a function of x, the maximum Raman linewidth coincides with the maximum entropy in the phase-transition model, but not in the on-site model.

I. INTRODUCTION

Recently Greene *et al.* fabricated metastable, substitutional, crystalline alloys of group-IV elements with III-V compounds such as $(GaAs)_{1-x}(Ge_2)_x$ (Ref. 1) and $(GaSb)_{1-x}(Ge_2)_x$ (Ref. 2). Unlike conventional quasibinary III-V alloys such as $GaAs_{1-x}P_x$, the characterization of the disorder in this new type of alloy is complicated by the fact that three electronically different elements are distributed on the lattice sites of two face-centered-cubic sublattices. The first theoretical approach to this problem, proposed by Newman *et al.*,[3,4] is based on a mean-field theory in which a zinc-blende–diamond (order-disorder) phase transition occurs at an alloy composition $x_c=0.3$. This approach successfully explains the observed V-shaped bowing of the direct band gap in $(GaAs)_{1-x}(Ge_2)_x$ as a function of composition x.[3] Furthermore, unambiguous confirmation of the predicted phase transition has been reported in the recent x-ray diffraction measurements of $(GaSb)_{1-x}(Ge_2)_x$ by Barnett *et al.*[5] Effects of the disorder in the alloy are also seen in the vibrational properties of $(GaAs)_{1-x}(Ge_2)_x$ and $(GaSb)_{1-x}(Ge_2)_x$, as shown in some recent measurements of their Raman spectra.[6–9]

In this paper, we consider the $(GaSb)_{1-x}(Ge_2)_x$ Raman data of Krabach *et al.*[6] and Beserman *et al.*,[7] which show anomalously broad linewidths for compositions in the range $x \simeq 0.2$–0.7. In order to understand how the disorder affects the vibrational properties of these metastable alloys, we calculate the vibrational densities of states of $(GaSb)_{1-x}(Ge_2)_x$ using the recursion method.[10–12] We employ two theoretical models of the disorder, one based on the predicted phase transition, and one (without an order-disorder transition) that assumes the Ge merely dilutes the GaSb without leading to the formation of any antisite defects or Sb–Sb bonds. The resulting densities of phonon states of $(GaSb)_{1-x}(Ge_2)_x$ are compared with the Raman data.[6,7]

The recursion method has been applied previously to theoretical studies of phonons in perfect crystals,[13] but, to our knowledge, the present work represents one of the first applications to semiconductive crystalline alloys,[14] and the first application to $(A^{III}B^{V})_{1-x}(C_2^{IV})_x$ metastable alloys. The recursion method is basically a continued-fraction scheme for calculating the Green's function and the local density of states at a central site of a cluster of atoms in real space. A major advantage of the method over other cluster schemes is that the continued-fraction truncation does not introduce spurious gaps (as a supercell scheme does) or surface states (as a finite-cluster method does). A clear and concise discussion of the method is given in the work of Nex[11] and Haydock;[10] the computer subroutines used to evaluate the densities of phonon states are the Cambridge Recursion Library;[11] main routines are

 ©1985 The American Physical Society

listed in Ref. 15. The method properly accounts for the local environment surrounding the central atom in a cluster, and thus it is ideally suited for obtaining local densities of states in any disordered system, especially one that is well described by a localized basis, such as a tight-binding model. The advantage of the recursion method over other commonly employed alloy theories (such as the coherent-potential approximation) is that the effects of clusters of minority atoms are explicitly included, and so the theory is accurate beyond orders x or $(1-x)$—while the coherent-potential approximation is not.

The main shortcoming of the method is that it is not well suited for handling long-ranged forces, such as the dipole-dipole interaction responsible for the Lyddane-Sachs-Teller splitting of longitudinal and transverse optic phonon modes at the center of the Brillouin zone.[16] Nevertheless a previous recursion calculation for the quasibinary semiconducting alloy $Al_xGa_{1-x}As$ (Ref. 14) showed that the method, despite its neglect of long-ranged forces, is useful for correlating peaks in the density of phonon states with major features of Raman and infrared spectra, and for associating those peaks with specific local atomic configurations in the alloy.

In this paper we compute the phonon spectral densities of states assuming two models of substitutional alloy disorder: (i) an on-site model, in which cations always remain on nominal cation sites, and (ii) a phase-transition model,[3,4] in which all atoms can occupy either nominal cation or nominal anion sites.

II. THEORY

There are three basic steps to our theory: (i) The development of an adequate theory for describing the dispersion relations and phonon densities of states of pure GaSb and Ge (Sec. II A); (ii) the modeling of the alloy statistics—that is, the determination of the probabilities that a site is occupied by each of the atoms Ga, Sb, and Ge (Sec. II B); and (iii) the proposing of prescriptions for determining the force constants in a specific cluster of the alloy in terms of the force constants of GaSb and Ge (Sec. II C). Details of the calculations are in the Appendices.

A. Phonons in GaSb and Ge

1. Force constants and dispersion relations

We assume the harmonic approximation and use a simple rigid-ion model to parametrize the phonon dispersion curves of GaSb and Ge, ignoring all but the first- and second-nearest-neighbor force constants; we also set the long-ranged Coulomb forces to zero. There are a number of disadvantages associated with this simple model. The most important is the neglect of all long-ranged Coulomb forces, which are responsible for (i) the nonzero splitting of the longitudinal optic (LO) and transverse optic (TO) modes at the center of the Brillouin zone for zinc-blende materials (due to the dipole-dipole interactions), (ii) additional contributions to the splitting of the LO and TO modes away from the Brillouin-zone center (due in part to the dipoles induced in vibrating polarizable atoms), and (iii) the flattening of the transverse acoustic (TA) branch near the Brillouin-zone boundary. These long-ranged effects have been incorporated in previous models, such as the shell model[17,18] or the bond-charge model,[19] but are neglected here because our primary interest is the study of the effects of local alloy disorder on the phonon spectra.

Justification for the neglect of Coulomb forces is afforded by the fact that (i) effective charges in these materials are typically very small (~ 0.1 eV) (Ref. 20) and (ii) Herscovici and Fibich[13] have shown that a model which includes second-nearest-neighbor force constants incorporates enough of the long-ranged forces to adequately simulate much of the essential qualitative physics resulting from those forces. With the rigid-ion model and without the long-ranged Coulomb forces, the application of the recursion method is tractable and even straightforward.

In the harmonic approximation (using Dirac notation), the time-independent eigenvalue equation of motion for the displacement of the lattice, $|\delta R\rangle$, is

$$\underline{\Phi}\,|\delta R\rangle = \underline{M}\Omega^2\,|\delta R\rangle\ , \tag{2.1}$$

where the displacement in the nth direction ($i=x$, y, or z) of the bth atom (b=anion or cation) in the nth unit cell (at R_n) is $\langle n,b,i\,|\delta R\rangle$, and M and Φ are the mass and force-constant matrices, respectively. Here we have the mass matrix (in the $|n,b,i\rangle$ basis):

$$\underline{M} = \sum_{n,b} |n,b,i\rangle M_b \langle n,b,i'|\ . \tag{2.2}$$

The force-constant matrix, $\langle n,b,i|\Phi|n',b',i'\rangle$, vanishes unless (n,b) and (n',b') refer to the same atom, nearest neighbors, or second-nearest neighbors. Taking the position of the bth atom in the nth unit cell to be R_n if b refers to an anion site of the underlying zinc-blende structure ($v_a=0$), and R_n+v_c if b refers to a cation site, we note that for the zinc-blende lattice, we have $v_c=(a_L/4)(1,1,1)$ with a_L being the lattice constant. Then we have, for the force constants between the anion-site atom and its nearest cation-site neighbor in the same unit cell:

$$\langle n,b=\text{anion},i\,|\,\Phi\,|\,n,b'=\text{cation},i'\rangle = \begin{array}{c|ccc} & \langle x| & \langle y| & \langle z| \\ \hline \langle x| & \alpha & \beta & \beta \\ \langle y| & \beta & \alpha & \beta \\ \langle z| & \beta & \beta & \alpha \end{array}_{i,i'}\ , \tag{2.3}$$

where the rows and columns are labeled by i and i'. The values of α and β depend on the two atoms at either end of the chemical bond between (n,b) and (n',b'); for example, in pure GaSb we have just two parameters: α(Ga—Sb) and β(Ga—Sb). The other nearest neighbors to an anion atom at R_n are in different unit cells, and the corresponding force-

constant matrices are obtained by symmetry operations, as given in Appendix A. The second-nearest-neighbor force-constant matrix between the anion atom at R_n and the one at $R_{n'}=R_n+(a_L/2)(0,1,1)$ is

$$\langle n,b=\text{anion},i\,|\,\underline{\Phi}\,|\,n',b'=\text{anion},i'\rangle=\begin{array}{c}\\ \langle x|\\ \langle y|\\ \langle z|\end{array}\begin{array}{c}\begin{array}{ccc}|x\rangle & |y\rangle & |z\rangle\end{array}\\ \begin{pmatrix}\lambda_a & 0 & 0\\ 0 & \mu_a & \nu_a\\ 0 & \nu_a & \mu_a\end{pmatrix}_{i,i'}\end{array} \tag{2.4}$$

Appendix A shows how the force-constant matrices for other second neighbors are related to Eq. (2.4); a similar expression holds for cations, with λ_c, μ_c, and ν_c.

The remaining nonzero elements of the force-constant matrix are the diagonal elements, which follow from the requirement of invariance of the equations under infinitesimal translation of the crystal:

$$\langle n,b,i\,|\,\underline{\Phi}\,|\,n,b,i'\rangle=-\sum_{n',b'}{}'\langle n,b,i\,|\,\underline{\Phi}\,|\,n',b',i'\rangle\ , \tag{2.5}$$

where the prime on the summation indicates that $(n',b')=(n,b)$ is excluded.

Taking advantage of the translation-group symmetry of the lattice, we change basis to Bloch states:

$$|\,\mathbf{k},b,i\rangle=N^{-1/2}\sum_n|\,n,b,i\rangle\exp[-i\mathbf{k}\cdot(\mathbf{R}_n+\mathbf{v}_b)]\ , \tag{2.6}$$

where N is the number of unit cells and k lies in the first Brillouin zone. In this Bloch-type basis set, the equation of motion, Eq. (2.1), can be written as

$$(\Omega^2\underline{I}-\underline{D})\,|\,\delta\mathbf{R}\rangle=0 \tag{2.7}$$

or

$$\sum_{i',b'}(\Omega^2\delta_{i,i'}\delta_{b,b'}-\langle\mathbf{k},b,i\,|\,\underline{D}\,|\,\mathbf{k},b',i'\rangle)\langle\mathbf{k},b',i'\,|\,\delta\mathbf{R}\rangle=0\ , \tag{2.8}$$

where we have

$$\langle\mathbf{k},b,i\,|\,\underline{D}\,|\,\mathbf{k},b',i'\rangle=(M_bM_{b'})^{-1/2}\sum_{n'}\langle n,b,i\,|\,\underline{\Phi}\,|\,n',b',i'\rangle\times\exp[i\mathbf{k}\cdot(\mathbf{R}_{n'}+\mathbf{v}_{b'}-\mathbf{R}_n-\mathbf{v}_b)]\ . \tag{2.9}$$

This is the dynamical matrix element in the Bloch-type basis set, $|\,\mathbf{k},b,i\rangle$. The phonon dispersion relation is determined from the secular equation (2.7) or (2.8), for a finite set of k values in a Brillouin zone. The crystal eigenstates are linear combinations of the above Bloch-type basis states:

$$|\,\mathbf{k},s\rangle=\sum_{b,i}\langle\mathbf{k},b,i\,|\,\mathbf{k},s\rangle\,|\,\mathbf{k},b,i\rangle\ . \tag{2.10}$$

Using the first- and second-nearest-neighbor force-constant matrices given in Eqs. (2.3) and (2.4) [and those in Eqs. (A4) and (A5) in Appendix A], as well as the "on-site force-constant matrices" determined by the relation in Eq. (2.5), the dynamical matrix in the present model is constructed as summarized in Appendix B.

Numerical values for the force-constant parameters for pure crystals with zinc-blende structure can be obtained by fitting the model to experimental data at several symmetry points in a Brillouin zone, and to elastic-constant data. Appendix C contains a detailed description of the fitting procedure. A tabulation of the resultant values of the force-constant parameters as well as the atomic masses may be found in Table I for GaSb and Ge, and in Ref. 15 for other semiconductors.

2. *Densities of states*

To obtain the densities of phonon states for GaSb and Ge, we employ 1000-atom clusters and execute the recursion method to 51 levels of continued fractions. These parameters are found to be necessary to obtain good representations of the densities of phonon states and to produce well-converged spectra. The phonon dispersion relations are compared with available data, and the densities of states calculated by the recursion method are compared with those evaluated using the Lehmann-Taut method[21] in Figs. 1 and 2 for GaSb and Ge,[22–25] respectively. As one can notice, the sharp Van Hove singularities[26] present in the Lehmann-Taut densities of states are absent from the densities of states obtained by the recursion method. Thus the recursion method is limited to representing densities of states of alloys, where the perfect-crystal Van Hove singularities are expected to be blurred by the disorder.

TABLE I. Force-constant parameters (in units of 10^3 dyn/cm) and masses (in units of 10^{-24} g).

	GaSb	Ge
α	−39.525[a]	−49.470
β	−34.000[a]	−37.159
λ_a	4.500[a]	5.821
λ_c	4.500[a]	5.821
μ_a	−3.697[a]	−3.092
μ_c	−4.467[a]	−3.092
ν_a	−3.697[a]	−3.092
ν_c	−4.467[a]	−3.092
M_a	202.1695	120.5379
M_c	115.7722	120.5379

[a]Parameters for GaAs obtained from Ref. 22. We used these parameters for GaSb as well.

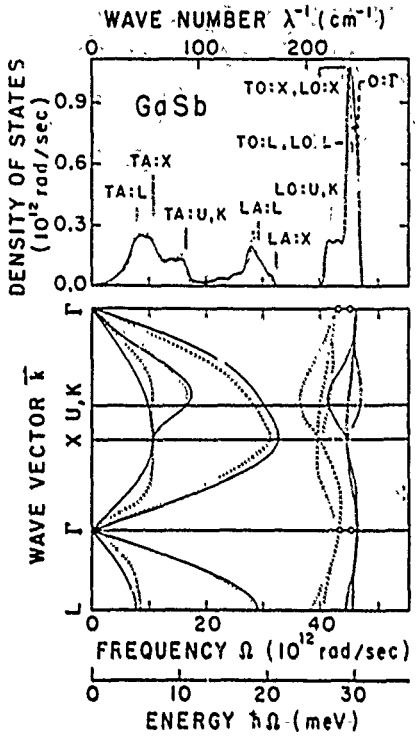


FIG. 1. Densities of phonon states times 10^{12} rad/sec (top) and phonon dispersion curves (bottom) for GaSb. The density of states obtained by the recursion method (solid line) is compared with the one obtained by the Lehmann-Taut method (dashed line). The phonon dispersion curves are obtained using the force-constant parameters for GaSb from Table I (solid lines), and are compared with the infrared reflection data from Ref. 23 at Γ point (circles), and with the neutron scattering data from Ref. 24 (dotted lines). The energies of various phonons are denoted on the density-of-states figure, as, for example, LO:U,K for the LO phonon at the U and K points of the Brillouin zone. The symmetry points of the Brillouin zone are $\Gamma=(0,0,0)$, $L=(2\pi/a_L)(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, $X=(2\pi/a_L)(1,0,0)$, $U=(2\pi/a_L)(1,\frac{1}{4},\frac{1}{4})$, and $K=(2\pi/a_L)(\frac{3}{4},\frac{3}{4},0)$.

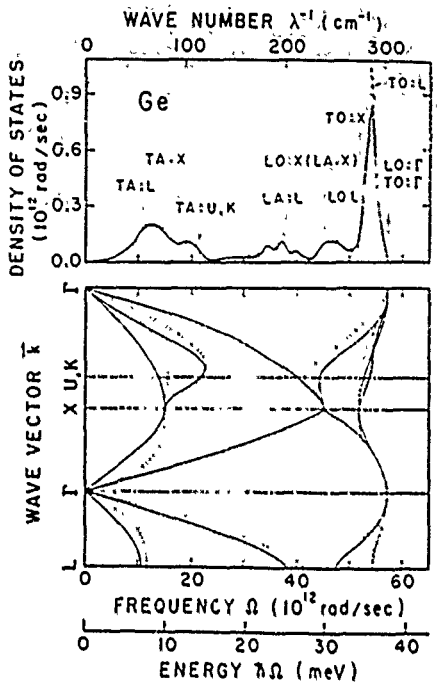


FIG. 2. Densities of phonon states times 10^{12} rad/sec (top) and phonon dispersion curves (bottom) for Ge. The density of states obtained by the recursion method (solid line) is compared with the one obtained by the Lehmann-Taut method (dashed line). The phonon dispersion curves are obtained in the present model (solid lines), and are compared with the neutron scattering data from Ref. 25 (dotted lines).

B. Alloy statistics

In order to treat the phonons in these alloys in a tractable manner, we first need to model the distribution of the three types of atoms on the two face-centered-cubic sublattices. The simplest approach is to view the substitutional alloy as a III-V compound with an underlying zinc-blende structure, but with the group-IV elements dispersed randomly. If at first we assume an "on-site model," the probability of occupancy of each atom (Ga, Sb, and Ge) on each nominal sublattice site (a=anion site, or c=cation site) is given by

$$\langle P_{Ga}\rangle_c=\langle P_{Sb}\rangle_a=1-x\ , \tag{2.11a}$$

$$\langle P_{Sb}\rangle_c=\langle P_{Ga}\rangle_a=0\ , \tag{2.11b}$$

and

$$\langle P_{Ge}\rangle_c=\langle P_{Ge}\rangle_a=x\ . \tag{2.11c}$$

Note that these probabilities allow only a limited type of alloy clusters: there is zero probability that a Ga atom will sit next to a Ga atom, for example—there are no antisite defects such as a Ga atom on an Sb site, and hence no Ga—Ga or Sb—Sb bonds.

Missing from the probabilities in Eqs. (2.11a)–(2.11c) is the idea of Newman *et al.*,[3,4] which postulates that $(GaSb)_{1-x}(Ge_2)_x$ [like $(GaAs)_{1-x}(Ge_2)_x$] undergoes a phase transition as a function of alloy composition x from a zinc blende to a diamond structure. They define an order parameter M which is nonzero in the zinc-blende phase and is zero in the diamond phase (since there is no distinction between "anion" and "cation" in the diamond structure):

$$M=(\langle P_{Ga}\rangle_c-\langle P_{Ga}\rangle_a+\langle P_{Sb}\rangle_a-\langle P_{Sb}\rangle_c)/2\ . \tag{2.12}$$

If we assume, for simplicity, that Ge occupies both types of sites with equal probability ($\langle P_{Ge}\rangle_c=\langle P_{Ge}\rangle_a$), and also that nominal zinc-blende-lattice sites are occupied by one of the three types of atoms, then we have simply

$$\begin{aligned}M&=\langle P_{Ga}\rangle_c-\langle P_{Ga}\rangle_a\\&=\langle P_{Sb}\rangle_a-\langle P_{Sb}\rangle_c\ .\end{aligned} \tag{2.13}$$

This order parameter can take on all the values from $x-1$ to $1-x$: If all the Ga atoms occupy nominal cation sites, then we have $M=1-x$; if the Ga atoms are distributed evenly over nominal cation sites and nominal anion sites, then we have $M=0$; if all the Ga atoms are on the nominal anion sites, then we have $M=x-1$; and if we have $|M|\neq 1-x$, then some anions and cations are

found on nominally "wrong" sublattice sites. The average occupancies of Ga, Sb, and Ge atoms on nominal cation sites c and on nominal anion sites a in the phase-transition model are summarized as

$$\langle P_{\mathrm{Ga}}\rangle_c = (1-x+M)/2 , \tag{2.14a}$$

$$\langle P_{\mathrm{Ga}}\rangle_a = (1-x-M)/2 , \tag{2.14b}$$

$$\langle P_{\mathrm{Sb}}\rangle_c = (1-x-M)/2 , \tag{2.14c}$$

$$\langle P_{\mathrm{Sb}}\rangle_a = (1-x+M)/2 , \tag{2.14d}$$

and

$$\langle P_{\mathrm{Ge}}\rangle_c = \langle P_{\mathrm{Ge}}\rangle_a = x . \tag{2.14e}$$

The order parameter $M(x)$ is determined within a mean-field approximation as the solution of[3,4]

$$M/(1-x) = \tanh[M/(1-x_c)] , \tag{2.15}$$

where the phase-transition composition x_c is extracted from data. We use $x_c=0.3$ as is suggested by both the optical absorption[3] and the x-ray diffraction data.[5] We note that the probabilities of atomic distribution in the on-site model, Eqs. (2.11a)–(2.11c), can be obtained simply from the phase-transition model probabilities, Eqs. (2.14a)–(2.14e), by replacing M with $1-x$.

It should be emphasized that the present calculations are based on the phase-transition theory of the alloy as evaluated in a mean-field approximation. To be sure, the mean-field approximation is not exact and an improved theory may be needed as the data for $(GaSb)_{1-x}(Ge_2)_x$ are refined.[27] Indeed Holloway and Davis[28] have recently suggested an alternative model of the similar alloys $(GaAs)_{1-x}(Ge_2)_x$ in which they forbid the formation of As–As and Ga–Ga bonds—and hence forbid the order-disorder phase transition. Their model assumes that the zinc-blende–diamond transition is percolative in nature, i.e., the transition is assumed to be controlled by geometry alone, with a transition composition x_c characteristic of a site-diluted diamond lattice, $x_c \simeq 0.5$.[29] Thus their model does not either include differences in clustering that can arise from different growth techniques or allow for temperature-dependent growth. (For example, the restriction of eliminating Ga–Ga and As–As bonds forces the probability of finding a Ga–Ge bond to be identical to the probability of finding a Ge–As bond.) Given the dependence of $(GaAs)_{1-x}(Ge_2)_x$ electronic structure on growth conditions seen experimentally,[1,30] and the existence of x-ray diffraction data supporting the idea of a zinc-blende–diamond phase transition for $(GaSb)_{1-x}(Ge_2)_x$,[5] we elect to concentrate here on the model which contains those features. We do, however, study the difference in phonon densities of states obtained between distributions given by Eqs. (2.11a)–(2.11c) and Eqs. (2.14a)–(2.14e).

C. Force constants

Based on the atomic distributions defined in Eqs. (2.11) and (2.14) above, we generate a 1000-atom cluster (using a random number generator) in which the distribution of masses is given either by Eqs. (2.11a)–(2.11c) for the case of $M=1-x$, or by Eqs. (2.14a)–(2.14e) for the case of M determined by the mean-field theory with the occurrence of the phase transition at $x_c=0.3$. Using the force-constant parameters for parent compounds in Table I, the first-nearest-neighbor force constants α for the six possible pairs of first-neighboring atoms in $(GaSb)_{1-x}(Ge_2)_x$ are given by

$$\alpha((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ga}-\mathrm{Ga}) = \alpha(GaSb) , \tag{2.16a}$$

$$\alpha((GaSb)_{1-x}(Ge_2)_x;\mathrm{Sb}-\mathrm{Sb}) = \alpha(GaSb) , \tag{2.16b}$$

$$\alpha((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ge}-\mathrm{Ge}) = \alpha(Ge) , \tag{2.16c}$$

$$\alpha((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ga}-\mathrm{Sb}) = \alpha(GaSb) , \tag{2.16d}$$

$$\alpha((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ga}-\mathrm{Ge}) = (1-x)\alpha(GaSb)+x\alpha(Ge) , \tag{2.16e}$$

$$\alpha((GaSb)_{1-x}(Ge_2)_x;\mathrm{Sb}-\mathrm{Ge}) = (1-x)\alpha(GaSb)+x\alpha(Ge) . \tag{2.16f}$$

We use the same relations for β for these pairs of atoms. The second-nearest-neighbor force constants λ for six possible pairs of second-neighboring atoms are given by

$$\lambda((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ga}-\mathrm{Ga}) = \lambda_c(GaSb) , \tag{2.17a}$$

$$\lambda((GaSb)_{1-x}(Ge_2)_x;\mathrm{Sb}-\mathrm{Sb}) = \lambda_a(GaSb) , \tag{2.17b}$$

$$\lambda((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ge}-\mathrm{Ge}) = \lambda(Ge) , \tag{2.17c}$$

$$\lambda((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ga}-\mathrm{Sb}) = [\lambda_c(GaSb)+\lambda_a(GaSb)]/2 , \tag{2.17d}$$

$$\lambda((GaSb)_{1-x}(Ge_2)_x;\mathrm{Ga}-\mathrm{Ge}) = (1-x)\lambda_c(GaSb)+x\lambda(Ge) , \tag{2.17e}$$

$$\lambda((GaSb)_{1-x}(Ge_2)_x;\mathrm{Sb}-\mathrm{Ge}) = (1-x)\lambda_a(GaSb)+x\lambda(Ge) . \tag{2.17f}$$

We use the same relations for the other second-nearest-neighbor force constants μ and ν, for these pairs of atoms.

D. Outline of the calculation

The local density of states $d(\mathbf{R};\Omega)$ in a disordered system depends sensitively on the local atomic environment. To obtain the total densities of states we sum over an ensemble of local densities of states as follows. (i) A specific five-atom minicluster is generated, e.g., a central Ga atom surrounded by two As, one Ga, and one Ge, for the center of the 1000-atom cluster; (ii) the probability $p(x)$ of this cluster occurring in an alloy with composition x is determined (see Appendix D); (iii) the remaining 995 atoms are added with the probability for each atom as defined in Eqs. (2.11a)–(2.11c) or Eqs. (2.14a)–(2.14e); and (iv) the local density of states is computed by the recursion method at the central site of the minicluster embedded in its alloy environment of the 995-atom cluster. The process is repeated for all possible miniclusters, each embedded in the 995-atom cluster. The total density of states per unit cell is then

$$D(\Omega;(GaSb)_{1-x}(Ge_2)_x)=\sum_\kappa p_\kappa(x)d_\kappa(R;\Omega)\,, \quad (2.18)$$

where the sum is over all possible miniclusters, and d_κ and p_κ are the local density of states and the probability of occurrence of the κth minicluster, respectively.

In addition to the densities of states by the recursion method for the above two types of disorder, we also obtained, for comparison, (i) the densities of states for the virtual-crystal-approximation alloy as well as (ii) the densities of phonon states in the persistent approximation,[31] which is a linear superposition of the densities of states of GaSb and Ge:

$$D(\Omega;(GaSb)_{1-x}(Ge_2)_x)=(1-x)D(\Omega;GaSb)+xD(\Omega;Ge)\,. \quad (2.19)$$

In the following section, we discuss our results for the densities of phonon states and interpret available Raman data for $(GaSb)_{1-x}(Ge_2)_x$.[6,7]

III. RESULTS AND DISCUSSION

A. Spectra

In Fig. 3 we display the calculated densities of states for various alloy compositions in the case of the on-site model, i.e., the model with $M=1-x$, as well as the density of states obtained by the virtual-crystal approximation for $x=0.5$. In this figure, we can see that the virtual-crystal approximation yields a single "amalgamated"[31] optic band. Thus, this approximation does not reproduce the two-mode behavior for the optic-phonon spectrum expected of semiconducting mixed crystals: two sets of long-wavelength optic phonons that are each energetically close to those of the parent compounds, GaSb and Ge. We conclude that the virtual-crystal approximation should not be used to describe phonons in these materials.

Figure 4 shows the densities of states calculated for the

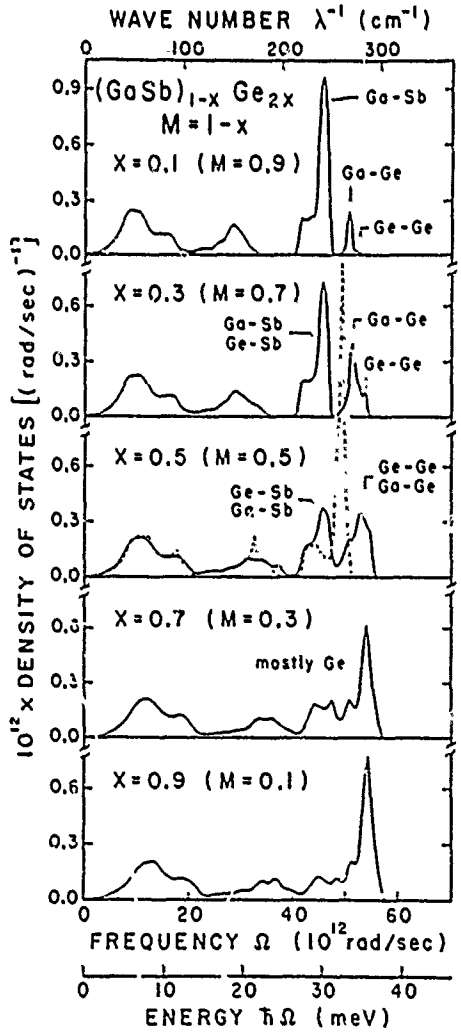


FIG. 3. Densities of phonon states for $(GaSb)_{1-x}(Ge_2)_x$ alloys obtained by the recursion method in the case of $M=1-x$ (solid lines), along with density of states obtained using the virtual crystal approximation for $x=0.5$ (dashed line). The assignments given to prominent peaks represent bonds that are responsible for the vibrations giving rise to the peaks (see text).

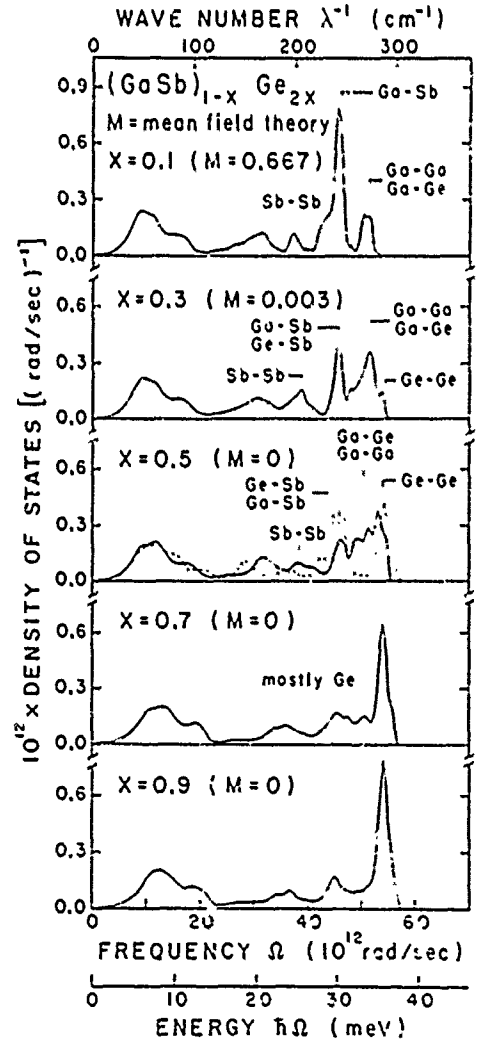


FIG. 4. Densities of phonon states for $(GaSb)_{1-x}(Ge_2)_x$ alloys obtained by the recursion method in the case of M determined by mean-field theory (solid lines), along with density of states obtained using the persistent approximation, Eq. (2.19), for $x=0.5$ (dashed line). The assignments given to prominent peaks represent bonds that are responsible for the vibrations giving rise to the peaks (see text).

case of M determined by the phase-transition model in the mean-field approximation, as well as the persistent spectra [Eq. (2.19)] for $x=0.5$. The persistent approximation is capable of describing at least two optic bands originating from the parent compounds, although it does not include alloy modes, which are characteristic vibrations of various atomic arrangements in alloys that do not exist in the parent compounds (e.g., characteristic vibrations of an Sb atom surrounded by three Ga atoms and one Ge atom). This approximation seems to work well for quasibinary alloys, such as $Al_xGa_{1-x}As$ (Ref. 14), where the substitutional disorder is limited to the atomic arrangements on one of the two face-centered-cubic sublattices; in this case, deviations of the recursion spectra from the persistent spectra are relatively minor and can be easily attributed to the alloy modes. However, in the case of $(GaSb)_{1-x}(Ge_2)_x$ alloys (see Fig. 4), there seems to be a major difference between the persistent spectrum (dashed line) and the recursion spectrum (solid line). This obviously indicates a high degree of disorder on *both* sublattices, which gives rise to many different types of alloy modes, and therefore to a variety of peaks in the density-of-states spectra.

To see how the different types of disorder affect the vibrational densities of states, we compare the recursion spectra obtained from the atomic distributions defined in Eqs. (2.11a)–(2.11c) and Eqs. (2.14a)–(2.14e). The difference between these two models is the type of disorder: in one case [the on-site model of Eqs. (2.11a)–(2.11c)], we set the order parameter M equal to $1-x$ and have no "antisite defects" such as Sb on a nominal Ga site, while in the other case [the phase-transition model of Eqs. (2.14a)–(2.14e)], the order parameter is determined by mean-field theory, and antisite defects occur in relatively high concentrations. The local densities of states obtained by the recursion method for each central atom in each configuration of the five-atom miniclusters are useful in identifying the origin of peaks in the total density of vibrational modes. Each of the statistically independent configurations of the miniclusters contributes to characteristic peaks in the density of states. We find, however, that it is most convenient to associate different contributions to the densities of states with various "bonds" rather that with entire miniclusters. In this scheme, we have four different types of bonds for the on-site case of $M=1-x$: Ga–Sb, Ga–Ge, Ge–Sb, and Ge–Ge; and two additional ones for the case of M determined by mean-field theory: Ga–Ga and Sb–Sb. We concentrate on the optic region since the optic modes are sensitive to local atomic order and thus exhibit interesting disorder effects.

In the densities of states calculated with $M=1-x$ (Fig. 3), we note the existence of some structure in each of the two main optic bands. The optic band at the higher frequency is "Ge-like" and increases in intensity with increasing alloy composition, x, and the optic band at the lower frequency is "GaSb-like" and decreases in intensity as x increases. Both longitudinal and transverse-acoustic bands gradually change in positions and intensities on going from GaSb to Ge. The structures in the main optic bands are assigned to vibrations of various bonds as shown in Fig. 3.

The impurity (local) mode at ~ 270 cm^{-1} for $x=0.1$ results mainly from vibrations of Ga–Ge bonds. Vibrations of Ge–Ge bonds also produce modes near this frequency; as the alloy composition x increases, the Ga–Ge and Ge–Ge vibrations gradually evolve to form the Ge-like optic band. At $x\simeq 0.7$ this Ge-like optic band is composed mostly of vibrations of Ge–Ge bonds.

The GaSb-like optic band at $x=0.1$ retains most of the features of pure GaSb, and thus the highest peak corresponds to the zone-boundary TO phonons (such as the transverse-optic modes at the X and L points of the Brillouin zone, denoted TO:X and TO:L) and the shoulder at ~ 220 cm^{-1} corresponds to the zone-boundary LO phonons (such as LO:U,K). (These can be interpolated from the k-space assignments for GaSb in Fig. 1.) As alloy composition increases, vibrations from Ge–Sb bonds take part and smooth the shape of the GaSb-like optic band. And at $x\simeq 0.7$ the modes from Ga–Sb and Ge–Sb bonds become highly degenerate with the nearby modes, and the density of states retains most of the features of pure Ge.

The densities of states with M determined by the mean-field theory (Fig. 4) show much richer structure than those with $M=1-x$. At $x=0.1$, in addition to the mostly GaSb-like features, we can see two new peaks. These are impurity modes: the peak at ~ 270 cm^{-1} is mainly composed of vibrations of Ga–Ga (and also Ga–Ge) bonds, and the peak at ~ 195 cm^{-1} is due to Sb–Sb bonds.

As alloy composition increases, vibrations of Ge–Ge bonds take part in the Ge-like optic band, and, along with the vibrations of Ga–Ga and Ga–Ge bonds, broaden this optic band in the range $0.3\leq x\leq 0.5$. In the GaSb-like optic band for $0.3\leq x\leq 0.5$, however, although the vibrations of Ge–Sb bonds as well as Ga–Sb bonds contribute to the band, the Ga–Sb plus Ge–Sb line shape in Fig. 4 appears narrower than the corresponding peak in Fig. 3, because of the spectral distribution of the modes.

The Sb–Sb peak is significant in the theory for $x=0.1$ and 0.3, but then it decreases in intensity, and is indistinguishable from nearby modes for $x>0.5$. Presumably this occurs because the number of Sb–Sb bonds rapidly decreases as alloy composition increases. As in the case of $M=1-x$ (for which there are no Sb–Sb bonds or peaks), the density of states retains most of the features of pure Ge for $x>0.7$.

We now make a comparison with data. In Figs. 5(a)–9(a), we display Raman spectra obtained by Krabach *et al.*[6] and Beserman *et al.*,[7] along with the densities of states for the case of M determined by mean-field theory [Figs. 5(b)–9(b)], and for the case of $M=1-x$ [Figs. 5(c)–9(c)]. In comparing our results with Raman data, we recognize that the experimental quantities depend on the transition matrix elements (involving electronic states) as well as on the densities of phonon states, and that the Raman lines are a subset of the density-of-states spectra. In particular, for certain experimental geometries and light polarizations, in the case of pure semiconducting compounds, many Raman matrix elements are expected to vanish because of selection rules;[32,33] for example, phonons excited by light scattering have $\mathbf{k}\simeq 0$. These selec-

tion rules are broken in alloys; certain modes are "disorder activated" and appear in the Raman spectra. Every mode that appears in a Raman spectrum should also appear in the density of phonon states, although the strength of the mode can be quite different in the Raman and density-of-states spectra.

Another caution should be taken while comparing the calculated densities of states with the data: The peak positions, as well as the peak motions as a function of alloy composition, are not accurately reproduced in the present calculations. These numerical inaccuracies are due to the neglect of long-ranged Coulomb forces and due to the use of the rigid-ion model. Even with these limitations, by mentally shifting or broadening the calculated densities of states by an amount comparable with the differences between the theory and the data for crystalline GaSb or Ge, the present calculations can be used to identify the principal spectral features and to associate them with vibrations of specific bonds.

Since the Raman spectra depend on the densities of phonon states as well as the Raman matrix elements, the effects of disorder on the Raman spectra are manifested in two ways: (i) new features in the densities of states, caused by various alloy modes which do not exist in the persistent spectra, may contribute to the Raman spectra; and (ii) due to a breakdown of the selection rules, certain modes which are not normally observed in the Raman spectra of pure compounds (such as acoustic modes and zone-boundary optic modes), as well as some of the new alloy modes, become Raman active. The alloy modes are likely responsible for the broad linewidths observed in the data of Krabach *et al.*[6] and Beserman *et al.*[7] To demonstrate this, we compare the Raman spectra with the densities of states in our two models. As mentioned earlier, the major difference in our two models is in the type of disorder allowed—the simpler on-site model (with $M=1-x$) does not include vibrational modes caused by antisite defects in the form of Ga–Ga bonds or Sb–Sb bonds. We examine the spectra in the regions (i) $x<0.15$ (ii) $0.15<x<0.75$, and (iii) $x>0.75$ below.

1. $x<0.15$

Since the Ge-atom concentration is very low in this region, the density-of-states spectra for $x=0.13$, shown in Figs. 5(b) and 5(c), retain most of the features found in GaSb. A new feature is the impurity (local) mode around 270 cm^{-1}, which is due to the vibrations of Ga–Ga and Ga–Ge bonds for the spectrum in Fig. 5(b), and is mainly due to the vibrations of Ga–Ge bonds for the spectrum in Fig. 5(c), as discussed earlier.

In the Raman spectrum of Fig. 5(a), we note that the Ge-like LO mode at ~ 260 cm^{-1} exhibits a linewidth that is relatively broad compared with that of typical impurity modes in quasibinary III-V alloys. A comparison with our calculated densities of states indicates that this feature could result from the vibrations of Ga atoms, within Ga–Ge bonds or Ga–Ga bonds, which would give rise to a broader linewidth than could be obtained from Ge–Ge bond vibrations only.

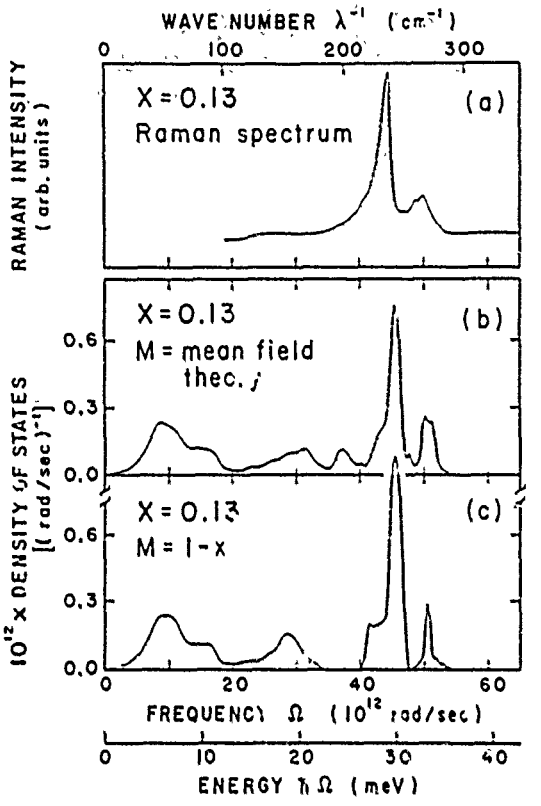


FIG. 5. (a) Raman spectrum for $(GaSb)_{0.87}(Ge_2)_{0.13}$ from Refs. 6 and 7; (b) calculated density of phonon states for the same alloy obtained by the recursion method in the case of M determined by mean-field theory; and (c) calculated density of states in the case of $M=1-x$.

Even at this low value of alloy composition, we notice a slight asymmetry in the line shape of the GaSb-like LO mode in the data. We speculate that this feature can be attributed to: (i) the disorder-activated zone-boundary LO phonons of GaSb, which produce a shoulder around 220 cm^{-1} in the densities of states in Figs. 5(b) and 5(c); or (ii) a long-wavelength GaSb-like TO mode resulting from the "leakage" due to a slight deviation from a perfect backscattering geometry. It was indicated in Ref. 7 that this weak shoulder is due to the long-wavelength GaSb-like TO mode, for $x<0.1$. Both modes, however, could coexist, discriminating between the two possibilities is difficult, because of the weakness of the feature involved.

The long tail in the frequency range between ~ 190 cm^{-1} and ~ 210 cm^{-1}, evident in the data, may have resulted from disorder-activated vibrational modes of Sb–Sb bonds [the peak around 195 cm^{-1} in Fig. 5(b)]. Note that a gap is predicted in this frequency range by the on-site model [Fig. 5(c)]. However, the density of states of the on-site model with $M=1-x$ could produce nearly the same filled-gap feature as seen in both the data and the phase-transition model [Fig. 5(b)], after a slight broadening of the GaSb-like optic band. Therefore, one cannot state unambiguously from comparing the theory and data of Fig. 5 whether or not Sb–Sb bonds are significant in the observed Raman spectrum. Nevertheless, the comparison of theory with data in Fig. 5(b) strongly

suggests that Sb–Sb bonds are likely present in significant numbers, and casts doubt on ar[illegible]lier conclusion to the contrary (which had assumed that Sb–Sb bonds would produce a sharp Raman peak).[7]

2. $0.15 < x < 0.75$

Figures 6–8 show the data for $x = 0.24$, 0.34, and 0.56, respectively, along with the densities of states obtained by the two different models: on-site and phase-transition. The two main peaks in the data exhibit broad linewidths: the total widths for alloy composition $x = 0.34$ are ~ 38 cm^{-1} for the GaSb-like LO mode, and ~ 35 cm^{-1} for the Ge-like LO mode. Asymmetry in these peaks is also evident. In addition, a disorder-activated longitudinal-acoustic mode (DALA) is observed in the data at ~ 150 cm^{-1} and this identification is supported by the theory.

The highest energy peak is associated largely with Ge vibrations, and is broad because Ga–Ge and Ga–Ga (for the phase-transition model) bonds contribute as well as Ge–Ge bonds. The lower peak is a combined Ga–Sb and Ge–Sb vibration with a low-energy shoulder associated (in the phase-transition model) with Sb–Sb bonds. Note that the widths of the two main peaks are larger in the phase-transition model than in the on-site ($M = 1-x$) model because of the additional modes due to Sb–Sb and Ga–Ga bonds. We shall see below that the systematic trends in these widths lend additional support to the phase-transition model.

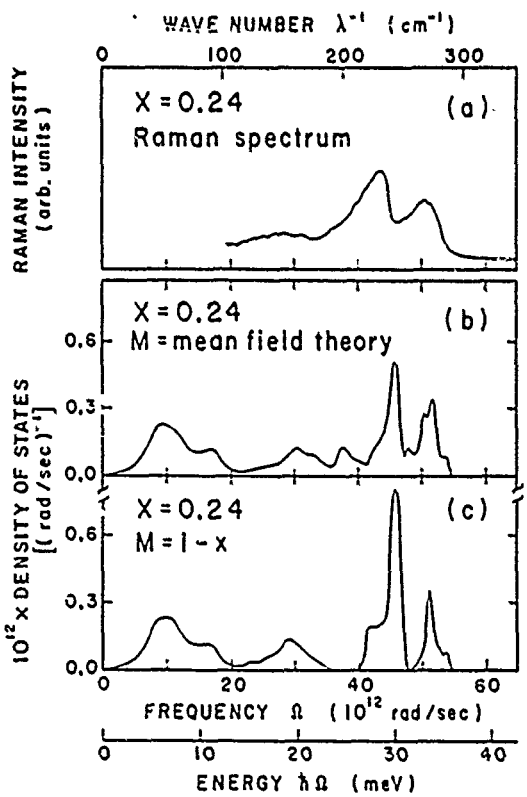


FIG 6 (a) Raman spectrum for $(GaSb)_{0.76}(Ge_2)_{0.24}$ from Refs 6 and 7, (b) calculated density of phonon states for the same alloy obtained by the recursion method in the case of M determined by mean-field theory, and (c) calculated density of states in the case of $M = 1-x$.

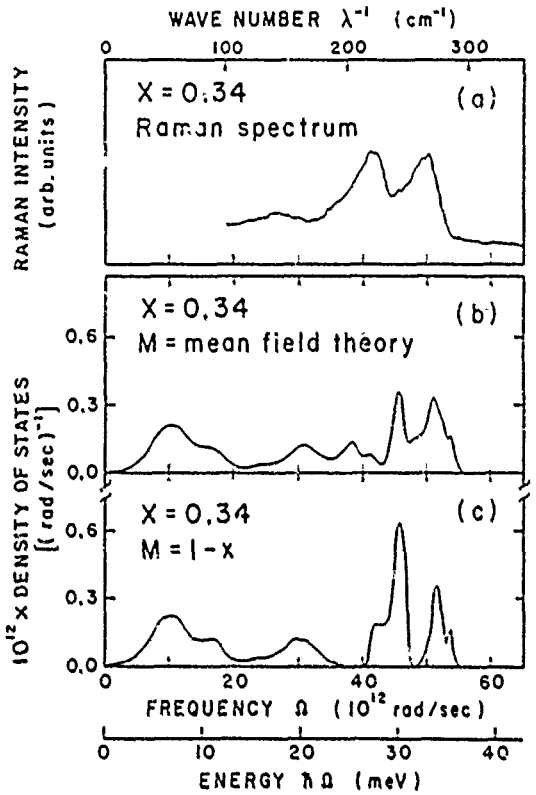


FIG. 7. (a) Raman spectrum for $(GaSb)_{0.66}(Ge_2)_{0.34}$ from Refs. 6 and 7; (b) calculated density of phonon states for the same alloy obtained by the recursion method in the case of M determined by mean-field theory; and (c) calculated density of states in the case of $M = 1-x$.

3. $x > 0.75$

Figure 9 shows the data for $x = 0.8$ along with the densities of states obtained by the two different models. We note that the densities of states retain most of the features of Ge [Figs. 9(b) and 9(c)], and that the impurity modes due to Ga and Sb atoms are not distinguishable from the zone-boundary phonon modes of Ge (such as LO:X, which is degenerate with LA:X, and LO:L). This is consistent with the data [Fig. 9(a)], in which the GaSb-like LO mode is no longer detectable for $x > 0.75$.

B. Alloy dependence of linewidths and entropy

In these alloys for $0.05 < x < 0.95$ the phonon Raman linewidths are due primarily to disorder, and not to damping. For example, the Ge-like LO mode (or Ge–Ge bond vibration) is nearly degenerate with vibrational modes of the Ga–Ge and Ga–Ga bonds, because Ge and Ga have similar masses. These Ga-related modes arise in the alloy and give the appearance of broadening the Ge-like LO mode. Thus the alloy linewidth is due primarily to inhomogeneous broadening by disorder-related alloy modes.

The linewidths in the phase-transition model appear broader than those in the on-site model (Figs. 3 and 4), because of the presence of the additional modes associated with antisite-defect Ga–Ga and Sb–Sb bond vibrations. Thus the alloy dependence of the Raman linewidths

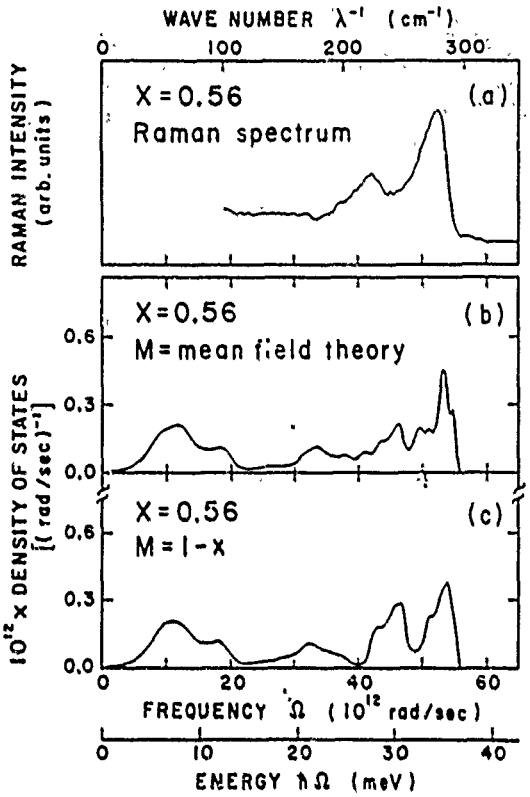


FIG. 8. (a) Raman spectrum for $(GaSb)_{0.44}(Ge_2)_{0.56}$ from Refs. 6 and 7; (b) calculated density of phonon states for the same alloy obtained by the recursion method in the case of M determined by mean-field theory; and (c) calculated density of states in the case of $M = 1-x$.

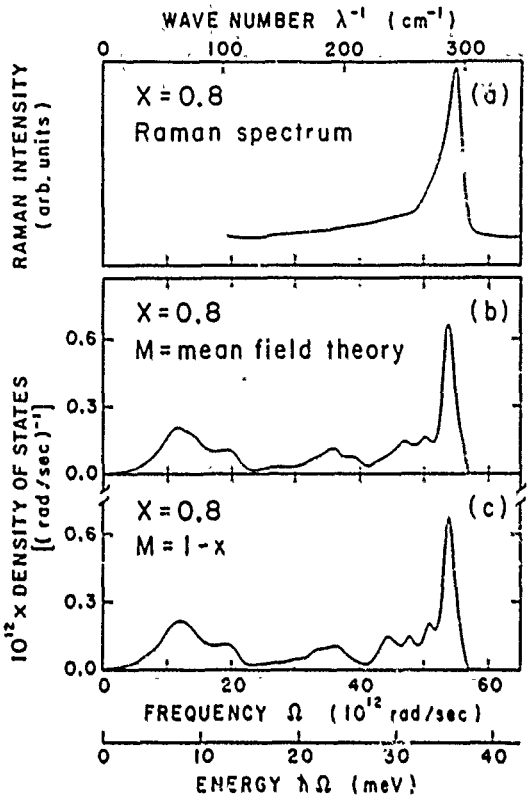


FIG. 9. (a) Raman spectrum for $(GaSb)_{0.20}(Ge_2)_{0.80}$ from Refs. 6 and 7; (b) calculated density of phonon states for the same alloy obtained by the recursion method in the case of M determined by mean-field theory, and (c) calculated density of states in the case of $M = 1-x$.

should be a good indicator of whether such antisite defects are present in significant concentrations.

The observed Raman peaks are due to a superposition of nearly degenerate modes and (according to the theory) would be highly structured in the limit of zero broadening and anharmonicity. Therefore it is difficult to define theoretically which substructures of a peak should be considered part of that peak when computing the peak's linewidth. Moreover the present theory does not include either Raman matrix elements or the effects of anharmonicity and broadening—making it impractical to define and predict quantitatively linewidths corresponding the observed Raman peak widths. Hence we seek a semiquantitative measure of the linewidths in terms of the *entropy*, which can be compared with the data.

Qualitatively the composite linewidth of the Ge-like LO mode and its adjacent alloy modes will be maximum when the disorder is maximum. In other words, the linewidth qualitatively reflects the disorder or entropy of the alloy. To illustrate this point, we compare in Fig. 10 the entropy per site[4] as a function of x calculated in the phase-transition model using mean-field theory:

$$S(x)/k_B = \langle P_{Ga}\rangle_c \ln\langle P_{Ga}\rangle_c + \langle P_{As}\rangle_c \ln\langle P_{As}\rangle_c + \langle P_{Ge}\rangle_c \ln\langle P_{Ge}\rangle_c$$
$$= [(1-x+M)/2]\ln[(1-x+M)/2] + [(1-x-M)/2]\ln[(1-x-M)/2] + x\ln(x)\,, \tag{3.1}$$

with the Raman linewidths of Ref. 7. Here k_B is Boltzmann's constant. Note the kink in the data [Fig. 10(a)] near $x \simeq 0.3$ for both the GaSb-like and the Ge-like LO modes. Furthermore, the width of the Ge-like LO mode has an x dependence similar to that of the entropy of the phase-transition model [Fig. 10(a)]. (The GaSb-like LO mode for $x > 0.3$ merges with the mostly Ge modes of Figs. 3 and 4 cannot be separated from the other modes, and hence a similar comparison of its linewidth with entropy cannot be made.)

For comparison, we also show in Fig. 10(b) the corresponding entropy per site computed for the phase-transition model with $x_c = 0.2$ and 0.7, respectively, and the entropy per site for the on-site model with no phase transition:

$$S(x)/k_B = (1-x)\ln(1-x) + x\ln(x)\,. \tag{3.2}$$

Only the entropy of the phase-transition model with $x_c \simeq 0.3$ has a kink and a maximum at the same composition as the maximum Raman linewidth.

In the absence of a phase transition, the entropy is a maximum at the composition x for which the probabili-

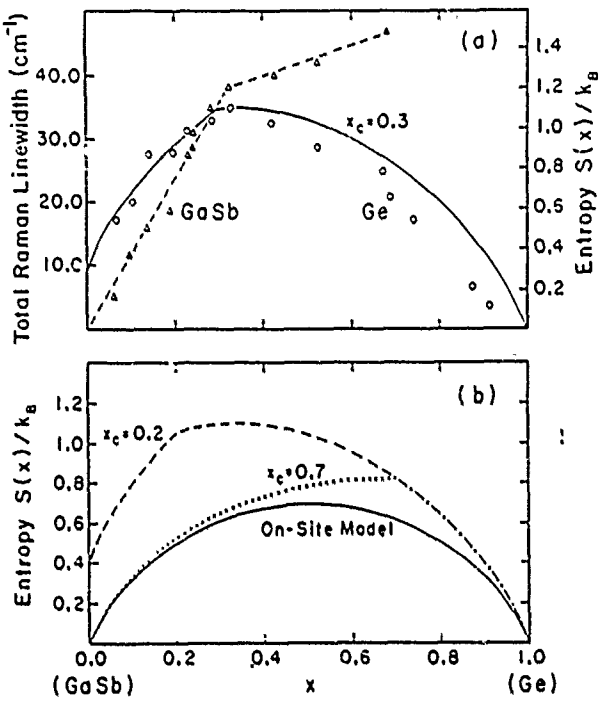


FIG. 10. (a) Total linewidth in cm^{-1} as a function of composition x for the GaSb-like (triangles) and Ge-like (circles) LO modes, after Refs. 7 and 3. The GaSb-like mode has been fit in Ref. 7 to two straight lines (dashed curve). The entropy S per site (divided by Boltzmann's constant) as a function of x, calculated using mean-field theory for $(GaSb)_{1-x}(Ge_2)_x$ and assuming $x_c=0.3$ (solid curve) is plotted on the scale on the right-hand side of the figure. The scales have been chosen such that the Ge LO mode linewidth and $S(x)/k_B$ coincide at their maxima. (b) Entropy S per site as a function of x, calculated using the on-site model of Eq. (2.11) (solid line) and the phase-transition model with $x_c=0.2$ (dashed line) and 0.7 (dotted line). Note [in (a)] the similar shapes of $S(x)$ and the linewidth of the Ge-like LO mode. Note also the kink discontinuities in $S(x)$ at x_c (characteristic of a phase transition) that are seen both in the theory and in the experimental linewidth curves at $x\simeq 0.3$ to 0.4. The maximum of $S(x)$ would not necessarily occur at $x\simeq 0.3$ if the critical composition were different (b). For $x_c<\frac{1}{3}$ the discontinuity in S occurs at x_c and the maximum occurs at $x=\frac{1}{3}$, as demonstrated for $x_c=0.2$ [dashed line of (b)]; for $x_c>\frac{1}{3}$, the maximum occurs at $x=x_c$, as demonstrated for $x_c=0.7$ [dotted line of (b)].

ties are equal for finding the various allowed substitutional atoms on a given site. For the on-site model (i.e., for $M=1-x$), the maximum is at $x=\frac{1}{2}$. For the phase-transition model, for $x_c<\frac{1}{3}$, the maximum is at $x=\frac{1}{3}$, the composition at which there are equal numbers of Ga, Sb, and Ge atoms on a given site. Additionally, for the phase-transition model, the entropy exhibits a kink at $x=x_c$ characteristic of a second-order transition, as calculated in a mean-field approximation (the second derivative of the Gibbs free energy is discontinuous).[34] If we have $x_c>\frac{1}{3}$, this kink represents the maximum of the entropy.

Clearly, of the calculated curves presented in Fig. 10, only the result for the phase-transition model with

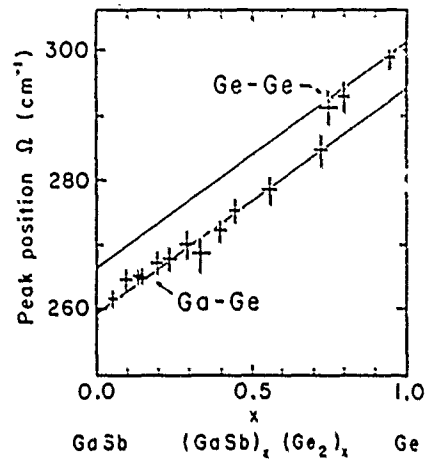


FIG. 11. Illustrating our interpretation of the discontinuity as a function of x in the Raman peak position of the Ge-like LO mode, as observed in Ref. 7: There are two principal bond vibrations, Ge–Ge and Ga–Ge. We have drawn parallel lines through the data for Ge–Ge and Ge–Ga modes, separated by $\simeq 7\ cm^{-1}$, the separation predicted for $x=0.5$. Note that the theory, which does not include long-ranged forces, does not accurately predict the positions or slopes of these lines, but only predicts the splitting between them.

$x_c=0.3$ [Fig. 10(a)] exhibits the same qualitative dependence on alloy composition x as the observed Ge-like LO mode width. The fact that the observed linewidth of the Ge-like LO mode exhibits the same qualitative dependence on x as the phase-transition model's entropy, namely that it has a maximum for $x\simeq 0.3$, is evidence supporting the phase-transition theory—and against the on-site model. Clearly $(GaSb)_{1-x}(Ge_2)_x$ grown under conditions such that x_c is greater than 0.3 (Ref. 30) should exhibit a maximum Raman linewidth at the different critical composition x_c.

C. Peak positions

Beserman *et al.*[7] noted that the peak position of their broad Raman line for the Ge-like LO mode, when plotted as a function of x, exhibits a discontinuity of $\simeq 7\ cm^{-1}$ near $x\simeq 0.8$ (see Fig. 11). The present theory (Figs. 3 and 4) provides a simple and natural explanation of this discontinuity: there are two types of bonds contributing significant but separate subpeaks to the Ge-like LO mode Raman line, Ge–Ge and Ge–Ga bonds. In GaSb-rich material ($x<0.7$) the main spectral feature observed is dominated by the Ge–Ga bonds, but in Ge-rich alloys ($x>0.7$) the Ge–Ge peak dominates. The theory shows that these two subpeaks are separated of order $7.4\ cm^{-1}$.

D. Asymmetries of Raman lines

The Raman data for $(GaSb)_{1-x}(Ge_2)_x$ are abnormally asymmetric, as shown in Ref. 7: the GaSb-like (Ge-like) LO mode peak has maximum asymmetry for $x\simeq 0.2$ ($x\simeq 0.8$) and these asymmetries are *not* smoothly varying functions of x.

These experimental facts are simply explained: the

GaSb-like Raman peak has contributions from the Sb–Sb bond subpeak on its low-energy side (see Figs. 6 and 7), which is especially visible for $x<0.5$. Likewise Ga–Ge bonds contribute on the low-energy side of the Ge-like LO Raman line for $x>0.7$ (see Fig. 11). The fact that the asymmetries are not smoothly varying functions of x indicates that different modes from the main LO modes are responsible.

E. LO-TO splitting

The zone-center optical phonons of a crystal experience a Lyddane-Sachs-Teller splitting that should be proportional to the square of the average dipole moment per unit cell.[16] In an alloy, this prescription has not been fully justified (except in a virtual-crystal sense); nevertheless, we employ it here to predict the x dependences of the $k=0$ LO and TO modes in $(GaSb)_{1-x}(Ge_2)_x$. The dipole moment per unit cell is proportional to the order parameter $M(x)$, and the proportionality constant can be determined by fitting the observed splitting for $x=0$. The order parameter, in a mean-field theory, is approximately proportional to $(x_c-x)^{1/2}$, indicating that in the mean-field approximation the LO-TO splitting of the GaSb-like $k=0$ optic mode should decrease approximately linearly from $x=0$ to $x=0.3$. Such a decrease in the LO-TO splitting is indeed observed, although the decrease is not necessarily linear in x_c-x (see Fig. 12). [One expects that more accurate calculations using the renormalization group or some similar technique would determine that the LO-TO splitting decreases as $(x_c-x)^{2\beta}$, with $\beta\simeq 0.325$.[35]] The bifurcation of the GaSb-like $k=0$ mode at $x\simeq 0.3$ is a consequence of the change of symmetry from a zinc-blende structure for $x<x_c$, in which there is a net dipole moment per unit cell, to a diamond structure for $x>x_c$, in which there is no average dipole moment and no distinction between anion and cation sites. This is a general feature of the phase-transition model.

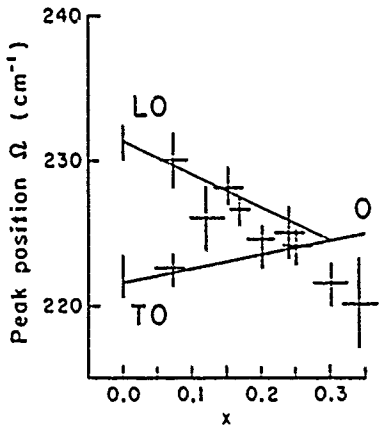


FIG. 12. Raman peak positions Ω (in cm^{-1}) of the GaSb $k=0$ LO and TO modes in $(GaSb)_{1-x}(Ge_2)_x$ versus alloy composition x, after Refs. 7 and 8, compared with the predictions of the phase-transition model plus a Lyddane-Sachs-Teller LO-TO splitting proportional to the square of the order parameter M. The bifurcation at $x_c\simeq 0.3$ of the diamond-phase optic mode (O) into LO and TO modes is characteristic of the order-disorder zinc-blende–diamond phase transition. Here the theory is shifted down $\simeq 4$ cm^{-1} to coincide with the data for GaSb, but the variation of the peak position with x, $d\Omega/dx$, is not adjusted to account for long-ranged forces omitted from the model, and so is not accurately predicted.

Certainly the bifurcation will not be as sharp as predicted in this simple theory, and alloy fluctuations will broaden the GaSb-like optic modes, especially for $x\simeq x_c$. Nevertheless, the data exhibit this splitting, which is evidence of the phase transition. Indeed, the zone-center LO-TO splitting offers a means to directly determine the square of the order parameter.

IV. SUMMARY

To our knowledge, the theory presented here is the first comprehensive treatment of phonons in $(A^{III}B^{V})_{1-x}(C_2^{IV})_x$ metastable, substitutional, crystalline alloys. We have presented predictions for both the phase-transition model and the on-site model, and have compared them with data. While the comparison favors the phase-transition model, the evidence is not compelling. We suspect that a correct model of these alloys embraces the essentials of the phase-transition model, but goes beyond the mean-field approximation to include correlations in the atomic positions. Such an improved theory will still have Sb–Sb and Ga–Ga bonds, but they will be fewer in number than for the present mean-field theory.[36]

The main features of the Raman data are well described by the theory, despite the fact that many of those features were previously thought to be anomalous. The Raman linewidths are a measure of the alloy disorder and vary with x in a similar fashion to the entropy of the phase transition model. The observed dependence of the Raman linewidth appears to be inconsistent with the on-site model. The apparent discontinuity, as a function of x, of the Ge-like LO Raman mode is attributed to the fact that separate but nearly degenerate modes associated with two distinct types of bonds, Ge–Ge and Ga–Ge, contribute to the Raman peak—and the dominant bond changes with x. The apparent anomalous asymmetry of the Raman lines is due to Sb–Sb and Ga–Ge sideband modes.[37] Finally, the order parameter of the zinc-blende–diamond phase transition can be extracted from the Lyddane-Sachs-Teller splitting of the GaSb LO and TO phonons at $k=0$.

In general the model accounts for all the essential physics of the spectra of these alloys. It demonstrates that virtual-crystal and persistence approximations to the phonon spectra are inadequate. Nevertheless, the theory itself requires improvement to make it truly quantitative, and long-ranged forces should be incorporated into the model. We hope that this work will stimulate further theoretical and experimental investigations of phonons in these interesting alloys.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support. We have enjoyed stimulating conversations with L. Abels, R. Besserman, T. Kuabach, T. McGlinn, and P. M. Raccah.

APPENDIX A: SYMMETRY OF FORCE-CONSTANT MATRICES

Choosing the origin at an ion, the positions of its four nearest neighbors are $\tau_1=(a_L/4)(1,1,1)$, $\tau_2=(a_L/4)(-1,-1,1)$, $\tau_3=(a_L/4)(1,-1,-1)$, and $\tau_4=(-1,1,-1)$. Similarly, the positions of the 12 second nearest neighbors are given by the nonzero combinations of $\tau_i-\tau_j$, for i and j ranging from 1 to 4.

Applying the symmetry of the crystallographic point group T_d, with an x, y, and z basis, the first-nearest-neighbor force constants between the bth ion at the origin and the b'th ion at the position τ_1 are parametrized as in Eq. (2.3), and the second-nearest-neighbor force constants between the bth ions at the origin and at the position $(a_L/2)(0,1,1)$ are parametrized as in Eq. (2.4). In this appendix, we use the notation $\underline{\Phi}^{(1)}$ and $\underline{\Phi}^{(2)}$ for the 3×3 force-constant matrices given by Eqs. (2.3) and (2.4), respectively.

The force-constant matrix between an ion at the origin and an ion at the position τ_2 can be obtained from $\underline{\Phi}^{(1)}$ by rotating the coordinate system by π around the z axis:

$$\underline{R}_z^{\dagger}\underline{\Phi}^{(1)}\underline{R}_z=\begin{vmatrix}\alpha & \beta & -\beta\\ \beta & \alpha & -\beta\\ -\beta & -\beta & \alpha\end{vmatrix}, \tag{A1}$$

where we have the rotation matrix in the form:

$$\underline{R}_z=\begin{vmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{vmatrix}. \tag{A2}$$

The force-constant matrices between an ion at the origin and ions at the positions τ_3 and τ_4 can be obtained by similar rotation operations.

The force-constant matrix between the bth ions at the origin and at the position $(a_L/2)(0,-1,1)$ is derived from $\underline{\Phi}^{(2)}$ by operating with $\underline{R}_z$:

$$\underline{R}_z^{\dagger}\underline{\Phi}^{(2)}\underline{R}_z=\begin{vmatrix}\lambda_b & 0 & 0\\ 0 & \mu_b & -\nu_b\\ 0 & -\nu_b & \mu_b\end{vmatrix}. \tag{A3}$$

The other ten force-constant matrices for the second-nearest neighbors can be easily obtained by similar symmetry operations.

These force-constant matrices can be written in a compact form using the direction cosines (l,m,n) of the vector connecting the two ion positions:

$$\underline{\Phi}^{(1)}(l,m,n)=\begin{vmatrix}\alpha & 3lm\beta & 3ln\beta\\ 3lm\beta & \alpha & 3mn\beta\\ 3ln\beta & 3mn\beta & \alpha\end{vmatrix} \tag{A4}$$

for first-nearest neighbors, and

$$\underline{\Phi}^{(2)}(l,m,n)=\begin{vmatrix}\lambda_b(1-2l^2)+2\mu_b l^2 & 2lm\nu_b & 2ln\nu_b\\ 2lm\nu_b & \lambda_b(1-2m^2)+2\mu_b m^2 & 2mn\nu_b\\ 2ln\nu_b & 2mn\nu_b & \lambda_b(1-2n^2)+2\mu_b n^2\end{vmatrix} \tag{A5}$$

for second-nearest neighbors. The direction cosines (l,m,n) take the values, for instance, $(-3^{-1/2},3^{-1/2},-3^{-1/2})$ for $(a_L/4)(-1,1,-1)$ and $(0,2^{-1/2},2^{-1/2})$ for $(a_L/2)(0,1,1)$.

In using the recursion method, wherein a finite cluster of atoms is generated, it is convenient to use the above 3×3 force-constant matrices in terms of the direction cosines, for each pair of atoms.

APPENDIX B: DYNAMICAL MATRIX

With use of the force-constant parameters introduced in Sec. II, the 6×6 dynamical matrix $\underline{D}(\mathbf{k})$ for each k vector in Eq. (2.9) can be written as

$$\underline{D}(\mathbf{k})=\begin{array}{c}\\ b/b'\\ a\\ c\end{array}\begin{array}{c}\begin{array}{cc}\quad a & \quad c\end{array}\\ \begin{vmatrix}\underline{D}_a & \underline{D}_{ac}\\ \underline{D}_{ac}^{\dagger} & \underline{D}_c\end{vmatrix}\end{array}. \tag{B1}$$

Each element of this matrix is a 3×3 matrix. The on-site matrix $\underline{D}_a$ for the anion is given by

$$(-M_a/4)\underline{D}_a=\begin{array}{c}\\ \langle a,x|\\ \langle a,y|\\ \langle a,z|\end{array}\begin{array}{c}\begin{array}{ccc}\quad |a,x\rangle & \quad |a,y\rangle & \quad |a,z\rangle\end{array}\\ \begin{vmatrix}\alpha+\lambda_a f_1(\mathbf{k})+\mu_a g_1(\mathbf{k}) & \nu_a h_1(\mathbf{k}) & \nu_a h_2(\mathbf{k})\\ \nu_a h_1(\mathbf{k}) & \alpha+\lambda_a f_2(\mathbf{k})+\mu_a g_2(\mathbf{k}) & \nu_a h_3(\mathbf{k})\\ \nu_a h_2(\mathbf{k}) & \nu_a h_3(\mathbf{k}) & \alpha+\lambda_a f_3(\mathbf{k})+\mu_a g_3(\mathbf{k})\end{vmatrix}\end{array}, \tag{B2}$$

where (with a_L being the lattice constant)

$$f_1(\mathbf{k})=1-\cos(\pi a_L k_y)\cos(\pi a_L k_z)\ ,$$
$$f_2(\mathbf{k})=1-\cos(\pi a_L k_z)\cos(\pi a_L k_x)\ ,$$
$$f_3(\mathbf{k})=1-\cos(\pi a_L k_x)\cos(\pi a_L k_y)\ ,$$
$$g_1(\mathbf{k})=2-\cos(\pi a_L k_x)[\cos(\pi a_L k_y)+\cos(\pi a_L k_z)]\ ,$$
$$g_2(\mathbf{k})=2-\cos(\pi a_L k_y)[\cos(\pi a_L k_z)+\cos(\pi a_L k_x)]\ ,$$
$$g_3(\mathbf{k})=2-\cos(\pi a_L k_z)[\cos(\pi a_L k_x)+\cos(\pi a_L k_y)]\ ,$$
$$h_1(\mathbf{k})=\sin(\pi a_L k_x)\sin(\pi a_L k_y)\ ,$$
$$h_2(\mathbf{k})=\sin(\pi a_L k_x)\sin(\pi a_L k_z)\ ,$$
$$h_3(\mathbf{k})=\sin(\pi a_L k_y)\sin(\pi a_L k_z)\ .$$

The on-site matrix, $\underline{D}_c$, for the cation, is identical in form to $\underline{D}_a$, except that the index a is everywhere replaced by c. The off-site matrix $\underline{D}_{ac}$ is given by

$$(M_a M_c)^{-1/2}\underline{D}_{ac}=\begin{array}{c} \\ \langle a,x| \\ \langle a,y| \\ \langle a,z| \end{array}\begin{array}{c}\begin{array}{cccc}|c,x\rangle & |c,y\rangle & |c,z\rangle\end{array}\\ \begin{vmatrix} \alpha s_1(\mathbf{k}) & \beta s_2(\mathbf{k}) & \beta s_3(\mathbf{k}) \\ \beta s_2(\mathbf{k}) & \alpha s_1(\mathbf{k}) & \beta s_4(\mathbf{k}) \\ \beta s_3(\mathbf{k}) & \beta s_4(\mathbf{k}) & \alpha s_1(\mathbf{k}) \end{vmatrix}\end{array}\ , \tag{B3}$$

where we have

$$s_1(\mathbf{k})=e_1+e_2+e_3+e_4\ ,$$
$$s_2(\mathbf{k})=e_1+e_2-e_3-e_4\ ,$$
$$s_3(\mathbf{k})=e_1-e_2-e_3+e_4\ ,$$
$$s_4(\mathbf{k})=e_1-e_2+e_3-e_4\ .$$

The e_i's ($i=1$, 2, 3, and 4) are given by

$$e_1=\exp[i\pi a_L(k_x+k_y+k_z)/2]\ ,$$
$$e_2=\exp[i\pi a_L(-k_x-k_y+k_z)/2]\ ,$$
$$e_3=\exp[i\pi a_L(k_x-k_y-k_z)/2]\ ,$$
$$e_4=\exp[i\pi a_L(-k_x+k_y-k_z)/2]\ .$$

Despite the importance of the long-ranged Coulomb forces, we ignore these forces and use the rigid-ion model with first- and second-nearest-neighbor force constants for describing the lattice dynamics of semiconducting compounds.

APPENDIX C: FITTING PROCEDURE

At points of high symmetry in the Brillouin zone, the dynamical matrix becomes block diagonal. We can express the frequencies at these points in terms of the force-constant parameters, and determine these parameters by fitting the frequencies with experimental data. The analytic forms at the Γ, X, and L points are

$$\Omega^2(\mathrm{LO};\Gamma)=\Omega^2(\mathrm{TO};\Gamma)=-4\alpha(1/M_a+1/M_c)\ , \tag{C1}$$

$$\Omega^2(\mathrm{LO};X),\ \Omega^2(\mathrm{LA};X)=\begin{cases}-4(\alpha+4\mu_a)/M_a \\ -4(\alpha+4\mu_c)/M_c\ ,\end{cases} \tag{C2}$$

$$\begin{aligned}\Omega^2(\mathrm{TO};X),\ \Omega^2(\mathrm{TA};X)=&-[2\alpha(1/M_a+1/M_c)+4(\lambda_a+\mu_a)/M_a+4(\lambda_c+\mu_c)/M_c)] \\ &\pm[\,|2\alpha(1/M_a-1/M_c)+4(\lambda_a+\mu_a)/M_a-4(\lambda_c+\mu_c)/M_c|^2 \\ &+16\beta^2/(M_aM_c)]^{1/2}\ \text{(twofold degeneracy)}\ ,\end{aligned} \tag{C3}$$

$$\begin{aligned}\Omega^2(\mathrm{LO};L),\ \Omega^2(\mathrm{LA};L)=&-[2\alpha(1/M_a+1/M_c)+2(\lambda_a/M_a+\lambda_c/M_c)+4(\mu_a/M_a+\mu_c/M_c)+4(\nu_a/M_a+\nu_c/M_c)] \\ &\pm[\,|2\alpha(1/M_a-1/M_c)+2(\lambda_a/M_a-\lambda_c/M_c) \\ &+4(\mu_a/M_a-\mu_c/M_c)+4(\nu_a/M_a-\nu_c/M_c)|^2+4(-\alpha+2\beta)^2/(M_aM_c)]^{1/2}\ ,\end{aligned} \tag{C4}$$

and

$$\begin{aligned}\Omega^2(\mathrm{TO};L),\ \Omega^2(\mathrm{TA};L)=&-[2\alpha(1/M_a+1/M_c)+2(\lambda_a/M_a+\lambda_c/M_c)+4(\mu_a/M_a+\mu_c/M_c)-2(\nu_a/M_a+\nu_c/M_c)] \\ &\pm[\,|2\alpha(1/M_a-1/M_c)+2(\lambda_a/M_a-\lambda_c/M_c) \\ &+4(\mu_a/M_a-\mu_c/M_c)-2(\nu_a/M_a-\nu_c/M_c)|^2 \\ &+4(\alpha+\beta)^2/(M_aM_c)]^{1/2}\ \text{(twofold degeneracy)}\ .\end{aligned} \tag{C5}$$

In addition, by taking the elastic-continuum limit of the equation of motion,[38] we have the following relations between elastic constants and force-constant parameters:

$$c_{11}=(-1/a_L)[\alpha+4(\mu_a+\mu_c)]\ , \tag{C6}$$

$$\begin{aligned}c_{12}=(-1/a_L)[&2\beta-\alpha+4(\nu_a+\nu_c)-2(\lambda_a+\lambda_c) \\ &-2(\mu_a+\mu_c)]\ ,\end{aligned} \tag{C7}$$

$$c_{44}=(-1/a_L)[\alpha+2(\lambda_a+\lambda_c)+2(\mu_a+\mu_c)-\beta^2/\alpha]\ . \tag{C8}$$

where a_L is the lattice constant.

In determining the parameter α from phonon dispersion curves, the squared frequency, $\Omega^2(\mathrm{LO};\Gamma)$ [$=\Omega^2(\mathrm{TO};\Gamma)$ in the present model], needs to be replaced by $[\Omega^2(\mathrm{LO};\Gamma)+2\Omega^2(\mathrm{TO};\Gamma)]/3$, since the LO-TO splitting exists in the experimental data for these mode frequencies. With this replacement in Eq. (C1), we have

$$\alpha=-[M_a M_c/4(M_a+M_c)][\Omega^2(\mathrm{LO};\Gamma)+2\Omega^2(\mathrm{TO};\Gamma)]/3 . \tag{C9}$$

The parameters μ_a and μ_c are determined from Eq. (C2) as

$$\begin{aligned}\mu_a&=(-M_a/16)\Omega^2(\mathrm{LA};X)-\alpha/4 ,\\ \mu_c&=(-M_c/16)\Omega^2(\mathrm{LO};X)-\alpha/4 ,\end{aligned} \tag{C10}$$

for $M_a > M_c$ and

$$\begin{aligned}\mu_a&=(-M_c/16)\Omega^2(\mathrm{LO};X)-\alpha/4 ,\\ \mu_c&=(-M_c/16)\Omega^2(\mathrm{LA};X)-\alpha/4 ,\end{aligned} \tag{C11}$$

for $M_a < M_c$.

The determination of the parameters λ_a, λ_c, ν_a, ν_c, and β, using the relations given by Eqs. (C1) to (C8), depends on the data available for each material. In some cases, a reduction of the number of parameters is necessary, because of lack of data. If we assume a central force operating between ions, we have

$$\mu_b=\nu_b+\lambda_b \quad (b=a \text{ and } c) . \tag{C12}$$

Banerjee and Varshni[22] have assumed the relation

$$\mu_b=\nu_b \quad (b=a \text{ and } c) \tag{C13}$$

in concurrence with the central-force model[39] and the angular-force model.[40]

Various combinations of the relations in Eqs. (C1) to (C8), with or without the simplifications given in Eq. (C12) or (C13), were attempted. (Our parameters were also compared with those obtained by Banerjee and Varshni.[22]) Judgment was made according to how well the phonon dispersion relations and global features of the densities of states were reproduced. For GaSb, the parameters for GaAs obtained by Banerjee and Varshni[22] were found to be excellent. Here we set one parameter, representing the long-range Coulomb forces in their model, to zero.) We use these parameters for GaSb as well, replacing the mass of As with the mass of Sb in the dynamical matrix. For Ge, using the relation in Eq. (C13), the four parameters α, β, μ ($=\nu$), and λ were determined by fitting $\Omega^2(\mathrm{LO};\Gamma)$ [$=\Omega^2(\mathrm{TO};\Gamma)$], $\Omega^2(\mathrm{TO};X)$, $\Omega^2(\mathrm{LO};X)$ [$=\Omega^2(\mathrm{LA};X)$], and $\Omega^2(\mathrm{TA};X)$ to neutron scattering data.[25] There is, of course, no distinction between anion and cation, and thus we have $\lambda_a=\lambda_c=\lambda$, for example.

APPENDIX D: MINICLUSTER PROBABILITIES

The probability of occurrence of a five-atom minicluster in which the central atom is a Ga atom on the nominal cation-site, with l Ga atoms, m Sb atoms, and n Ge atoms distributed over the four nearest-neighbor nominal anion sites of the central Ga atom is

$$\langle P_{\mathrm{Ga}}\rangle_c^l(\langle P_{\mathrm{Ga}}\rangle_a)^l(\langle P_{\mathrm{Sb}}\rangle_a)^m(\langle P_{\mathrm{Ge}}\rangle_a)^n . \tag{D1}$$

Such a cluster of neighbors occurs $4!/(l!m!n!)$ times. The probabilities of occurrence of various nearest-neighbor clusters were calculated using Eqs. (2.14a)–(2.14e). In case Ga atoms are allowed to occupy the nominal anion sites, as in the model using mean-field theory, the probabilities are identical in form, except that the indices a and c need to be interchanged everywhere. The probabilities of occurrence of five-atom miniclusters with a Sb atom on the central site, and with a Ge atom on the central site, are obtained by a cyclic permutation of the indices specifying atoms: from (Ga,Sb,Ge) to (Sb,Ge,Ga), and to (Ge,Ga,Sb) respectively. There are, therefore, $15\times2\times3$ statistically independent configurations in the zinc-blende phase, and 15×3 statistically independent configurations in the diamond phase (since we do not distinguish between anions and cations in this phase). The probabilities for the model with no antisite defects ($M=1-x$) can be obtained from Eqs. (2.11a)–(2.11c).

---

*Present address: University of Maryland, College Park, MD 20742.

[1]J. E. Greene, J. Vac. Sci. Technol. B 1, 229 (1983); J. L. Zilko and J. E. Greene, J. Appl. Phys. 51, 1549 (1980); 51, 1560 (1980).

[2]K. C. Cadien, A. H. Eltoukhy, and J. E. Greene, Appl. Phys. Lett. 38, 773 (1981).

[3]K. E. Newman, A. Lastras-Martinez, B. Kramer, S. A. Barnett, M. A. Ray, J. D. Dow, J. E. Greene, and P. M. Raccah, Phys. Rev. Lett. 50, 1466 (1983).

[4]K. E. Newman and J. D. Dow, Phys. Rev. B 27, 7495 (1983).

[5]S. A. Barnett, B. Kramer, L. T. Romano, S. I. Shah, M. A. Ray, S. Fang, and J. E. Greene, *Layered Structures, Epitaxy, and Interfaces*, edited by J. M. Gibson and L. R. Dawson (North-Holland, Amsterdam, 1984).

[6]T. N. Krabach, N. Wada, M. V. Klein, K. C. Cadien, and J. E. Greene, Solid State Commun. 45, 895 (1983).

[7]R. Beserman, J. E Greene, M. V. Klein, T. N. Krabach, T. C. McGlinn, L. T. Romano, and S. I. Shah, in *Proceedings of the 17th International Conference on the Physics of Semiconductors, San Francisco* (Springer, New York, 1985), p. 961.

[8]See also K. E. Newman, J. D. Dow, A. Kobayashi, and R. Beserman (unpublished).

[9]L. Abels and P. M. Raccah (unpublished).

[10]R. Haydock, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35, p. 215, and references therein.

[11]C. M. M. Nex, J. Phys. A 11, 653 (1978); Comput. Phys. Commun. 34, 101 (1984).

[12]M. J. Kelly, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35, p. 296; V. Heine, in *Solid State Physics*, edited by H. Ehren-

reich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35, p. 1.
[13]C. Herscovici and M. Fibich, J. Phys. C 13, 1635 (1980).
[14]A. Kobayashi, J. D. Dow, and E. P. O'Reilly, Superlattices Microstruct. 1, 530 (1985).
[15]A. Kobayashi, Ph.D. thesis, University of Illinois, 1985.
[16]R. Lyddane, R. G. Sachs, and E. Teller, Phys. Rev. 59, 218 (1941).
[17]B. G. Dick, Jr. and A. W. Overhauser, Phys. Rev. 112, 90 (1958).
[18]W. Cochran, Proc. R. Soc. London, Ser. A 253, 260 (1959).
[19]W. Weber, Phys. Rev. B 15, 4789 (1977).
[20]P. Vogl, J. Phys. C 11, 251 (1978).
[21]G. Lehmann and M. Taut, Phys. Status Solidi B 54, 469 (1972).
[22]R. Banerjee and Y. P. Varshni, Can. J. Phys. 47, 451 (1969).
[23]M. Hass and B. W. Henvis, J. Phys. Chem. Solids 23, 1099 (1962).
[24]M. K. Farr, J. G. Traylor, and S. K. Sinha, Phys. Rev. B 11, 1587 (1975).
[25]G. Nelin and G. Nilsson, Phys. Rev. B 5, 3151 (1972).
[26]L. Van Hove, Phys. Rev. 89, 1189 (1953).
[27]The mean-field theory treats $k=0$ properties adequately. But the short-range order, to be treated properly, requires a more complete treatment of correlations that will reduce the number of Sb—Sb and Ga—Ga bonds below the mean-field-theory value.
[28]H. Holloway and L. C. Davis, Phys. Rev. Lett. 53, 1510 (1984).
[29]The correct percolation composition is $x_c=0.572$. See, for example, R. Zallen, *The Physics of Amorphous Solids* (Wiley, New York, 1983).
[30]Z. I. Alferov, R. S. Vartanyan, V. I. Korol'kov, I. I. Mokan, V. P. Ulin, B. S. Yavich, and A. A. Yakovenko, Fiz. Tekh. Poluprovodn. 16, 887 (1982) [Sov. Phys.—Semicond. 16, 567 (1982)].
[31]Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968).
[32]W. Hayes and R. Loudon, *Scattering of Light by Crystals* (Wiley, New York, 1978).
[33]R. M. Martin and F. L. Galeener, Phys. Rev. B 23, 3071 (1981).
[34]See, for example, L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980), Chap. 9.
[35]J. C. Le Guillou and J. Zinn-Justin, Phys. Rev. B 21, 3976 (1980).
[36]B. Koiller, M. A. Davidovitch, and R. Osorio (unpublished).
[37]Since the existence of Ga—Ga and Sb—Sb bonds is an important feature in the phase-transition model, it is crucial to ascertain whether the tail near 195 cm^{-1} in the data does arise from vibrations of Sb—Sb bonds. In order to clarify the issue, we suggest that the feature near 195 cm^{-1} be investigated in detail by resonant Raman scattering measurements. We also emphasize that the mean-field approximation may overestimate the number of Sb—Sb bonds.
[38]P. Brüesch, *Phonons. Theory and Experiments I* (Springer, Berlin, 1982).
[39]H. M. J. Smith, Phys. Trans. R. Soc. London, Ser. A 241, 105 (1948).
[40]R. Braunstein, F. Herman, and A. R. Moore, Phys. Rev. 109, 695 (1958).

Reprinted from

Materials Science Forum

Volume 4 (1985)

CHEMICAL TRENDS OF SCHOTTKY BARRIERS

John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556

and

Otto F. Sankey
Department of Physics, Arizona State University
Tempe, Arizona 85287

and

Roland E. Allen
Department of Physics, Texas A&M University
College Station, Texas 77843

TRANS TECH PUBLICATIONS
SWITZERLAND - GERMANY - UK - USA

Materials Science Forum, Volume 4 (1985), pp. 39-50.
Copyright© 1985 by Trans Tech Publications Ltd., Switzerland

CHEMICAL TRENDS OF SCHOTTKY BARRIERS

John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556

and

Otto F. Sankey
Department of Physics, Arizona State University
Tempe, Arizona 85287

and

Roland E. Allen
Department of Physics, Texas A&M University
College Station, Texas 77843

ABSTRACT

The observed chemical trends in Schottky barrier heights (i.e., the variation in the barrier height ϕ_B as a function of the alloy composition x, or the dependence of the barrier height on the anion or cation species) are explained by Fermi-level pinning due to defects. Microscopic calculations of surface defect levels, rather than phenomenological arguments, are presented to support this viewpoint. We find that the slope of the pinning defect level as a function of alloy composition (dE/dx) is a signature of the detect type. In the case of $Al_xGa_{1-x}As$/Au contacts for all compositions x, and for $Al_xGa_{1-x}As$/Al and $Al_xGa_{1-x}As$/In contacts for large x, the Schottky barrier heights are attributed to Fermi-level pinning by cation-on-anion-site antisite defects (|dE/dx| is large). $Al_xGa_{1-x}As$/Al and $Al_xGa_{1-x}As$/In Schottky barriers, for small x, are attributed to Fermi-level pinning by anion-on-the-cation-site antisite defects (|dE/dx| is small). This interpretation is supported by both detailed calculations and the results of a simple four-atom model.

I. INTRODUCTION

Over the years, there have been many attempts to understand the observed chemical trends in Schottky barrier heights ϕ_B — the dependence of ϕ_B on the anion or cation species, or on the alloy composition x. In the past, such attempts have had the disadvantage that no fundamental microscopic foundation has been available. The introduction of Bardeen's concept of Fermi-level pinning [1] and Spicer's defect model [2-5], however, have provided a general framework that makes it possible to understand chemical trends in ϕ_B from a microscopic point of view: Since the Schottky barrier height in the Fermi-level pinning model is approximately equal to the difference between a band edge (conduction band edge for an n-type semiconductor and valence band edge for p-type) and the relevant defect level (lowest acceptor level for n-type; highest donor level for p-type), chemical trends in barrier heights are explained by the combined chemical trends in band edges and "deep" defect levels at the semiconductor/metal contact.

Recently we have reported theoretical predictions of Schottky barrier heights for Au contacts to various III-V alloys [6] and for transition-metal contacts to Si_xGe_{1-x} alloys [7]. The III-V/Au barriers are attributed to Fermi-level pinning by cation-on-anion-site III-V surface antisite defects [6]. The Si_xGe_{1-x} barriers are attributed to Fermi-level pinning by interfacial dangling bonds [7]. For both systems, the theory is in quite good agreement with the measured barrier heights, with the observed chemical trends being particularly well described by the theory.

Here we extend the simple theory of Ref. [6] and consider both antisite defects, including the the anion-on-cation-site antisite defect. Our principal motivation is to compare the alloy dependences for the Schottky barriers that result from Fermi-level pinning by the two different types of antisite defects. As discussed below, we find that dE/dx is very different for the two defects in some cases, where E is a Fermi-level pinning defect energy level and x is the alloy composition. This appears to explain the different dependences on x of observed barrier heights of $Al_xGa_{1-x}As$ with Al and In contacts on the one hand, and with Au contacts on the other hand.

II. Simple Four-Atom Model

Before giving the results of our detailed calculations -- which employ the sp^3s^* model of Vogl et al. [8] for the bulk electronic structure, the scaled-atomic energy model of Hjalmarson et al. [9] for the impurity potentials, and the analytic Green's function technique [10] -- let us consider a very simple four-atom model for each of the two surface antisite defects: the antisite atom at a surface and its three nearest-neighbors. We will find that this model provides a remarkably good description of the chemical trends, and tends to increase our faith in the central results of the much more complicated calculations.

The simple four-atom model can be constructed by first considering a five-atom model consisting of an antisite impurity in the bulk and its four neighbors, and then replacing one of the four neighboring atoms by a vacancy -- to simulate the semiconductor surface. In the bulk, an anion or cation antisite defect is tetrahedrally coordinated, which leads to deep level

electronic states of A_1 (s-like) or T_2 (p-like) symmetry. The symmetry is reduced at the surface, and the states of A_1 and T_2 symmetry mix.

A. Bulk antisite Defects

For concreteness, consider the anion-site bulk antisite defect (Ga_{As}) in GaAs. Take as a basis (i) the s- and p-orbitals of the antisite defect atoms $|s\rangle$ and $|p\rangle$ (with energies ε_s and ε_p) and (ii) the main s-like (or A_1-symmetric) and p-like (T_2-symmetric) orbitals of the rest of the solid without the central atom — namely the A_1 and T_2 orbitals of a vacancy (with energies $E(A_1;v)$ and $E(T_2;v)$). In a model which considers only the defect and its four neighbors, the vacancy A_1 orbital is

$$|A_1;v\rangle = (|1\rangle + |2\rangle + |3\rangle + |4\rangle)/2, \tag{1}$$

where $|i\rangle$ is the inward-directed sp^3-hybrid centered on the i-th neighboring site [11]. Similarly the relevant T_2-vacancy orbital is

$$|T_2;v\rangle = (12)^{-1/2} (|1\rangle + |2\rangle + |3\rangle - 3\,|4\rangle). \tag{2}$$

The s orbital of the antisite impurity only interacts with $|A_1;v\rangle$; and the p orbital which is polarized toward atom 4 interacts only with $|T_2;v\rangle$. Notice that the wavefunction is equally distributed among the four hybrids for the $|A_1;v\rangle$ orbital, but is more heavily weighted on hybrid $|4\rangle$ for the $|T_2;v\rangle$ orbital. The model bulk antisite Hamiltonian can be simply written as a direct sum:

$$H_{bulk} = \begin{pmatrix} H(A_1) & 0 \\ 0 & H(T_2) \end{pmatrix} \tag{3}$$

where we have (in the basis $|A_1;v\rangle$ and $|s\rangle$)

$$H(A_1) = \begin{pmatrix} E(A_1;v) & -t(A_1) \\ -t(A_1) & \varepsilon_s \end{pmatrix} \tag{4}$$

and (in the basis $|T_2;v\rangle$ and $|p\rangle$)

$$H(T_2) = \begin{pmatrix} E(T_2;v) & -t(T_2) \\ -t(T_2) & \varepsilon_p \end{pmatrix} \tag{5}$$

The vacancy energies $E(A_1;v)$ and $E(T_2;v)$ are obtained from Green's function calculations [9] of ideal vacancy energies, and are the eigenvalues of $H(A_1)$ and $H(T_2)$ in the limit of ε_s and ε_p being infinite [12]. The energies ε_s and ε_p are determined from atomic energy tables; for example, ε_s is 80% of the difference is s-orbital energies of Ga and As for Ga on the As site in GaAs [8,9]. The coupling parameters $t(A_1)$ and $t(T_2)$ are obtained by fitting calculated [9] bulk antisite defect levels.

B. Surface antisite Defects

We next change one of the four neighbors (atom 4) surrounding the antisite into a vacancy. This is accompl shed by allowing the antisite only to interact with the a_1 or σ-like molecular orbital, which has no amplitude on atom 4:

$$|a_1\rangle = (\sqrt{3}\,|A_1;v\rangle + |T_2;v\rangle)/2. \tag{6}$$

The surface Hamiltonian of the antisite interacting with only three neighbors becomes (in a basis $|a_1\rangle$, $|s\rangle$, and $|p\rangle$)

$$H_{surface} = \begin{pmatrix} \varepsilon(a_1) & -t_1 & -t_2 \\ -t_1 & \varepsilon_s & 0 \\ -t_2 & 0 & \varepsilon_p \end{pmatrix} \tag{7}$$

where $\varepsilon(a_1)$ is the self-energy of the remaining three sp^3-hybrid orbitals, and is given by

$$\varepsilon(a_1) = [3\,E(A_1;v) + E(T_2;v)]/4. \tag{8}$$

The s and p orbitals of the antisite interact with the remaining sp^3-hybrids with reduced strengths $t_1 = \sqrt{3}\,t(A_1)/2$ and $t_2 = t(t_2)/2$.

In this simple model, the changes due to the surface are contained in the facts that (a) the s and p orbitals of the antisite interact with an "average" hybrid orbital of its neighbors, having "average" energy $\varepsilon(a_1)$, and (b) the strengths of the interaction for the surface are reduced from those of the bulk. Both effects, in particular (b), can markedly shift the surface antisite levels from those of the bulk.

The results of this simple model are compared with the full surface Green's function calculation in Figs. 1 and 2. For the cation-on-the-anion-site defect (e.g., Ga_{As}), the model yields only one level in or very near the band gap — an acceptor level that can produce Fermi-level pinning and Schottky barrier formation for an n-type semiconductor. The detailed calculations [6] also produce only a single prominent level in the band gap for this defect — again an acceptor level. When the results of the

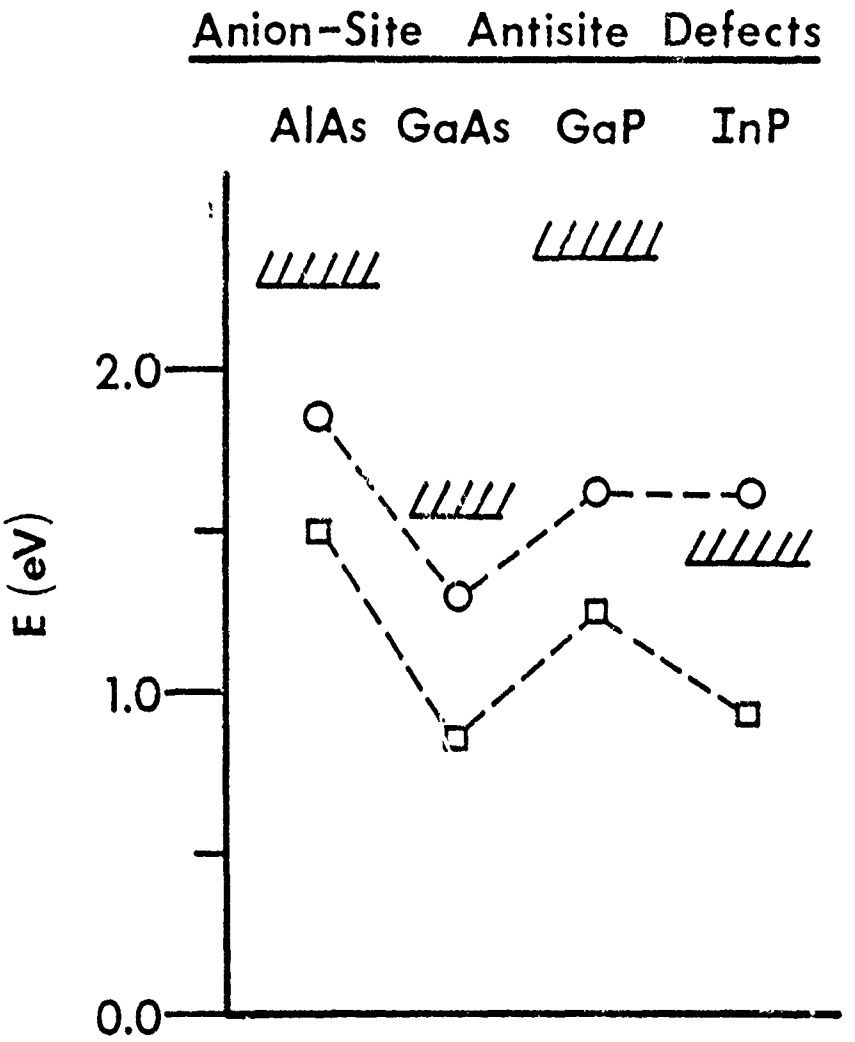


Fig. 1. Predictions of simple four-atom model for acceptor levels associated with cation-on-anion-site antisite defects (open circles) compared with predictions of detailed calculations (open squares) [6].

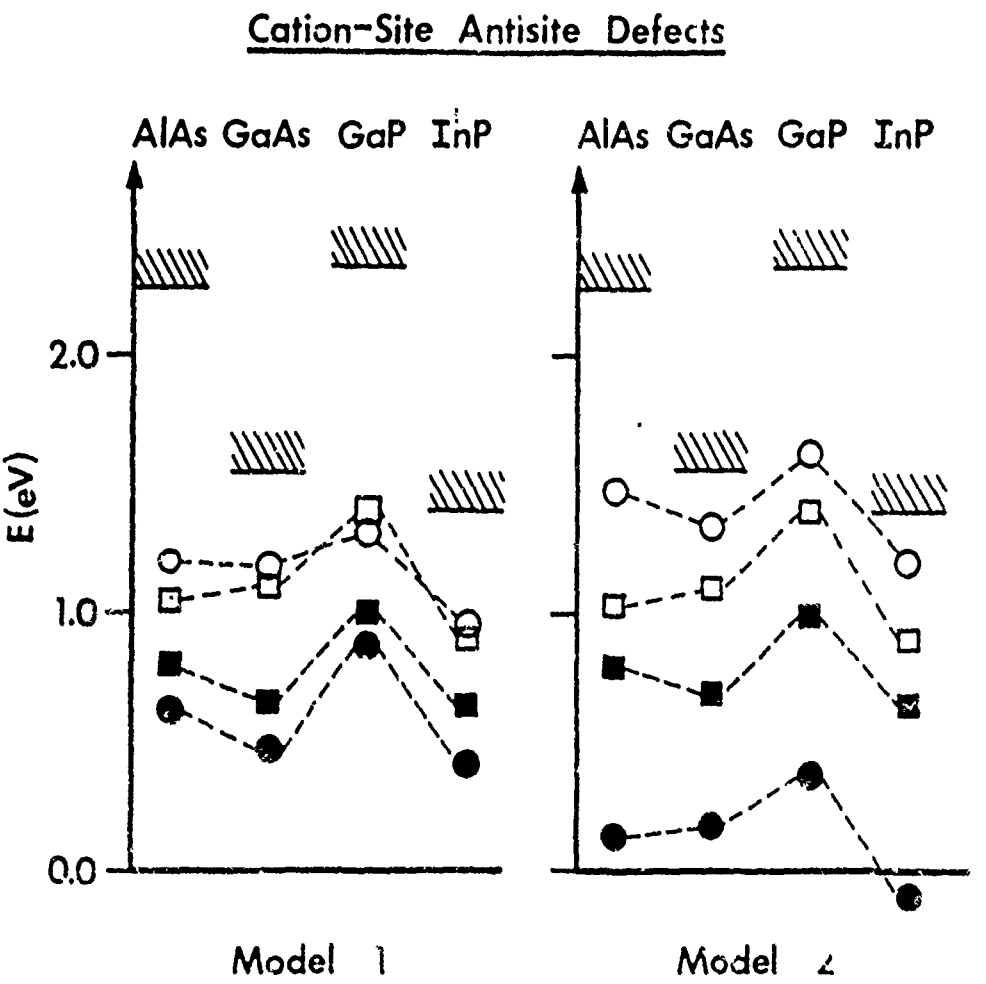


Fig. 2. Predictions of simple four-atom model for acceptor and donor levels associated with anion-on-cation-site antisite defects (circles) compared with predictions of detailed calculations (squares) [6]. Open circles and squares are acceptor levels (empty for neutral defect), and solid circles and squares are donor levels (filled for neutral defect). "Model 2" represents the "exact" calculation for the four-atom model, and "Model 1" represents a calculation in which the indirect coupling between s-orbital and dangling-bond p-orbital on the defect site is neglected. (See text.)

simple model (open circles) and detailed calculations (open squares) are compared in Fig. 1, it can be seen that the chemical trends are in remarkably good agreement.

A similar comparison for the other antisite, the anion-on-the-cation-site defect, is shown in Fig. 2. As described above, two versions of the simple model were used: In "model 1," the 3×3 problem of Eq. (7) was artificially decoupled to yield the two 2×2 problems of Eqs. (4) and (5). This amounts to neglecting the indirect interaction between the defect-site s-orbital and dangling-bond p-orbital via the direct interaction of each of these orbitals with neighboring orbitals on the adjacent anion atoms. (See Eq. (7).) In "model 2," the full 3×3 problem is solved. In both models, one acceptor level and one donor level are produced in (or very near) the band gap. (The "better" model, model 1, gives the "worse" results because of the hybridization of s-orbital and dangling-bond p-orbital on the antisite defect; this is not the relevant point, however.) As can be seen in Fig. 2, either of these versions of the simple 4-atom model (open and solid circles) yields chemical trends almost identical to those of the detailed calculations (open and closed squares).

The agreement between the simple models of both defects and the detailed calculations indicates that both approaches provide a reliable description of the chemical trends. It also indicates that these trends have a simple physical origin, principally involving the dangling-bond p-orbital for the anion-site defect (e.g., Ga_{As}) and both the s-orbital and dangling-bond p-orbital for the cation-site defect (e.g., As_{Ga}).

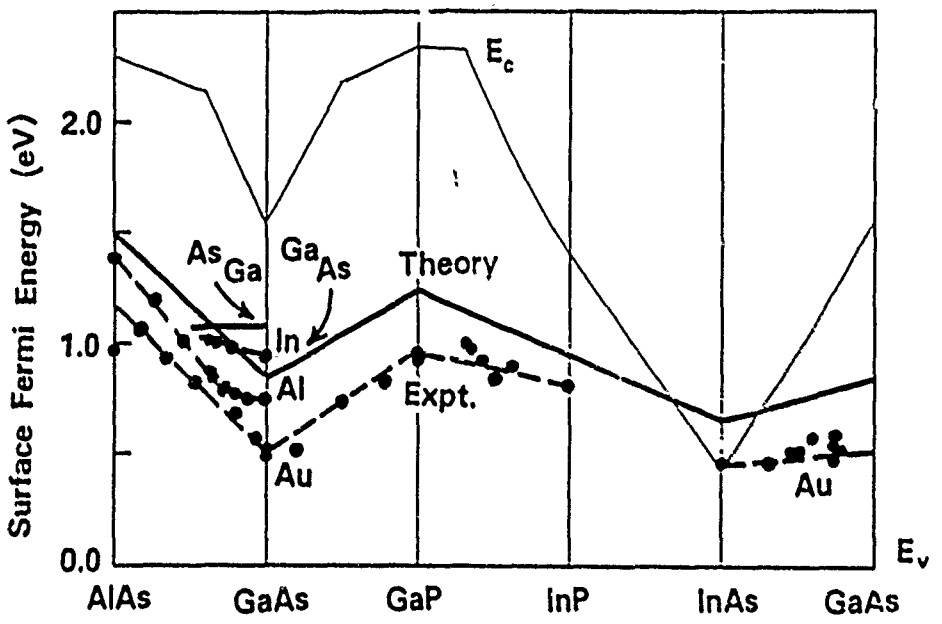


Fig. 3. Predictions of detailed calculations for the cation-on-anion-site defect (e.g., Ga_{As}) and the anion-on-cation-site defect (e.g., As_{Ga}) at relaxed (110) surfaces of III-V semiconductors and their alloys. Only the acceptor levels, relevant to Fermi-level pinnings on n-type semiconductors, are shown. For $Al_xGa_{1-x}As$, note that the slope of the acceptor level (dE/dx) is large for the cation-on-the-anion-site defect and small for the anion-on-the-cation-site defect. The experimental data for Au contacts to various alloys, and Al and In contacts to $Al_xGa_{1-x}As$, are also shown. (References to the experimental papers are given in Refs. [3] and [6]. We attribute the data for Au contacts and for Al contacts to Al-rich $Al_xGa_{1-x}As$, to Fermi-level pinning by the cation-on-the-anion-site antisite defect. The anion-on-cation-site defect is identified as responsible for pinning the Fermi level at In and Al contacts to Ga-rich $Al_xGa_{1-x}As$. These results indicate that the slope of the Fermi-level pinning position as a function of alloy composition (dE/dx) can serve as a signature of the defect type. The conduction band edge is denoted E_c.

In Fig. 3, we show predictions of our detailed calculations for several III-V alloys, compared with experimental Fermi-level pinning positions inferred from measurements of Schottky barriers and MOS (metal-oxygen-semiconductor) structures. (The sources of the experimental data are cited in Ref. [6] and the review of Mönch [3].) For Au contacts to all alloys, the data appear to be well described by the cation-on-anion-site defect level (e.g., Ga_{As}). This defect state is cation dangling-bond-like in character and draws its strength mainly from the conduction band. Hence its energy changes considerably as the alloy composition varies.

The data for In and Al for small x in $Al_xGa_{1-x}As$ however show only a modest change with alloy composition. In fact, Al appears to produce a kink in the Fermi-level pinning position as a function of x. The defect model readily explains this behavior in terms of a "switching" of the dominant defect from the cation-on-anion-site defect (e.g., Ga_{As}) for large x to the anion-on-cation-site defect (e.g., As_{Ga}) for small x. The anion-on-cation-site defect level has anion dangling-bond character, is valence-band-like, and hence shows little change as the alloy composition x varies.

III. SUMMARY

Thus the simple picture of Fermi-level pinning by deep energy levels associated with defects accounts for the chemical trends in the Schottky barrier data. Indeed, the essential physics is contained in the simple four-atom model which, in a hybrid basis, can be easily evaluated.

We thank the Office of Naval Research (Contract Nos. N0014-84-K-0352 and N00014-82-K-0447) for their generous support.

REFERENCES

[1] J. Bardeen, Phys. Rev. 71, 717 (1947).

[2] W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. 16, 1422 (1979); W. E. Spicer, I. Lindau, P. R. Skeath, and C. Y. Su, J. Vac. Sci. Technol. 17, 1019 (1980); W. E. Spicer, I. Lindau, P. R. Skeath, C. Y. Su, and P. W. Chye, Phys. Rev. Letters 44, 520 (1980); and references therein.

[3] W. Mönch, Surf. Sci. 132, 92 (1983) and references therein.

[4] R. H. Williams, Surf. Sci. 132, 122 (1983) and references therein.

[5] H. H. Wieder, Inst. Phys. Conf. Ser. 50, 234 (1980); Appl. Phys. Lett. 38, 170 (1980).

[6] R. E. Allen, T. J. Humphreys, J. D. Dow, and O. F. Sankey, J. Vac. Sci. Technol. B2, 491 (1984). See also R. E. Allen and J. D. Dow, Phys. Rev. B25, 1423 (1982).

[7] O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1984).

[8] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[9] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Letters 44, 810 (1980).

[10] R. E. Allen, Phys. Rev. B 20, 1454 (1979).

[11] S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B26, 951 (1982).

[12] M. Lannoo and P. Lenglart, J. Phys. C 30, 2409 (1969).

Deep impurity levels in semiconductor superlattices

Shang Yuan Ren
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
*and Department of Physics, University of Science and Technology of China, Hefei, China**

John D. Dow and Jun Shen
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 29 February 1988; revised manuscript received 20 October 1988)

A theory of sp^3-bonded substitutional deep impurity levels in periodic $N_1 \times N_2$ GaAs/$Al_xGa_{1-x}As$ superlattices predicts that as the thickness t(GaAs) of each GaAs layer is reduced below a critical value ($\lesssim 17$ Å or $N_1 \lesssim 6$ for $x = 0.7$) common shallow donor impurities such as Si cease donating electrons to the conduction band and instead become deep traps. This happens because the deep levels associated with point defects in either GaAs or $Al_xGa_{1-x}As$ layers (when measured relative to the valence-band maximum of GaAs) are much less sensitive to changes of the alloy composition or layer thicknesses of the superlattice than the superlattice band edges, particularly the conduction-band edge. For some compositions x, dopants such as Si are shallow donors in $N \times N$ GaAs/$Al_xGa_{1-x}As$ superlattices but deep traps in $Al_{x'}Ga_{1-x'}As$ alloy (the alloy obtained by disordering the superlattice). The band gap and band edges of the superlattice, and hence the ionization energies of deep levels, depend strongly on the layer thickness t(GaAs) but only weakly on $t(Al_xGa_{1-x}As)$. The T_2- and A_1-derived deep levels (of the bulk point group T_d) are split and shifted, respectively, near a GaAs/$Al_xGa_{1-x}As$ interface: the p-like T_2 level splits into an a_1 [p_z-like] level, a b_1 [(p_x+p_y)-like] level, and a b_2 [(p_x-p_y)-like] level of the point group for any general superlattice site (C_{2v}), whereas the s-like A_1 bulk level becomes an a_1 (s-like) level of C_{2v}. The order of magnitude of the shifts and splittings of deep levels at a GaAs/$Al_xGa_{1-x}As$ interface is 0.1 eV, depends on x, and becomes very small for impurities more than about three atomic planes away from an interface. Deep levels in the GaAs quantum wells experience level shifts due to (i) penetration of their wave functions into the more electropositive $Al_xGa_{1-x}As$ layers, (ii) the band offset, and (iii) quantum confinement. The cation vacancy, when brought close to a GaAs/$Al_xGa_{1-x}As$ interface, may undergo a shallow-deep transition. These predictions are based on a periodic superslab calculation for unit superslabs with total thickness t(GaAs)+$t(Al_xGa_{1-x}As)$ as large as 102 Å or $N_1+N_2=36$ two-atom-thick layers. The Hamiltonian is a tight-binding model in a hybrid basis that is a generalization of the Vogl model and properly accounts for the nature of interfacial bonds. The deep levels are computed by using the theory of Hjalmarson *et al.* and the special-points method. Our results indicate that some normally shallow donors, such as Si can become deep levels at certain sites in the superlattice as a result of local fluctuations in alloy composition x or layer thickness t(GaAs).

I. INTRODUCTION

Modulation doping of GaAs/$Al_xGa_{1-x}As$ superlattices,[1,2] by which Si impurities are inserted into the large-band-gap $Al_xGa_{1-x}As$ layers of a superlattice but donate their electrons to the small-band-gap GaAs layers, has already played a role in the development of high-speed III-V-compound semiconductive devices. However, practical devices based on $Al_xGa_{1-x}As$ often are limited to alloy compositions $x < 0.3$ because of the inability of Si to dope n type for $x > 0.3$,[3] apparently because of the formation of Si-related centers that are deep traps.[3-11] Furthermore, some devices, such as HEMT's (high-electron-mobility transistors) operate using quantum-well structures[12] at or near GaAs/$Al_xGa_{1-x}As$ interfaces and the performance of these devices depends on the doping (e.g., by Si), the alloy composition x, and the superlattice structure.[13] Clearly, a theory of impurity levels in superlattices and heterostructures is needed to understand the conditions under which a specific impurity produces shallow donor levels and "dopes" the semiconductor versus the conditions for deep-trap formation and the trapping of carriers. That is the purpose of this paper. Das Sarma and Madhukar[14] have previously discussed the deep levels of vacancies in some superlattices, but, to our knowledge, the present work is one of the first systematic studies of the chemical trends of deep impurity levels in superlattices.[15,16] We do know that Hjalmarson,[17] Nelson *et al.*,[18] and Lannoo and Bourgoin[19] are studying deep levels in parallel with our effort, however, although we are not fully aware of the current state of their work. This is also the first treatment of defects in large-layer superlattices. We consider unit supercells typically 40 atomic planes thick. The theory developed here is an extension to superlattices of the theory of Hjalmarson *et al.*[20] of deep impurity levels

 © 1988 The American Physical Society

in the bulk, which has successfully predicted deep levels and their wave functions.[21–23]

II. FORMALISM

A. Host Hamiltonian

We treat a periodic GaAs/$Al_xGa_{1-x}As$ superlattice whose layers are perpendicular to the [001] direction. We employ a nearest-neighbor tight-binding Hamiltonian with an s^*sp^3 basis of five orbitals at each site. Our Hamiltonian, in the limit $x=0$, is identical to the Vogl model for GaAs. Some differences are introduced because of the superlattice, which we treat using a superhelix or supercell method. The superlattice we consider has N_1 two-atom-thick layers of GaAs and N_2 two-atom-thick layers of $Al_xGa_{1-x}As$ repeated periodically; the GaAs and $Al_xGa_{1-x}As$ are assumed to be perfectly lattice matched. We denote this superlattice either as a $(GaAs)_{N_1}/(Al_xGa_{1-x}As)_{N_2}$ superlattice, or as a GaAs/$Al_xGa_{1-x}As$ superlattice with N_1 GaAs layers and N_2 $Al_xGa_{1-x}As$ layers, or as an $N_1 \times N_2$ GaAs/$Al_xGa_{1-x}As$ superlattice.

We first define (for the case $x=1$) a *superhelix* or supercell as a helical string with axis aligned along the [001] or z direction consisting of $2N_1+2N_2$ adjacently bonded atoms As, Ga, As, Ga, As, . . . , Ga, As, Al, As, Al, As, Al, . . . , As, Al. (For $x \neq 1$, replace Al, by Al_xGa_{1-x}.) The center of this helix is at L and each of the atoms of the helix is at position $\mathbf{L}+\mathbf{v}_\beta$ (for $\beta = 0, 1, 2, \ldots, 2N_1+2N_2-1$). A *superslab* of GaAs/$Al_xGa_{1-x}As$ consists of all such helices with the same value of L_z and all possible different values of L_x and L_y, and the superlattice is a stacked array of these superslabs. If the origin of coordinates is taken to be at an As atom, the x and y axes are oriented such that a neighboring cation is at $(\frac{1}{4},\frac{1}{4},\frac{1}{4})a_L$, where a_L is the lattice constant. At each site there are s^*sp^3 basis orbitals $|n,\mathbf{L},\mathbf{v}_\beta\rangle$, where $n=s^*$, s, p_x, p_y, or p_z, and $\beta=0,1,2,\ldots,2N_1+2N_2-1$. From these basis orbitals we form the sp^3 hybrid orbitals at each site $\mathbf{R}=(\mathbf{L},\mathbf{v}_\beta)$.

The hybrid orbitals are

$$|h_1,\mathbf{R}\rangle=[\,|s,\mathbf{R}\rangle+\lambda|p_x,\mathbf{R}\rangle+\lambda|p_y,\mathbf{R}\rangle+\lambda|p_z,\mathbf{R}\rangle]/2\,,$$

$$|h_2,\mathbf{R}\rangle=[\,|s,\mathbf{R}\rangle+\lambda|p_x,\mathbf{R}\rangle-\lambda|p_y,\mathbf{R}\rangle-\lambda|p_z,\mathbf{R}\rangle]/2\,,$$

$$|h_3,\mathbf{R}\rangle=[\,|s,\mathbf{R}\rangle-\lambda|p_x,\mathbf{R}\rangle+\lambda|p_y,\mathbf{R}\rangle-\lambda|p_z,\mathbf{R}\rangle]/2\,,$$

and

$$|h_4,\mathbf{R}\rangle=[\,|s,\mathbf{R}\rangle-\lambda|p_x,\mathbf{R}\rangle-\lambda|p_y,\mathbf{R}\rangle+\lambda|p_z,\mathbf{R}\rangle]/2\,, \tag{1}$$

where $\lambda=+1$ (-1) for atoms at anion (cation) sites. Next, we introduce the label $\nu=s^*$, h_1, h_2, h_3, or h_4, and our hybrid basis orbitals are $|\nu,\mathbf{R}\rangle$. In terms of these orbitals we form the tight-binding orbitals[24]

$$|\nu,\beta,\mathbf{k}\rangle=N_s^{-1/2}\sum_{\mathbf{L}}\exp(i\mathbf{k}\cdot\mathbf{L}+i\mathbf{k}\cdot\mathbf{v}_\beta)|\nu,\mathbf{L},\mathbf{v}_\beta\rangle\,, \tag{2}$$

where $\mathbf{k}$ is (in a reduced-zone scheme) any wave vector of the minizone or (in an extended-zone scheme) any wave vector of the zinc-blende Brillouin zone. Here, N_s is the number of supercells.

The minizone wave vector is a good quantum number, and so the tight-binding Hamiltonian is diagonal in $\mathbf{k}$. Evaluation of the matrix elements $\langle\nu,\beta,\mathbf{k}|H|\nu',\beta',\mathbf{k}\rangle$ leads to a tight-binding Hamiltonian of the block-tridiagonal form. For different β and β', the first three rows of block matrices are

$$\begin{vmatrix} H(0,0) & H(0,1) & 0 & \cdots & \cdots & \cdots & 0 & H(0,2N_1+2N_2-1) \\ H^*(0,1) & H(1,1) & H(1,2) & 0 & 0 & \cdots & 0 & 0 \\ 0 & H^*(1,2) & H(2,2) & H(2,3) & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & & & & \vdots \end{vmatrix}. \tag{3}$$

The last row of blocks is

$$\begin{vmatrix} \vdots & & \vdots & \vdots & & & & \vdots \\ H^*(0,2N_1+2N_2-1) & 0 & 0 & \cdots & 0 & H^*(2N_1+2N_2-2,2N_1+2N_2-1) & H(2N_1+2N_2-1,2N_1+2N_2-1) \end{vmatrix}. \tag{4}$$

Here, $H(\beta,\beta')$ depends on $\mathbf{k}$ and is given in terms of various 5×5 matrices for different ν and ν'.

The diagonal (in β) 5×5 matrix, $H(\beta,\beta)$ at site β, is

$$H(\beta,\beta)=\langle\nu,\beta,\mathbf{k}|H|\nu',\beta,\mathbf{k}\rangle=\begin{vmatrix} \varepsilon_{s^*} & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_h & T & T & T \\ 0 & T & \varepsilon_h & T & T \\ 0 & T & T & \varepsilon_h & T \\ 0 & T & T & T & \varepsilon_h \end{vmatrix}, \tag{5}$$

where

$$\varepsilon_h=(\varepsilon_s+3\varepsilon_p)/4 \tag{6}$$

is the hybrid energy, and

$$T=(\varepsilon_s-\varepsilon_p)/4 \tag{7}$$

is the hybrid-hybrid interaction; the energies ε_{s^*}, ε_h, and T in $H(\beta,\beta)$ refer to the atom at the βth site, and may be obtained from the energies as tabulated by Vogl *et al.*[25] To account for the observed[26] valence-band-edge discon-

tinuity of 32% of the direct band gap,[27] a constant is added to ε_{x^*} and ε_h for $Al_xGa_{1-x}As$; this constant is adjusted to give the valence-band maximum of $Al_xGa_{1-x}As$ below the valence-band maximum of GaAs by 32% of the direct-band-gap difference in the limit $N_1=N_2\to\infty$.

The off-diagonal matrix elements $(\nu,\beta,\mathbf{k}\,|\,H\,|\,\nu',\beta',\mathbf{k})$ or $H(\beta,\beta')$ are best expressed in terms of matrix elements of H between s^*, s, and p orbitals. This is accomplished by the transformation

$$(\nu,\beta,\mathbf{k}\,|\,H\,|\,\nu',\beta',\mathbf{k})=\sum_{n,n'}C(\nu,n;\beta)C(\nu',n';\beta')(n,\beta,\mathbf{k}\,|\,H\,|\,n',\beta',\mathbf{k}), \tag{8}$$

where we have the s^*sp^3 tight-binding functions

$$|\,n,\beta,\mathbf{k})=N_\beta^{-1/2}\sum_{\mathbf{L}}\exp(i\mathbf{k}\cdot\mathbf{L}+i\mathbf{k}\cdot\mathbf{v}_\beta)\,|\,n,\mathbf{L},\mathbf{v}_\beta), \tag{9}$$

and the 5×5 matrices $C(\nu,n;\beta)$ are

$$C(\nu,n;\beta)=\begin{array}{c}\\ s^*\\ h_1\\ h_2\\ h_3\\ h_4\end{array}\begin{array}{c}\begin{array}{ccccc}s^* & s & p_x & p_y & p_z\end{array}\\ \begin{pmatrix}1 & 0 & 0 & 0 & 0\\ 0 & \frac{1}{2} & +\lambda/2 & +\lambda/2 & +\lambda/2\\ 0 & \frac{1}{2} & +\lambda/2 & -\lambda/2 & -\lambda/2\\ 0 & \frac{1}{2} & -\lambda/2 & +\lambda/2 & -\lambda/2\\ 0 & \frac{1}{2} & -\lambda/2 & -\lambda/2 & +\lambda/2\end{pmatrix}\end{array}, \tag{10}$$

where $\lambda=+1$ (-1) if β refers to an anion (cation).

There are several distinct cases for which the off-diagonal (in β) matrix elements $(n,\beta,\mathbf{k}\,|\,H\,|\,n',\beta',\mathbf{k})$ are nonzero (for $\beta\neq\beta'$).

1. Intramaterial matrix elements

If β and β' both refer to nearest-neighbor sites in the same material (either the GaAs or the $Al_xGa_{1-x}As$), we have (assuming β and β' are in material number 1, the GaAs), for example,

$$(n,\beta,\mathbf{k}\,|\,H\,|\,n',\beta',\mathbf{k})=H_{c1a1}, \tag{11}$$

if β refers to a cation and β' refers to an anion. H_{c1a1} is a 5×5 matrix whose rows and columns are labeled by n, which ranges over the values s^*, s, p_x, p_y, and p_z. Similarly, we have matrix elements H_{a1c1}, H_{a2c2}, and H_{c2a2}. These matrix elements are

$$H_{c1a1}=\begin{pmatrix}0 & 0 & -V(s^*c,pa)g^*_{1ca} & -V(s^*c,pa)g^*_{1ca} & -V(s^*c,pa)g^*_{0ca}\\ 0 & V(s,s)g^*_{0ca} & -V(sc,pa)g^*_{1ca} & -V(sc,pa)g^*_{1ca} & -V(sc,pa)g^*_{0ca}\\ V(s^*a,pc)g^*_{1ca} & V(sa,pc)g^*_{1ca} & V(x,x)g^*_{0ca} & V(x,y)g^*_{0ca} & V(x,y)g^*_{1ca}\\ V(s^*a,pc)g^*_{1ca} & V(sa,pc)g^*_{1ca} & V(x,y)g^*_{0ca} & V(x,x)g^*_{0ca} & V(x,y)g^*_{1ca}\\ V(s^*a,pc)g^*_{0ca} & V(sa,pc)g^*_{0ca} & V(x,y)g^*_{1ca} & V(x,y)g^*_{1ca} & V(x,x)g^*_{0ca}\end{pmatrix} \tag{12}$$

and

$$H_{a1c1}=\begin{pmatrix}0 & 0 & V(s^*a,pc)g_{1ac} & -V(s^*a,pc)g_{1ac} & -V(s^*a,pc)g_{0ac}\\ 0 & V(s,s)g_{0ac} & V(sa,pc)g_{1ac} & -V(sa,pc)g_{1ac} & -V(sa,pc)g_{0ac}\\ -V(s^*c,pa)g_{1ac} & -V(sc,pa)g_{1ac} & V(x,x)g_{0ac} & -V(x,y)g_{0ac} & -V(x,y)g_{1ac}\\ V(s^*c,pa)g_{1ac} & V(sc,pa)g_{1ac} & -V(x,y)g_{0ac} & V(x,x)g_{0ac} & V(x,y)g_{1ac}\\ V(s^*c,pa)g_{0ac} & V(sc,pa)g^*_{0ac} & -V(x,y)g^*_{1ac} & V(x,y)g_{1ac} & V(x,x)g_{0ac}\end{pmatrix}. \tag{13}$$

All of the matrix elements V are those tabulated by Vogl *et al.*[25] for material number 1 (viz., GaAs). Identical expressions exist for H_{a2c2} and H_{c2a2}, with $Al_xGa_{1-x}As$ matrix elements. (The $Al_xGa_{1-x}As$ matrix elements are obtained by a virtual-crystal average of the Vogl matrix elements for AlAs and GaAs. x times the corresponding AlAs matrix elements plus $1-x$ times the GaAs matrix elements.)

2. Intermaterial matrix elements

At the interface between GaAs and $Al_xGa_{1-x}As$ there will be nonzero matrix elements of H for each bond between nearest neighbors. These are H_{a2c1} and H_{c2a1}:

$$H_{a2c1}=\begin{vmatrix} 0 & 0 & V(s^*a,pc)g_{1ac} & -V(s^*a,pc)g_{1ac} & -V(s^*a,pc)g_{0ac} \\ 0 & V(s,s)g_{0ac} & V(sa,pc)g_{1ac} & -V(sa,pc)g_{1ac} & -V(sa,pc)g_{0ac} \\ -V(s^*c,pa)g_{1ac} & -V(sc,pa)g_{1ac} & V(x,x)g_{0ac} & -V(x,y)g_{0ac} & -V(x,y)g_{1ac} \\ V(s^*c,pa)g_{1ac} & V(sc,pa)g_{1ac} & -V(x,y)g_{0ac} & V(x,x)g_{0ac} & V(x,y)g_{1ac} \\ V(s^*c,pa)g_{0ac} & V(sc,pa)g_{0ac} & -V(x,y)g_{1ac} & V(x,y)g_{1ac} & V(x,x)g_{0ac} \end{vmatrix} \tag{14}$$

and

$$H_{c2a1}=\begin{vmatrix} 0 & 0 & -V(s^*c,pa)g^*_{1ac} & V(s^*c,pa)g^*_{1ac} & V(s^*c,pa)g^*_{0ac} \\ 0 & V(s,s)g^*_{0ac} & -V(sc,pa)g^*_{1ac} & V(sc,pa)g^*_{1ac} & V(sc,pa)g^*_{0ac} \\ V(s^*a,pc)g^*_{1ac} & V(sa,pc)g^*_{1ac} & V(x,x)g^*_{0ac} & -V(x,y)g^*_{0ac} & -V(x,y)g^*_{1ac} \\ -V(s^*a,pc)g^*_{1ac} & -V(sa,pc)g^*_{1ac} & -V(x,y)g^*_{0ac} & V(x,x)g^*_{0ac} & V(x,y)g^*_{1ac} \\ -V(s^*a,pc)g^*_{0ac} & -V(sa,pc)g^*_{0ac} & -V(x,y)g^*_{1ac} & V(x,y)g^*_{1ac} & V(x,x)g^*_{0ac} \end{vmatrix} . \tag{15}$$

Here the Vogl matrix elements are those for the bond in question: If the cation is Al_xGa_{1-x} and the anion is As, then the matrix element is the $Al_xGa_{1-x}As$ matrix element obtained by a virtual-crystal average of the AlAs and GaAs values. We also have

$$g_{0ca}=\exp(i\mathbf{k}\cdot\mathbf{d}_1)+\exp(i\mathbf{k}\cdot\mathbf{d}_4)\,,$$

$$g_{1ca}=\exp(i\mathbf{k}\cdot\mathbf{d}_1)-\exp(i\mathbf{k}\cdot\mathbf{d}_4)\,,$$

$$g_{0ac}=\exp(i\mathbf{k}\cdot\mathbf{d}_2)+\exp(i\mathbf{k}\cdot\mathbf{d}_3)\,,$$

and

$$g_{1ac}=\exp(i\mathbf{k}\cdot\mathbf{d}_2)-\exp(i\mathbf{k}\cdot\mathbf{d}_3)\,, \tag{16}$$

where we have $4\mathbf{d}_i/a_L=(1,1,1)$, $(1,-1,-1)$, and $(-1,1,-1)$, and $(-1,-1,1)$ for $i=1$, 2, 3, and 4, respectively. Here, a_L is the room-temperature lattice constant of GaAs, 5.653 Å, which we assume is equal to that of $Al_xGa_{1-x}As$.[28]

In this work we study deep impurity levels in superhelices as large as $N_1+N_2=20$; that is, in 40-atom-thick superslabs.[29] The dimension of the Hamiltonian matrix at each value of k is $5(2N_1+2N_2)$, because there are five orbitals per site. We diagonalize this Hamiltonian numerically for each (special point) k, finding its eigenvalues $E_{\gamma,\mathbf{k}}$ and the projections of the eigenvectors $|\gamma,\mathbf{k})$ on the $|\nu,\beta,\mathbf{k})$ hybrid basis: $(\nu,\beta,\mathbf{k}|\gamma,\mathbf{k})$. Here, γ is the band index (and ranges from 1 to 200 for $N_1=N_2=10$) and k lies within the mini-Brillouin zone in a reduced-zone scheme or within the GaAs Brillouin zone in an extended-zone scheme. (We assume that GaAs and $Al_xGa_{1-x}As$ are perfectly lattice-matched.[28])

B. Deep levels

The theory of deep levels is based on the Green's-function theory of Hjalmarson *et al.*,[20] which solves the secular equation for the deep-level energy E,

$$\det[1-G(E)V]$$

$$=0=\det\left[1-\mathrm{P}\int\left[\delta(E'-H)\frac{V}{E-E'}\right]dE'\right]\,, \tag{17}$$

Here, V is the defect potential matrix,[20] which is zero except at the defect site and diagonal on that site, $(0,V_s,V_p,V_p,V_p)$, in the Vogl s^*sp^3 local basis centered on each atom. We also have $G=(E-H)^{-1}$, where H is the host tight-binding Hamiltonian operator. The spectral density operator is $\delta(E'-H)$ and P denotes the principal-value integral over all energies. In the fundamental band gap of the superlattice, G is real.

1. Point-group analysis

A substitutional point defect in either bulk GaAs or bulk $Al_xGa_{1-x}As$ has tetrahedral (T_d) point-group symmetry (assuming a virtual-crystal approximation for $Al_xGa_{1-x}As$). Each such sp^3-bonded defect normally has one s-like (A_1) and one triply-degenerate p-like (T_2) deep defect level near the fundamental band gap. If we imagine breaking the symmetry of bulk GaAs by making it into a GaAs/GaAs superlattice along the [001] direction, we reduce the T_d symmetry to D_{2d}. If, in addition, we change alternating slabs of GaAs to virtual-crystal $Al_xGa_{1-x}As$, forming a GaAs/$Al_xGa_{1-x}As$ superlattice, then the point-group symmetry of a general substitutional defect is C_{2v}. (For selected sites, e.g., at the center of a material, the symmetry can be higher.) Note that the C_{2v} symmetry we find differs from that of Ref. 14. In the GaAs/$Al_xGa_{1-x}As$ superlattice the A_1 and T_2 deep levels of the bulk GaAs or $Al_xGa_{1-x}As$ produce two a_1 levels (one s-like, derived from the A_1 level, and one T_2-derived p_z-like), one b_1 level [p_x-p_y-like], and one b_2

level [(p_x-p_y)-like]. Of course, for impurities far from a GaAs/$Al_xGa_{1-x}As$ interface, the s-like a_1 level will have an energy very close to the energy of a bulk A_1 level, and the p_z-like a_1 level and the b_2 and b_2 levels will lie close to the bulk T_2 level also. We normally expect to find the T_2-derived a_1 level between the b_1 and b_2 levels, but if the level lies close to the valence-band maximum, then the splitting of the valence-band edge into a (p_x+p_y)- and (p_x-p_y)-like maximum with a p_z-like edge at slightly lower energy (because of the smaller effective mass) may cause the a_1 deep level to lie lower in energy than the b_1 and b_2 levels by a comparable energy [see Eq. (17)]. Note that this splitting exists even for defects distant from the interface and is a consequence of the different host spectral densities in the superlattice for a_1 and b_1 and b_2 states.

2. Secular equations

The secular equation, Eq. (12), is reduced by symmetry to the following three equations:

$$G(b_1;E)=V_p^{-1} \quad \text{for } b_1 \text{ levels}, \tag{18}$$

$$G(b_2;E)=V_p^{-1} \quad \text{for } b_2 \text{ levels}, \tag{19}$$

and

$$\det\begin{vmatrix} G(s,s;E)V_s-1 & G(s,z;E)V_p \\ G(z,s;E)V_s & G(z,z;E)V_p-1 \end{vmatrix}=0 , \tag{20}$$

for a_1 levels, where we have

$$G(b_1;E)=\sum_{\gamma,\mathbf{k}} \frac{|\langle h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle-\langle h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle|^2}{2(E-E_{\gamma,\mathbf{k}})} , \tag{21}$$

$$G(b_2;E)=\sum_{\gamma,\mathbf{k}} \frac{|\langle h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle-\langle h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle|^2}{2(E-E_{\gamma,\mathbf{k}})} , \tag{22}$$

$$G(s,s;E)=\sum_{\gamma,\mathbf{k}} \frac{|\langle h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle|^2}{4(E-E_{\gamma,\mathbf{k}})} , \tag{23}$$

$$G(z,z;E)=\sum_{\gamma,\mathbf{k}} \frac{|\langle h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle-\langle h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle-\langle h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle|^2}{4(E-E_{\gamma,\mathbf{k}})} , \tag{24}$$

and

$$G(s,z;E)=\sum_{\gamma,\mathbf{k}} [\langle h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle] \times[\langle h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle-\langle h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle-\langle h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle+\langle h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle]^*[4(E-E_{\gamma,\mathbf{k}})]^{-1} . \tag{25}$$

Here, $G(z,s;E)$ is the Hermitian conjugate of $G(s,z;E)$ and β is the site of the defect in the superlattice.

For each β the relevant host Green's functions, Eqs. (21)–(25), are evaluated using the special-points method,[30] and the secular equations (18)–(20) are solved, yielding $E(b_1;V_p)$, $E(b_2;V_p)$, and two values of $E(a_1;V_s,V_p)$. The defect-potential matrix elements V_s and V_p are obtained using a slight modification of Hjalmarson's approach.[31,32] For $N_1=N_2=10$, there are 40 possible sites β, each with four relevant deep levels: two a_1, one b_1, and one b_2; thus there are 160 levels.

3. Special points

For our studies of deep levels in the band gaps of superlattices, we consider only superlattices such that $(N_1+N_2)/4$ is an integer. In such cases the sums over k in Eqs. (21)–(25) can be performed using 12 special points $\mathbf{k}=(\pi/4a_L)\mathbf{u}$, where we have the value u and the weight w of each special point (u;w):[30] (7,3,1;$\frac{1}{2}$), (5,3,1;$\frac{1}{2}$), (7,1,1;$\frac{1}{2}$), (3,1,1;$\frac{1}{2}$), (3,3,1;$\frac{1}{16}$), (1,1,1;$\frac{1}{16}$), (−7,3,−1;$\frac{1}{2}$), (−5,3,−1;$\frac{1}{16}$), (−7,1,−1;$\frac{1}{16}$), (−3,1,−1;$\frac{1}{2}$), (−3,3,−1;$\frac{1}{16}$), and (−1,1,−1;$\frac{1}{16}$). For $(N_1+N_2)/4$ an integer and a GaAs/GaAs superlattice, either the first six or the last six special points would be sufficient to give the Green's function at any site in the superlattice with the same accuracy as Chadi's and Cohen's[33] ten special points for bulk GaAs, which are known to give an adequately accurate bulk Green's function. However, for the GaAs/$Al_xGa_{1-x}As$ superlattice, 12 special points are necessary because of the reduction of symmetry from D_{2d} to C_{2v} from GaAs/GaAs to GaAs/$Al_xGa_{1-x}As$.

III. RESULTS

A. Host band gap

Our calculations produce $E_{\gamma,\mathbf{k}}$, the superlattice band structure, including the band gap, which exhibits a particularly interesting behavior as the thickness of the GaAs slabs, $t(\text{GaAs})=N_1a_1/2$ (where $a_1=5.653$ Å is the lattice constant of GaAs), or the number of GaAs layers, N_1, becomes small in comparison with the thickness, $t(Al_xGa_{1-x}As)=N_2a_2/2$, or the number of layers, N_2,

of $Al_xGa_{1-x}As$ slabs: The small-band-gap GaAs layers become quantum wells surrounded by large-band-gap $Al_xGa_{1-x}As$ (Fig. 1). As a result, the band gap of the superlattice increases from the GaAs band gap toward the $Al_xGa_{1-x}As$ band gap as N_1 decreases (for N_2 large). (Qualitatively similar results for smaller $N_1 \div N_2$ superlattices have been reported by Schulman and McGill for GaAs/AlAs superlattices.[24]) This is demonstrated by the calculated results of Figs. 2 and 3. The band gap was taken to be the smallest gap found at one of the following $k=(2\pi/a_L)S$ points of the mini-Brillouin-zone: $S=(0,0,0)$, $(0,0,\tau)$, $(\frac{1}{2},\frac{1}{2},\tau)$, $(\frac{1}{2},\frac{1}{2},0)$, $(1,0,\tau)$, and $(1,0,0)$, where we have $\tau=(N_1+N_2)^{-1}$. For N_1 large, the gap was invariably at $k=\Gamma=(0,0,0)$, but for small N_1 and large x (somewhat larger than 0.3) it was sometimes found at $S=(\frac{1}{2},\frac{1}{2},\tau)$ or $(\frac{1}{2},\frac{1}{2},0)$. For example, in Fig. 2 we see the results for $N_1 \times N_2$ GaAs/$Al_{0.7}Ga_{0.3}As$ superlattices: the valence-band maximum is at Γ and the conduction-band minimum for thick GaAs layers ($N_1 \geq 8$) is also at Γ. However, for thinner layers ($2 \leq N_1 < 8$) the wave vector of the superlattice conduction-band minimum is at $(2\pi/a_L)(\frac{1}{2},\frac{1}{2},0)$ and the superlattice states associated with this minimum are largely derived from GaAs conduction-band states from near the L point, $(2\pi/a_L)(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, of the bulk Brillouin zone. This is the case because the L point in the bulk, $(2\pi/a_L)(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, is the sum of $(2\pi/a_L)(\frac{1}{2},\frac{1}{2},0)$ and $(2\pi/a_L)(0,0,\frac{1}{2})$, and the bulk point $(2\pi/a_L)(0,0,\frac{1}{2})$ corresponds to $\Gamma=(0,0,0)$ in the superlattice Brillouin zone for N_1+N_2 even. In the thinnest superlattices, $N_1 \simeq 1$, the $Al_xGa_{1-x}As$ X point, $(2\pi/a_L)(0,0,1)$, is reflected in the superlattice conduction-band minimum. These conduction-band minima away from the Γ point of the superlattice Brillouin zone have severe consequences for the optical properties of small-period GaAs/$Al_{0.7}Ga_{0.3}As$ superlattices: the superlattices will be indirect-band-gap materials, and hence will be poor candidates for light-emitting devices. For $N_1=1$ in Fig. 2, the band gap is direct and the superlattice conduction-band minimum is at Γ, although this minimum is derived from the X point of the $Al_{0.7}Ga_{0.3}As$ band structure, and so one should expect the direct transitions associated with it to be weaker than those associated with the GaAs Γ minimum for $N_1 > 8$, for example.

The band gap is somewhat more sensitive to changes of the GaAs layer thickness than to changes in the $Al_xGa_{1-x}As$ layer thickness, as demonstrated in Fig. 3. This sensitivity of the superlattice band gap to the GaAs layer thickness is important for the physics of deep levels in superlattices because, as we shall see below, the deep levels have energies relative to the GaAs valence-band maximum that vary relatively little with the thickness of the GaAs layers. Hence a deep level that is near the conduction-band edge but within the gap in a GaAs quantum well can be "covered up" and autoionized by the conduction-band edge when the GaAs layer thickness is increased and the conduction-band edge descends in energy, while the deep level remains at a relatively constant energy (Fig. 4). Here it is important to remember that we use the new definition of a "deep" level[20] as one that originates from the perturbation caused by the central-cell potential. (This contrasts with the old definition as a level that lies within the band gap by at least 0.1 eV.) As a result there are "deep resonances" that lie in the conduction band above the band gap. We shall see below that Si in GaAs, although producing only shallow levels in the band gap of bulk GaAs (i.e., its deep levels are all resonances that lie in the conduction band) is a candidate for producing a deep level in the band gap of a GaAs/$Al_xGa_{1-x}As$ superlattice, in the GaAs quantum-well limit. For neutral Si this level, when in the

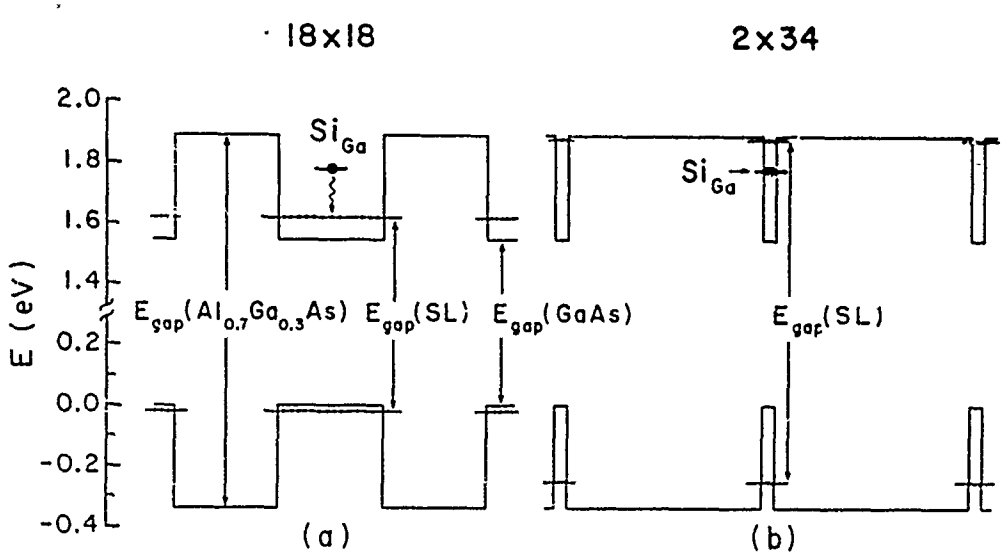


FIG. 1. Illustrating the quantum-well effect on the band gap $E_{gap}(SL)$ of an $N_1 \times N_2$ GaAs/$Al_xGa_{1-x}As$ superlattice: (a) $N_1=N_2=18$ and (b) $N_1=2$, $N_2=34$. The band edges of the superlattice are denoted by chained lines. For this alloy composition the superlattice gap is indirect for case (b), with the conduction-band edge at $k=2\pi/a_L(\frac{1}{2},\frac{1}{2},0)$. Note the broken energy scale. The zero of energy is the valence-band maximum of GaAs.

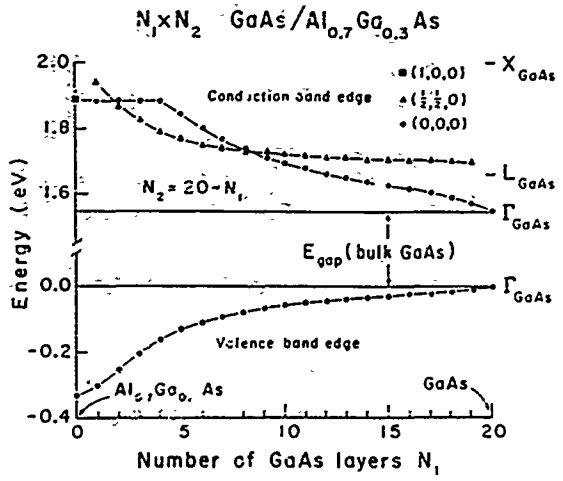


FIG. 2. Predicted energies (in eV) of the superlattice conduction-band minima and valence-band maximum with respect to the valence-band maximum *of bulk GaAs* for a GaAs/$Al_xGa_{1-x}As$ [001] superlattice vs reduced layer thicknesses N_1 and N_2 for various $N_1 \times N_2$ [001] GaAs/$Al_xGa_{1-x}As$ superlattices, with $x = 0.7$ and $N_1 + N_2$ fixed to be 20. The calculations are based on the low-temperature band structures of GaAs and $Al_{0.7}Ga_{0.3}As$, with bulk band gaps of 1.55 and 2.22 eV, respectively. The conduction-band minimum of the superlattice is at $(2\pi/a_L)(\frac{1}{2},\frac{1}{2},0)$ for the triangular points, at k = 0 for the circles, and at $(2\pi/a_L)(1,0,0)$ for the rectangle. The superlattice valence-band maximum is at k = 0. Note the broken scale on the ordinate. The positions of the band extrema of bulk GaAs at Γ, L, and X are shown on the right of the figure, at $N_1 = 20$.

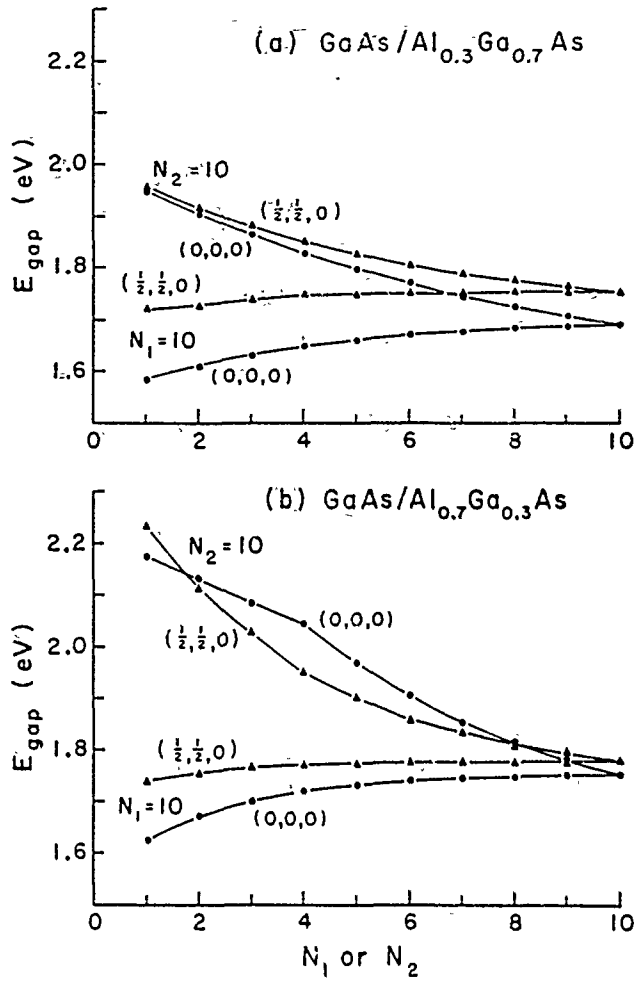


FIG. 3. Predicted fundamental energy band gaps E_{gap} at k = 0 (circles) and $k = (2\pi/a_L)(\frac{1}{2},\frac{1}{2},0)$ (triangles) of a $(GaAs)_{N_1}(Al_xGa_{1-x}As)_{N_2}$ superlattice as functions of reduced GaAs layer thickness N_1 or $Al_xGa_{1-x}As$ layer thickness N_2 for (a) $x = 0.3$ and (b) $x = 0.7$. Note that the variation of the gap with decreasing N_2 from, say, 8 to 4, is less than the variation associated with changing N_1 from 8 to 4. Note also that the k = 0 conduction-band extremum of the superlattice in (b) is derived from the X point of the $Al_{0.7}Ga_{0.3}As$ band structure for $N_1 < 4$ and from the Γ point of the GaAs band structure for $N_1 > 4$.

fundamental band gap, will be occupied by one electron. When the GaAs layer thickness increases, this level is covered up by the falling conduction-band edge and becomes a resonance. The electron in the resonant level is autoionized and relaxes (e.g., by phonon emission) to the conduction-band edge where it is a shallow donor electron, donated by the Si. (In the present theory, which neglects the long-ranged Coulomb potential of the donor, the binding energy of a shallow donor is zero; in a more complete theory, the Coulomb potential would trap this electron at zero temperature in a hydrogenic orbit.) A gratifying feature of the band-gap calculation is that we obtain for $N_1 = N_2 = 1$ a fundamental gap of 2.11 eV for a GaAs/AlAs superlattice, in good agreement with the measured value of Ref. 35.

B. Defect levels

The substitutional defect energy levels for sp^3-bonded impurities can be evaluated using the techniques of Hjalmarson *et al.*,[20] as described above for superlattices. When interpreting the predictions, one should remember that the absolute energy levels predicted by this theory have a theoretical uncertainty of a few tenths of an eV. This is, of course, comparable with the uncertainties of the other sophisticated theories of deep levels that have been presented to date.[36] Nevertheless, the theoretical uncertainty is a significant fraction of the band-gap energy, and so one must not use the theory in a futile attempt to predict absolute energy levels with high precision. Rather, the theory should be employed to understand the chemical trends in the deep energy levels, to study qualitative changes in level structures (such as a deep resonance descending into the fundamental band gap—the shallow-deep transition), or to suggest experiments for testing hypotheses about impurity states. One of the reasons that the Hjalmarson model[20] has been so successful is that the tight-binding Hamiltonian[25] has been constructed with manifest chemical trends in its parameters, following ideas developed originally by Harrison.[37] Earlier theories sometimes obtained tight-binding parameters by performing least-squares fits to the band structures of the semiconductors being studied. Such fits, while having

given impressive band structures, often lacked the essential chemistry that determines deep levels, and, as a result, those theories have not been as successful as the Hjalmarson theory. Indeed, because the Hamiltonian employed in the Hjalmarson theory has manifest chemical trends and also has (by construction) the correct band gaps, the Hjalmarson method is comparably accurate with far more cumbersome pseudopotential theories of deep levels.[36,38]

1. Dependence on layer thickness

Figure 4 displays the dependence on GaAs reduced layer thickness N_1 of the deep Ga-site A_1 level of a Si impurity[31] *in the middle of a GaAs layer* in a $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice. As the size, N_1, of the GaAs layer shrinks, the *deep levels remain relatively constant in energy* with respect to the GaAs valence-band maximum, while the conduction-band edge of the superlattice increases in energy—progressively uncovering the once-resonant deep level of Si and converting this shallow donor impurity into a deep trap.[39] This shallow-deep transition as a function of GaAs well size N_1, to our knowledge, has not been anticipated in the literature—and has consequences for $GaAs/Al_xGa_{1-x}As$ superlattice and quantum-well devices, because it implies that the most common dopant, Si, may become a deep electron trap rather than a shallow donor in GaAs.

Also note (Fig. 4) that when Si in a GaAs quantum well becomes a deep impurity with its deep level in the fundamental gap of the superlattice, this level (with respect to the GaAs valence-band maximum) generally lies at a higher energy than the bulk GaAs band gap, at lower energy than the Si deep level in an $Al_xGa_{1-x}As$ layer, and below the superlattice and bulk $Al_xGa_{1-x}As$ alloy conduction-band edges. Because Si in an $Al_xGa_{1-x}As$ layer lies at higher energy than Si in a GaAs layer, it is possible to move the conduction-band edge up by reducing the width of the GaAs layers and to achieve a situation such that Si in a GaAs layer is a deep level, but that Si in an $Al_xGa_{1-x}As$ layer is a shallow donor with respect to the superlattice (but not with respect to bulk $Al_xGa_{1-x}As$) because its deep level lies above the superlattice conduction-band edge, but below the bulk $Al_xGa_{1-x}As$ conduction-band edge. In the more common case for very thin GaAs layers, Si in both GaAs and $Al_xGa_{1-x}As$ layers will produce deep levels below the superlattice conduction-band edge, that is, since the superlattice band edge lies below the $Al_xGa_{1-x}As$ band edge, for Si to be deep in a thin GaAs layer requires Si to be a deep level in bulk $Al_xGa_{1-x}As$ as well. Because of this requirement, and the fact that Si successfully modulation-dopes $GaAs/Al_xGa_{1-x}As$ superlattices for $x < 0.3$ (suggesting that Si is a shallow donor in, for example, $Al_{0.2}Ga_{0.8}As$), we doubt that the shallow-deep transition as a function of N_1 will be observable in $GaAs/Al_xGa_{1-x}As$ for $x \lesssim 0.2$. For $x > 0.3$, however, Si in bulk $Al_xGa_{1-x}As$ is almost certainly a deep level,[8] modulation doping with Si in $GaAs/Al_xGa_{1-x}As$ superlattices should require thermal activation or tunneling (presumably because Si in an $Al_xGa_{1-x}As$ layer is a deep impurity, not a shallow donor), and the predicted shallow-deep transition for Si in a GaAs layer as a function of GaAs layer thickness should occur.

Because of uncertainties in the theory, we cannot estimate with precision the layer thickness N_1 at which the Si level in GaAs should undergo the shallow-deep transition. Based on the general structure of the curves of Fig. 4, it probably occurs for $N_1 \sim 6$ and a GaAs layer thickness of order $\simeq 17$ Å or less.

A similar analysis can be made of the behavior of other

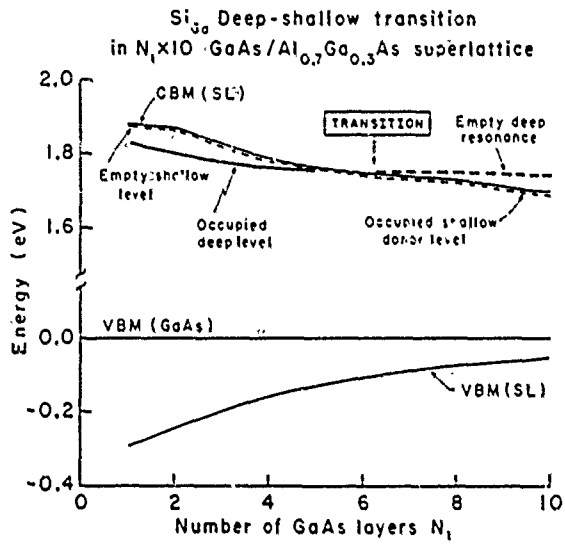


FIG. 4. Illustrating the deep-to-shallow transition as a function of GaAs layer thickness N_1 in a $GaAs/Al_xGa_{1-x}As$ $N_1 \times 10$ superlattice (SL) with $x = 0.7$ for a Si impurity on a column-III site in the center of a GaAs layer of the superlattice host. The conduction-band minimum (CBM) and valence-band maximum (VBM) are indicated by light solid lines. The Si deep level is denoted by a heavy line, which is solid when the level is in the gap but dashed when the level is resonant with the conduction band. The deep level in the band gap for $N_1 < 6$ is covered up by the conduction band as a result of changes in the host for $N_1 > 6$. The impurity's deep level lies in the gap for $N_1 < 6$ and is occupied by the extra Si electron; the Si, in this case, is thus a "deep impurity." For $N_1 > 6$ the deep level lies above the conduction-band edge as a resonance. The daughter electron from the Si impurity which was destined for this deep level is autoionized, spills out of the deep resonance level, and falls to the conduction-band edge (light solid line) where it is subsequently bound (at low temperature) in a shallow level associated with the long-ranged Coulomb potential of the donor (indicated by the short-dashed line). It is important to realize that *both* the deep level and the shallow levels coexist and are distinct levels with qualitatively different wave functions. The issue of whether an *impurity* is "deep" or "shallow" is determined by whether or not a deep level associated with the impurity lies in the band gap. The computed deep-shallow transition occurs for $N_1 \sim 6$ layers. While the qualitative physics is completely reliable, it is possible that the transition layer thickness may differ somewhat from $N_1 = 6$ in real superlattices. The predicted fundamental band gap of the superlattice is indirect for $1 < N_1 < 8$. All energies are with respect to the valence-band maximum *of GaAs*.

impurities,[16] for example, N in GaAs. Nitrogen has a resonant state in the conduction band of GaAs. We have predicted that this resonant state can be driven out of the conduction band into the band gap by hydrostatic pressure,[40] an effect that has been observed by Wolford *et al.*[41] and Zhao *et al.*[42] Here we predict that N in GaAs layers of a $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice will have an impurity state in the band gap if the thickness of the GaAs layers is thin enough, for example, ≤ 48 Å or 17 molecular layers.

2. *As vacancy levels*

Figure 5 displays the deep energy levels in the band gap of an As vacancy in an $N_1=N_2=10$ $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice, as a function of β, the z component of the position of the vacancy in the superhelix or superslab. A vacancy is simulated here by letting the defect-potential elements V_s and V_p at the vacancy site approach infinity, making the defect into an "atom" with infinite orbital energies, and thereby decoupling it from the host by virtue of the fact that all energy denominators in a perturbation expansion are infinite.[43] Several features of the results in Fig. 5 are worth noting: The valence-band maximum of the superlattice splits, with the p_z-like edge moving to lower energy, owing to its lighter effective mass. Near an interface ($\beta=0$, 20, or 40) the p-like T_2 bulk As vacancy level splits into three levels, a_1, b_1, and b_2. For any point defect, one of the b_1 or b_2 levels corresponds to a p-like level with its orbital composed of hybrids directed toward the GaAs layer and has an energy almost the same as the bulk GaAs T_2 level; the other is composed of hybrids directed toward the $Al_xGa_{1-x}As$ and is virtually an $Al_xGa_{1-x}As$ T_2 level. The b_1 and b_2 levels reverse ordering from $\beta=0$ to $\beta=20$ (because of the defined orientation of the x and y axes), and a_1 lies between them in most cases such that the splitting of the host valence-band edge can be neglected. The splittings between the b_1 and b_2 levels at the interface are small, of order 0.1 eV, and become negligible when the vacancy is more than three or so atomic layers from the interface (a fact noted first for deep levels near surfaces by Daw and Smith[44]). The splitting between the a_1 level and the b_1 and b_2 levels is comparably small, but may not vanish even if the defect is distant from the interface, as a result of the splitting of the valence-band edge in the superlattice and the resulting changes of the host spectral density.

The energy of a deep level is determined by a balance between the conduction-band states, which push the level down in energy, and the valence-band states, which repel it upward. Since the GaAs valence-band maximum is almost at the same energy as the valence-band maximum for the 10×10 $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice, the conduction band has the primary influence on the change of the energies of the As-vacancy deep level from bulk

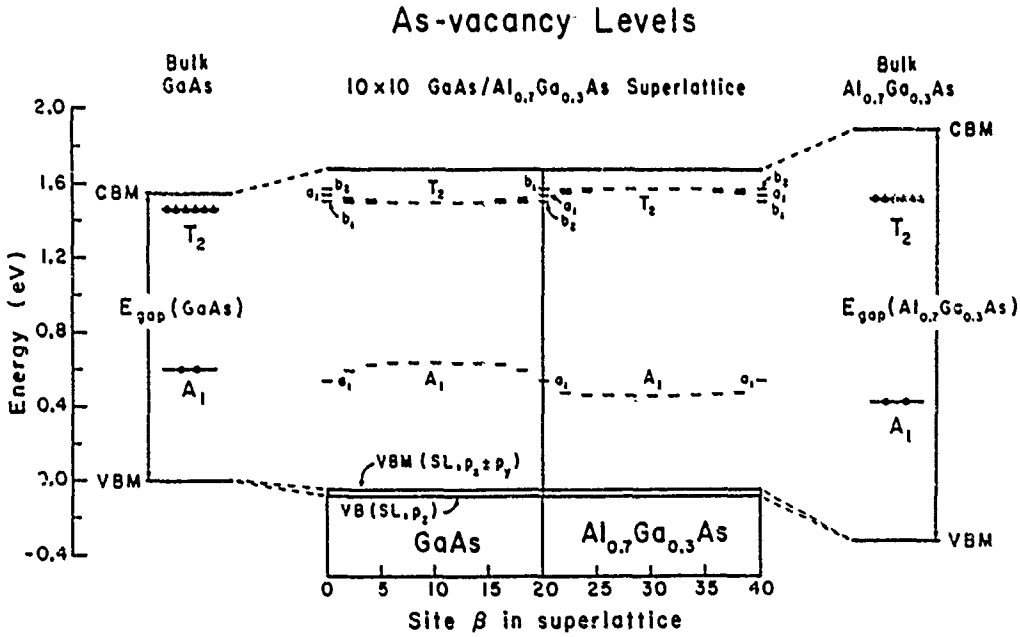


FIG. 5. Predicted energy levels of an As vacancy in a $(GaAs)_{10}/(Al_{0.7}Ga_{0.3}As)_{10}$ superlattice, as a function of β, the position of the vacancy (even values of β correspond to As sites). Note the splitting of the T_2 levels at and near the interfaces ($\beta=0$, 20, and 40), and that the T_2-derived vacancy levels lie at higher energy in an $Al_{0.7}Ga_{0.3}As$ layer than in a GaAs layer. The b_1 and b_2 ordering changes at successive interfaces. The zero of energy is the valence-band maximum of bulk GaAs, and the corresponding valence band (VBM) and conduction-band (CBM) edges and deep levels in bulk GaAs and bulk $Al_{0.7}Ga_{0.3}As$ are given to the left and right of the central figure, respectively. The top of the central figure is the conduction-band edge of the superlattice, and the bottom corresponds to the split valence band in the superlattice—the valence-band maximum of the superlattice being of b_1 or b_2 symmetry ($p_x \mp p_y$) and the split-off (a_1, p_z) band maximum lying 0.062 eV lower in energy. The A_1 level in the $Al_xGa_{1-x}As$ layer of the superlattice is lower than the corresponding level in the GaAs layer because of the band offset of 0.334 eV. The electron (hole) occupancies of the deep levels in bulk GaAs and bulk $Al_xGa_{1-x}As$ are denoted by solid circles (open triangles).

GaAs to a GaAs layer in the superlattice, and the levels move up in energy because the conduction-band minimum does. In going from an $Al_{0.7}Ga_{0.3}As$ layer in the superlattice to bulk $Al_{0.7}Ga_{0.3}As$, the levels move down in energy, because the valence-band maximum moves down. (The valence-band maximum is T_2-like.) The T_2 As vacancy levels lie at somewhat higher energy in an $Al_xGa_{1-x}As$ layer than in a GaAs layer, owing to the fact that Al is more electropositive than Ga. The A_1 bulk As-vacancy level in $Al_xGa_{1-x}As$ is slightly shifted at the interface, and is at lower energy in an $Al_xGa_{1-x}As$ layer than in a GaAs layer because the band offset causes the nearest A_1 states in the valence band to lie at lower energy than in the superlattice. The T_2 level in GaAs ($Al_xGa_{1-x}As$) is both split in the superlattice and, on the average, is shifted up (down) in energy at the interface. Because of the band offset, the A_1 level for the As vacancy in GaAs ($Al_xGa_{1-x}As$) moves down (up) at the interface. Still another effect does influence the relative ordering of the a_1, b_1, and b_2 levels: As the impurity approaches the interface in GaAs (but not in $Al_xGa_{1-x}As$) its wave function suffers quantum confinement and its energy levels shift—and the shifts for b_1 and b_2 levels are generally different from those for a_1 levels because the a_1 valence-band edge in the 10×10 superlattice lies lower by 0.062 eV than the b_1 and b_2 edges. The conduction-band minimum in the 10×10 superlattice has s symmetry and does not split.

For the neutral vacancy, the A_1-derived deep level is filled by two electrons, and one electron occupies the lowest of the T_2-derived states.

Similar behavior to that found for the As-vacancy levels is to be expected for all sp^3-bonded deep impurity levels in $GaAs/Al_xGa_{1-x}As$ superlattices, although the issue of whether a specific deep level lies in the fundamental band gap or not depends on the defect potential for that impurity and on N_1, N_2, and x.

3. *Cation-vacancy levels*

The A_1 bulk levels for a Ga vacancy in GaAs and for a cation vacancy in $Al_xGa_{1-x}As$ all lie very deep in the host valence bands and are not near the fundamental band gap, either in the bulk or at an interface of a $GaAs/Al_xGa_{1-x}As$ superlattice.

The T_2-derived cation-vacancy levels produce deep levels near the valence-band maxima of bulk GaAs, bulk $Al_xGa_{1-x}As$,[45,46] and the superlattice (see Fig. 6). In the Hjalmarson-type theory, the uncertainty in the predictions of absolute energies is typically a few tenths of an

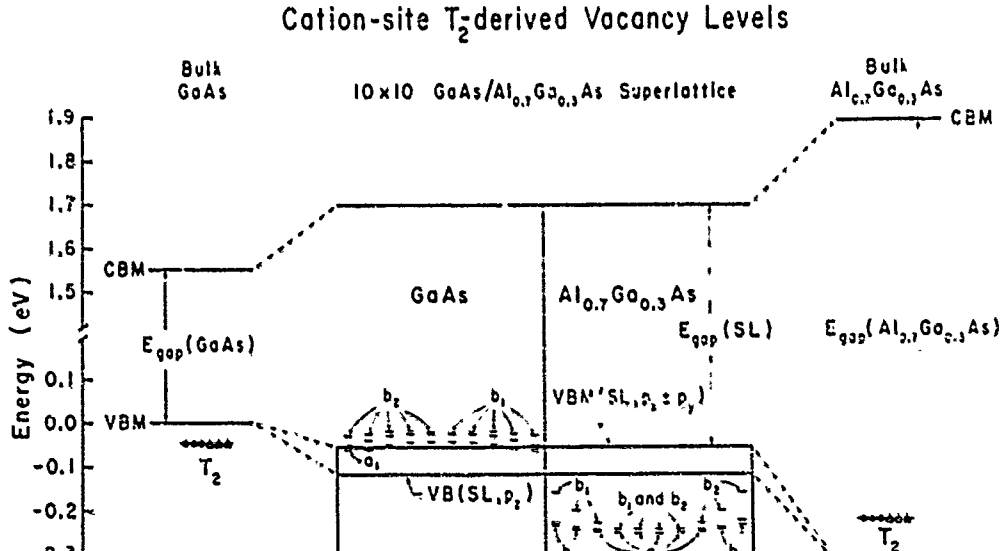


FIG. 6. Predicted T_2-derived vacancy levels of a cation vacancy in a $(GaAs)_m(Al_{0.7}Ga_{0.3}As)_n$ superlattice, as a function of site index β, the position of the vacancy. The zero of energy is the valence-band edge of bulk GaAs, and the corresponding valence- and conduction-band edges and deep levels in bulk GaAs and bulk $Al_{0.7}Ga_{0.3}As$ are given to the left and right of the central figure, respectively. The top of the central figure is the conduction-band edge of the superlattice, and the two bottom lines correspond to the b_1- and b_2-symmetric [$p_x \pm p_y$-like] valence-band maxima and the split a_1 (p_z-like) edge below it. Electrons occupying the bulk levels are denoted by solid circles. Holes are denoted by open triangles. When, as in bulk GaAs, the holes are in deep levels *below the valence-band maximum*, they bubble up to the valence-band maximum where the long-ranged Coulomb potential can trap them into shallow acceptor levels not shown. The cation vacancy in bulk GaAs and in an $Al_xGa_{1-x}As$ layer in the superlattice is predicted to be a triple shallow acceptor, providing three such holes to the valence band. In bulk $Al_xGa_{1-x}As$ and in a GaAs layer of the superlattice, neutral cation vacancies are predicted to produce deep traps for either electrons or holes. In the superlattice, the lowest-energy level is often of a_1 symmetry, and the highest is typically either of b_1 or b_2 symmetry, but exceptions to this rule do occur, as indicated on the figure.

eV and tends to be somewhat larger for the T_2 levels than for the A_1 levels. Therefore, in what follows, the reader should not interpret our predictions for the cation vacancy too literally or too quantitatively. Rather, the predictions illustrate the very interesting properties of a cation-site deep T_2 level that lies near but below the valence-band maximum of GaAs. Since this defect never lies at the interface (which, by definition in a $GaAs/Al_xGa_{1-x}As$ superlattice, is at an As layer), its level splittings are smaller (<0.05 eV) than the interfacial anion vacancy's splittings. In bulk GaAs, the T_2 cation-vacancy level is predicted[20,46] to lie $\simeq 0.03$ eV below the valence-band maximum. In bulk $Al_{0.7}Ga_{0.3}As$, the predicted vacancy level is in the gap, 0.11 eV above the $Al_{0.7}Ga_{0.3}As$ valence-band maximum. Near a $GaAs/Al_xGa_{1-x}As$ interface, the T_2 cation-vacancy level splits into a_1, b_1, and b_2 sublevels. For $x \simeq 0.7$ some or all of these sublevels may lie in the gap of the superlattice.

If the predictions are taken literally, then near the interface the cation vacancy produces a very interesting level structure, depending on the site of the vacancy. To begin with, in a GaAs layer the b_1 and b_2 vacancy levels lie in the gap of the superlattice, but in an $Al_{0.7}Ga_{0.3}As$ layer the vacancy levels are resonant with the valence band. This is due mainly to the band offset and the fact that, roughly speaking, the deep levels do not move (much), whereas the valence-band edges do, as one goes from bulk GaAs to the superlattice to bulk $Al_{0.7}Ga_{0.3}As$. The b_1 and b_2 levels in a GaAs layer lie typically $\simeq 0.03$ eV above the a_1 level, when we might have expected the a_1 level to lie between them. This expectation is not met because the T_2 levels are near the valence-band maximum and, in the superlattice, the T_2-like valence-band maximum is split into a_1 and b_1 and b_2 edges. Hence the a_1 valence-band edge has a stronger quantum-well-confinement effect: the band edge for a_1 (p_z-like) states lies 0.062 eV below the edge for b_1 and b_2 states. The a_1 defect states lie lower because the valence-band states that repel them are at lower energy in the superlattice. The largest splitting between b_1 and b_2 states is of order $\simeq 0.02$ eV, much smaller than the ($\simeq 0.1$–0.2)-eV splittings deduced from the anion vacancy—because the cation vacancy always lies at least one layer from an interface. In particular, the p_z-like a_1 level decreases in energy as the vacancy moves from the center of the GaAs layer toward the interface. The initial decrease is due to quantum-well confinement, and begins when the vacancy wave function significantly overlaps the $Al_{0.7}Ga_{0.3}As$. However, as the vacancy becomes quite close to the interface, its wave function penetrates thoroughly into the $Al_{0.7}Ga_{0.3}As$ layer and feels the electropositivity of the

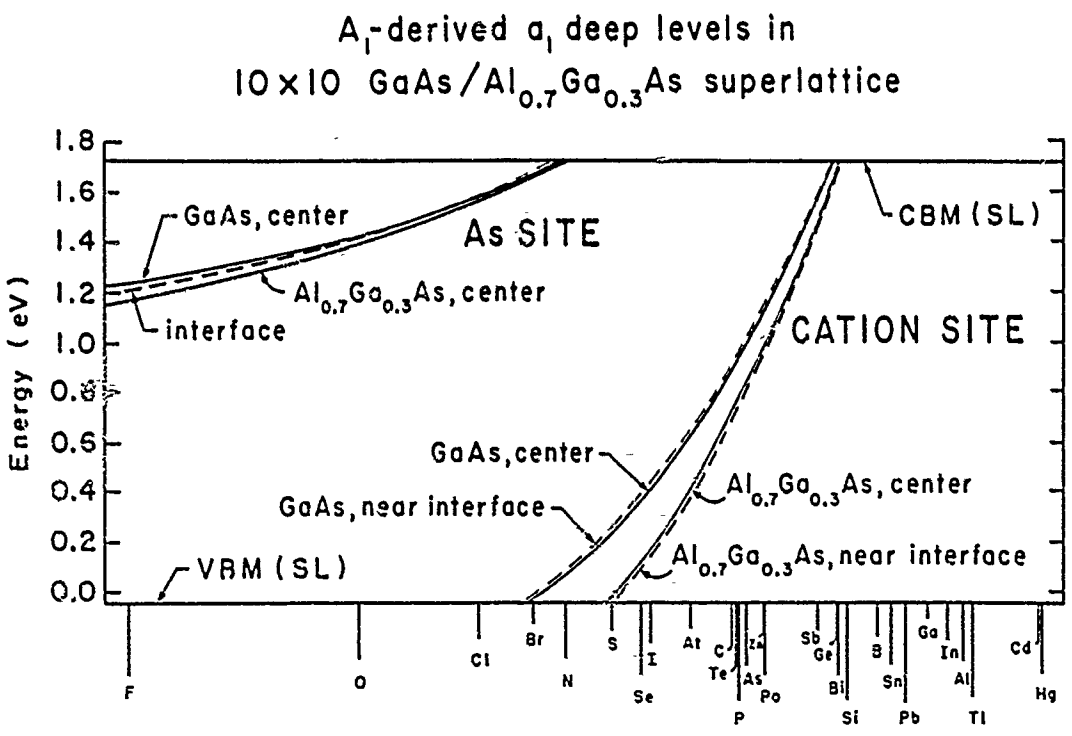


FIG. 7. Predicted energies of A_1-derived s-like a_1 deep levels of the indicated defects [Ref. 47] for a $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice with $N_1 = 10$ GaAs layers and $N_2 = 10$ $Al_{0.7}Ga_{0.3}As$ layers in its unit supercell. The dashed lines denote the predictions for defects at or near a $GaAs/Al_{0.7}Ga_{0.3}As$ interface, in the $\beta = 0$ layer for the interface, levels for As-site defects (upper left-hand corner), and, in the $\beta = 1$ and $\beta = 21$ layers adjacent to interface layers, for Ga-site defects in GaAs and cation-site defects in $Al_{0.7}Ga_{0.3}As$, respectively (lower right-hand corner). The solid lines denote the same levels for defects at or near the center of GaAs and $Al_{0.7}Ga_{0.3}As$ layers, respectively. In viewing this figure, remember that the theory is not precise, but that the general shape of the figure is reliable. The zero of energy is the valence-band maximum of *bulk* GaAs. The valence- and conduction-band edges of the superlattice are denoted VBM(SL) and CBM(SL), respectively.

Al. If the valence-band offset were smaller, say 0.2 eV, then this effect would cause the level to increase in energy near the interface.

The neutral Ga vacancy in the bulk of GaAs (assuming its deep level lies in the valence band) is a triple acceptor (Fig. 6). (A Ga vacancy is created by adding three holes—to remove the three Ga electrons—and letting the defect potential become infinite.[43]) Consider this vacancy at the near-interfacial site $\beta=1$ in a GaAs layer of the superlattice; its levels, in order of decreasing energy, are b_2, b_1, and a_1. In all of the cases of Fig. 6, the cation vacancy is either a triple shallow acceptor or a deep trap for both an electron and a hole, having at least one partially filled deep level in the gap. If it were the case that a cation vacancy near an interface had only one of its sublevels in the gap and two sublevels resonant with the valence band of the superlattice, then the vacancy would be a single acceptor, because only three electrons are available for four spin orbitals: the hole in the fourth spin orbital would "bubble up" to the valence-band maximum. Relatively small amounts of lattice relaxation or charge-state splittings of the defect levels could alter the predictions. Nevertheless, we think that studies of the character of the Ga-vacancy wave function in $GaAs/Al_xGa_{1-x}As$ superlattices are warranted—and could possibly reveal this shallow-deep transition in the character of the cation vacancy.

4. *Deep levels of sp^3-bonded impurities*

The predicted deep energy levels of substitutional sp^3-bonded impurities in a $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice are given in Figs. 7 and 8 for a superlattice with $N_1=N_2=10$.[31,47] These levels' energies E are obtained by solving the secular equation (17) with Hjalmarson's defect potential V slightly modified.[31] In these figures we display results for impurities at or near a $GaAs/Al_{0.7}Ga_{0.3}As$ interface and near the center of the GaAs and $Al_{0.7}Ga_{0.3}As$ layers.

For the A_1- or s-like a_1 levels on the As site (Fig. 7), most interfacial impurities have energy levels roughly midway between the levels for impurities near the center of the GaAs and the $Al_xGa_{1-x}As$ layers. There are two types of cation sites: Ga (in GaAs) and Al_xGa_{1-x} (in

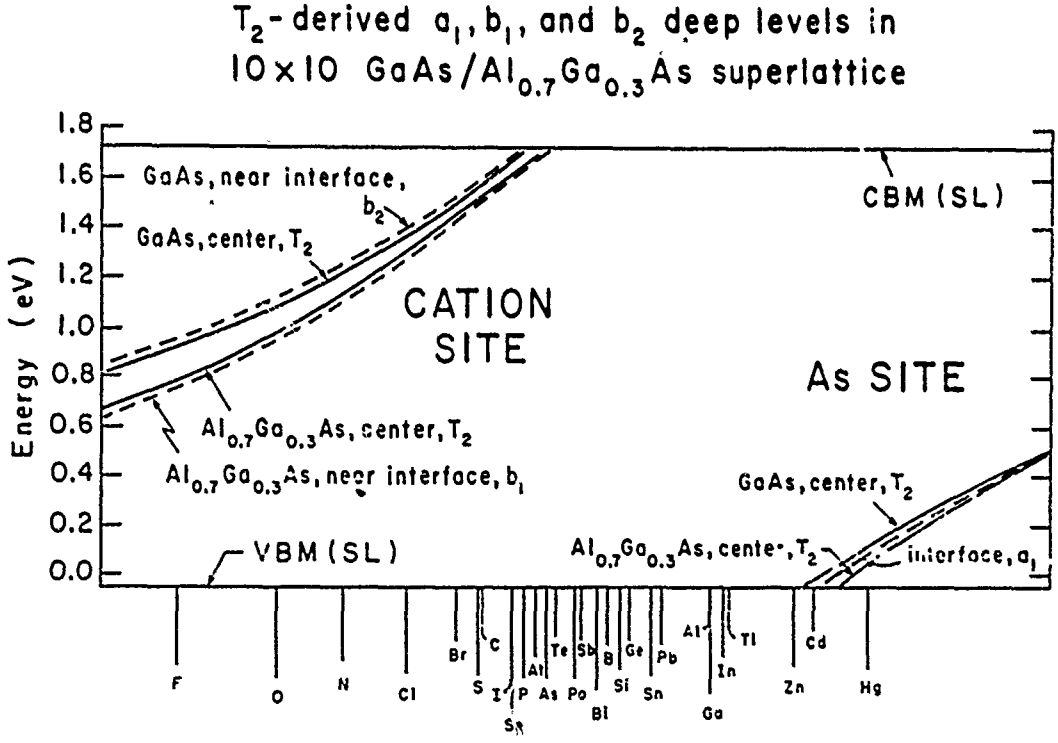


FIG. 8. The predicted T_2-derived deep levels for the indicated defects (Ref. 47) in a $(GaAs)_{10}/Al_xGa_{1-x}As)_{10}$ superlattice. For As-site defects (lower right-hand corner) the two solid lines correspond to T_2-derived b_1 levels (which are almost degenerate with b_2 levels) in the center of GaAs and $Al_{0.7}Ga_{0.3}As$ layers. The corresponding T_2-derived a_1 levels lie slightly (of order 0.01 eV) below the b_1 and b_2 levels and are not shown. The dashed line is an interfacial $\beta=0$ a_1 level. The corresponding b_2 level (not shown) is nearly degenerate with it, and the b_1 level (not shown) is above it of order 0.01 eV. For cation-site defects (upper left-hand corner) the solid lines correspond to T_2-derived b_1 levels (which are almost degenerate with b_2 and a_1 levels not shown). The dashed lines correspond to interfacial levels for site $\beta=19$ in GaAs of b_2 symmetry and for site $\beta=21$ of b_1 symmetry. The $\beta=19$ a_1 and b_1 levels are almost degenerate with the T_2 levels of defects in the center of GaAs and are not shown. Similarly, the $\beta=21$ ($Al_xGa_{1-x}As$) a_1 and b_2 (which are polarized perpendicular to the interface) interfacial levels are not shown. Near the $\beta=0$ interface, the b_1 and b_2 levels are interchanged. Simply stated, the levels b_1 or b_2 are split off from the bulk T_2 energy (in the material occupied by the defect) and the remaining two levels (almost) are at the T_2 bulk energy. The zero of energy is the valence-band maximum *of bulk GaAs*. The valence- and conduction-band edges of the superlattice are denoted VBM(SL) and CBM(SL), respectively.

$Al_xGa_{1-x}As$); the cation-site defects near the GaAs/$Al_xGa_{1-x}As$ interface have nearly the same deep levels as defects at the center of the layers—because those defects are at least one layer distant from the interface (which coincides with an As layer), and hence are only weakly perturbed by the atoms on the other side of the interface.

The T_2-derived levels in the GaAs layer (Fig. 8) are split when they lie near or at the interface. These are dangling-bond-p-like levels, and the p orbitals composed of hybrids oriented most toward the interface are split most [b_1 or (p_x+p_y)-like and a_1 or p_z-like on the Ga site; b_2 or (p_x-p_y)-like and a_1 at the As site], whereas the orbital directed away from the interface is least perturbed and normally has a level closest to that of the bulk T_2 level. The same physics hold for defects in the $Al_xGa_{1-x}As$ layer (Fig. 8), but the signs of the splittings and the orderings of b_1, a_1, and b_2 levels are normally reversed.

In general, three factors influence the relative positions of deep levels in GaAs and $Al_xGa_{1-x}As$ layers. (i) The more electropositive character of Al with respect to Ga pushes levels up in energy, so that the same defect has a tendency to exhibit higher-energy deep levels in $Al_xGa_{1-x}As$ than in GaAs. (See the T_2-derived As-vacancy levels in Fig. 5.) (ii) However, the Al also widens the band gap, causing a band offset, and rearranging the spectral distribution of host states somewhat (differently for the b_1 and b_2 states than for the a_1 states)—which often has the opposite effect on those impurity levels most affected by the host states near the valence-band maximum. (See the A_1-derived As-vacancy levels in Fig. 5.) (iii) In addition to the electropositivity and band-gap-widening effects, there is a quantum-well-confinement effect. Impurities in GaAs close enough to an $Al_xGa_{1-x}As$ layer that their wave functions overlap the $Al_xGa_{1-x}As$ barrier will experience level shifts due to confinement (that depend on the symmetry of the level). States near and above the conduction-band minimum with considerable conduction-band character are expected to move up in energy and valence-band states should move down due to confinement (see Fig. 6). The confinement effects should be more severe for a_1 states than for b_1 or b_2 states, due to the fact that their wave functions are polarized in the superlattice growth direction.

5. *Dependence on alloy composition*

In Fig. 9 we show our predictions for the A_1-symmetric deep level of cation-site Si (Ref. 31) in bulk $Al_xGa_{1-x}As$ as a function of x, which are similar to those first obtained by Hjalmarson.[48]

Figure 10, in comparison with Fig. 9, illustrates that a Si impurity in a GaAs/$Al_xGa_{1-x}As$ superlattice with $N_1=N_2$ may not produce a deep level in the superlattice band gap, although the same impurity in the alloy obtained by disordering the superlattice, $Al_{x/2}Ga_{1-(x/2)}As$, does produce a deep level in the gap of the alloy. To see this, consider $x<0.5$, for which the Si theoretical level does not lie in the gap of the 2-2

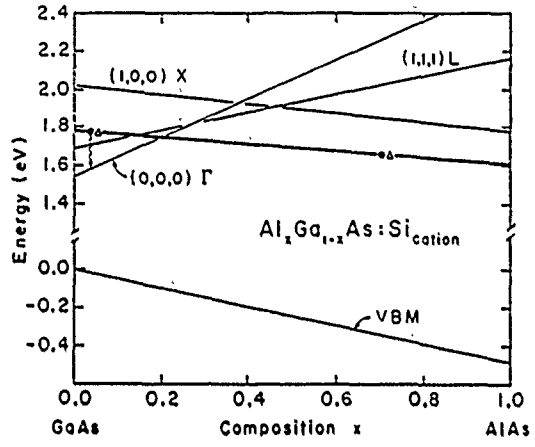


FIG. 9. Chemical trends with alloy composition x in the energies (in eV) of principal conduction-band edges Γ, L, and X, and the valence-band maximum of the alloy with respect to the valence-band maximum *of GaAs*, in the alloy $Al_xGa_{1-x}As$, as deduced from the Vogl model (Ref. 25). Also shown is the predicted energy of the A_1-symmetric cation-site deep level of Si (heavy line), similar to the predictions of Hjalmarson (Refs. 31 and 48). The Vogl model is known to obtain very little band bending. Moreover, the L minimum for $x\simeq 0.45$ is known to be at a bit too low an energy in this model. When the deep level of neutral Si lies below the conduction-band minimum, it is occupied by one electron (solid circle) and one hole (open triangle). When this level is resonant with the conduction band, the electron spills out and falls (wavy line) to the conduction-band minimum, where it is trapped (at zero temperature) in a shallow donor level (not shown).

GaAs/$Al_xGa_{1-x}As$ superlattice. For $x=0.25$ the Si level is in the gap of the alloy, however.

This situation is much more common in larger-period superlattices,[13] where the band edge of the superlattice is almost the band gap of the small-band-gap material (GaAs) (see Fig. 1). To a good approximation, if an im-

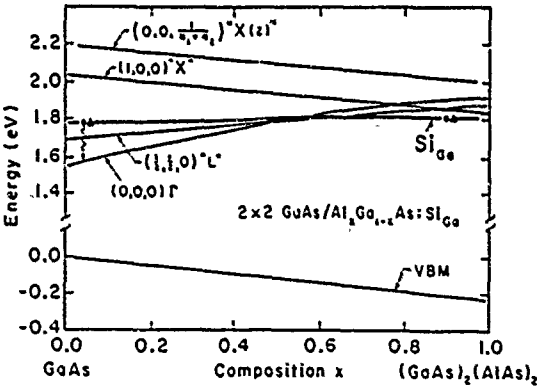


FIG. 10. Chemical trends with alloy composition x in the energies (in eV) of principal conduction- and valence-band extrema of a $(GaAs)_2/(Al_xGa_{1-x}As)_2$ superlattice, with respect to the valence-band maximum *of GaAs*. Compare with Fig. 9 for the alloy. The superlattice wave vectors of the minima are $k=0$, $k=(2\pi/a)(\frac{1}{2},\frac{1}{2},0)$, which has states derived from the L point of the bulk Brillouin zone, and the points derived from the bulk X point, $(2\pi/a)(1,0,0)$, $(2\pi/a)(0,0,1)$, which is p_z-like, and $(2\pi/a)(1,0,0)$, which is p_x- and p_y-like.

purity[20] (i) produces a deep level in the gap when the impurity lies in the large-band-gap bulk material ($Al_xGa_{1-x}As$), and (ii) has its corresponding deep level resonant with the host bands when the impurity occupies a site in the small-band-gap bulk material (GaAs), then the impurity's deep level will not lie in the gap of the large-period superlattice for the case that the impurity occupies a site in a small-band-gap (GaAs) layer. Thus in an 18×18 $GaAs/Al_{0.7}Ga_{0.3}As$ superlattice (Fig. 1), Si on a Ga site in a GaAs layer has its A_1-symmetric deep level in the conduction band of the superlattice, and so is a shallow donor. However, in the corresponding alloy with $x=0.35$ (Fig. 9) the cation-site Si level lies within the fundamental band gap and is a deep trap. Similar physics holds for other defects. This physics might be the origin of the improved transport characteristics of GaAs/AlAs superlattices over $Al_{0.5}Ga_{0.5}As$ alloys, as observed by Fujiwara *et al.*[13]

6. *Fluctuations*

For some time it has been known that fluctuations in alloy composition x of $Al_xGa_{1-x}As$ can cause normally shallow donors such as Si (for $x<0.3$) to produce deep levels in regions where the Al mole fraction is considerably larger than the average value x. Here we have shown that in thin quantum wells of GaAs in $GaAs/Al_xGa_{1-x}As$ superlattices for $x>0.3$ the Si can produce deep levels as well. Thus fluctuations in layer thickness can produce previously unanticipated deep traps. Since deep-level wave functions have large amplitudes only within a radius of $\simeq5$ Å of the impurity, a fluctuation in GaAs layer thickness (in the z direction) down to, say, 20 Å, occurring within a small circle of radius 5 Å in the x-y plane, would lead to Si deep levels in that region. Hence, special care may be necessary during superlattice growth to prevent nonuniformities in layer thickness, which could lead to such deep-level formation.

7. *Relationship between cation-site Si and the DX center*

For some time the technologically important native defect *DX* in $Al_xGa_{1-x}As$ has been the subject of discussion, with Lang[4] first proposing that it is a donor-vacancy complex, and Hjalmarson[4-20,49] setting forth the model of a Si impurity on a cation site. Clearly, the Si center discussed here has the energy levels and dependencies on alloy composition x required of a *DX* center. Our opinion is that the *DX* center is normally either a Si impurity, a complex containing a Si impurity (perhaps Si-Si pairs in some cases), or, in some instances, other impurities such as Sn that are similar to Si. The complexes should have spectra very close to the spectra of their constituents.[50] The persistent photoconductivity associated with the *DX* center appears to be best explained by the Hjalmarson-Drummond phonon-coupling model.[4] However, we note that Li *et al.*[51] dispute the Hjalmarson-Drummond conclusion and there remain unanswered questions concerning the *DX* center. We hope that the present work, which shows how the deep levels associated with Si should behave in the superlattice, may help in solving the mysteries surrounding this interesting defect.

IV. SUMMARY

The calculations presented here call into question the common assumption that the character of an impurity in a superlattice will always be the same as in the bulk. We have presented calculations which indicate that the normal shallow dopant Si in GaAs may become a deep trap in a GaAs quantum well of a $GaAs/Al_xGa_{1-x}As$ superlattice. This prediction should be tested for $x\gtrsim0.7$ superlattices, where it is likely to be most reliable.[52]

We have elucidated the physics of deep levels in superlattices, and find splittings of T_2 bulk levels and shifts of A_1 levels of order 0.1–0.2 eV for defects at the interface and less for impurities within two or three atomic planes of an interface. For impurities more distant from an interface the effect of the superlattice is primarily to change the *window of observability* of the deep level: If one imagines the deep levels as being relatively fixed in energy, the role of the superlattice is to provide the band gap; superlattices with small GaAs quantum wells have sufficiently large band gaps that deep levels which are covered up by the bands in bulk GaAs are uncovered and observable in the superlattice.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research (Contract No. N00014-84-K0352) for their generous support. We thank K. Hess for alerting us to the technological importance of this problem, and we have benefited from illuminating conversations with D. Wolford, W. Wang, H. Hjalmarson, and R. Hong.

---

*Permanent address.

[1] L. Esaki and R. Tsu, IBM Research Note RC-2418, 1969 (unpublished).

[2] R. Dingle, H. L. Störmer, A. C. Gossard, and W. Wiegmann, Appl. Phys. Lett. 33, 665 (1978).

[3] D. K. Maude, J. C. Portal, L. Dmowski, T. Foster, L. Eaves, M. Nathan, M. Heiblum, J. J. Harris, and R. B. Beall, Phys. Rev. Lett. 59, 815 (1987); M. D. Sturge, Appl. Phys. Lett. 33, 665 (1978).

[4] H. P. Hjalmarson and T. J. Drummond, Appl. Phys. Lett. 48, 657 (1986); see also Ref. 5.

[5] J. C. M. Henning and J. P. M. Ansems, Mater. Sci. Forum 10-12, 429 (1986); Semicond. Sci. Technol. 2, 1 (1987); A. K. Saxena, Solid State Electron. 25, 127 (1982).

[6] D. V. Lang and R. A. Logan, Phys. Rev. Lett. 39, 635 (1977); D. V. Lang, R. A. Logan, and M. Jaros, Phys. Rev. B 19, 1015 (1979).

[7] T. Ishibashi, S. Tarucha, and H. Okamoto, Jpn. J. Appl. Phys. 21, 476 (1982).

[8] L. G. Salmon and I. J. D'Haenens, J. Vac. Sci. Technol. B 2,

197 (1984).

[9]M. Mizuta, M. Tachekawa, H. Kukimoto, and S. Minomura, Jpn. J. Appl. Phys. 24, L143 (1985).

[10]P. M. Mooney, E. Calleja, S. L. Wright, and M. Heiblum, Mater. Sci. Forum 10, 417 (1986).

[11]See also, K. J. Malloy, Ph.D. thesis, Stanford University, 1984.

[12]T. Mimura, S. Hiyamizu, T. Fujii, and K. Nambu, Jpn. J. Appl. Phys. 19, L225 (1980); K. Hess, Physica 117&118B, 723 (1983); M. Inoue, S. Hiyamizu, M. Inayama, and Y. Inuishi; Jpn. J. Appl. Phys. Suppl. 22-1, 357 (1983); H. Sakaki, *ibid.* 21, 381 (1982).

[13]K. Fujiwara, H. Oppolzer, and K. Ploog, Inst. Phys. Conf. Ser. No. 74, 351 (1985); in *Gallium Arsenide and Related Compounds, 1984*, edited by B. de Cremoux (Hilger, Bristol, 1985).

[14]S. Das Sarma and A. Madhukar, J. Vac. Sci. Technol. 19, 447 (1981).

[15]A preliminary account of this work was given by J. Shen, S. Y. Ren, and J. D. Dow, Bull. Am. Phys. Soc. 30, 630 (1985). See also Ref. 16.

[16]R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Proc. Mater. Res. Soc. 77, 545 (1987).

[17]A preliminary account was given by H. P. Hjalmarson, Bull. Am. Phys. Soc. 32, 814 (1987). See also J. W. Farmer, H. P. Hjalmarson, and G. A. Samara, Proc. Mater. Res. Soc. (to be published).

[18]J. S. Nelson, C. Y. Fong, I. P. Batra, W. E. Pickett, and B. M. Klein (unpublished).

[19]M. Lannoo and J. Bourgoin (private communication).

[20]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980); see also W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977), for the concepts that form the foundation of this work.

[21]J. D. Dow, in *Highlights of Condensed Matter Theory*, Proceedings of the International School of Physics "Enrico Fermi," Course 89, Varenna, 1983, edited by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna/North-Holland, Amsterdam, 1985), pp. 465 *et seq.*

[22]S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).

[23]S. Y. Ren, Sci. Sin. 27, 443 (1984).

[24]The tight-binding formalism we use here is hybrid based. For bulk semiconductors the hybrid-based tight-binding formalism is equivalent to the widely used atomic-orbital-based tight-binding formalism. However, for superlattices these two formalisms are different: Because the atomic-orbital-based tight-binding parameters are obtained by fitting to the band structure of bulk GaAs and bulk AlAs, the tight-binding parameters for the As atom at $\beta=0$ (which is considered to be an As atom in GaAs) are usually different from the corresponding tight-binding parameters for the atom at $\beta=2N_1$ (which is considered to be an As atom in AlAs). However, both of these As atoms are interfacial As atoms and physically they are completely equivalent to each other. The way we correct for this problem is that we assign the tight-binding parameters to bonds rather than to atoms. We consider the hybrids h_1 and h_4 of the As atoms at $\beta=0$ to be hybrids of GaAs, while h_2 and h_3 are taken to be AlAs hybrids. Similarly, h_1 and h_4 of the As atom at $\beta=2N_1$ are considered to be hybrids in GaAs, and h_2 and h_3 are AlAs hybrids. We believe that this property accounts for the correct physics and the nature of interfacial bonds.

[25]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[26]D. J. Wolford, T. F. Kuech, J. A. Bradley, M. A. Grell, D. Ninno, and M. Jaros, J. Vac. Sci. Technol. B 4, 1043 (1986).

[27]In our model, which uses low-temperature band gaps, we take the valence-band offset for $Al_{0.7}Ga_{0.3}As$ to be 0.334 eV below the valence-band edge of GaAs. R. Dingle, W. Wiegmann, and C. H. Henry, Phys. Rev. Lett. 33, 827 (1974), had found a 15% offset in earlier work.

[28]*Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology*, New Series, Vol. 17a, *Semiconductors: Physics of Group IV and III-V Compounds*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer-Verlag, Berlin, 1982). The room-temperature lattice constant of AlAs is actually 5.660 Å, but here we take it to be equal to the GaAs lattice constant of 5.653 Å.

[29]We have also calculated the band structure, but not the deep levels, in 2×34 and 18×18 superlattices. This requires diagonalization of 360×360 matrices.

[30]S. Y. Ren and J. D. Dow, Phys. Rev. B 38, 1999 (1988).

[31]The defect-potential matrix elements are related to the difference in atomic energies of the impurity and the host atom it replaces:

$$V_l=\beta_l[w_{\text{impurity}}(l)-w_{\text{host}}(l)]+C_l\ ,$$

where we have $l=s$ or p, $\beta_s=0.8$, and $\beta_p=0.6$, and the atomic-orbital energies w can be found in Table III of Ref. 25. The constant C_l is zero in Refs. 20 and 25, but here we take $C_s=1.434$ eV in order to have the Si deep level in GaAs appear 0.234 eV above the conduction-band minimum, where it is observed (Ref. 3), instead of 0.165 eV below it, thereby compensating for the small theoretical uncertainty. In order to simplify the presentation of the results for the a_1 states, we use Hjalmarson's rule:[32] $V_s-C_s=2V_p$, thereby displaying E for the a_1 states as a function of V_s alone for the s-like state and as a function of V_p alone for the p_z-like state.

[32]H. P. Hjalmarson (private communication). This rule has been widely used in other contexts, most often with $C_l=0$: See, for example, J. D. Dow, R. E. Allen, and O. F. Sankey, in *The Chemistry and Physics of Solid Surfaces V*, Vol. 35 of *Springer Series in Chemical Physics*, edited by R. Vanselow and R. Howe (Springer-Verlag, Berlin, 1984).

[33]D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).

[34]J. N. Schulman and T. C. McGill, Phys. Rev. B 19, 6341 (1979).

[35]A. C. Gossard, P. M. Petroff, W. Weigmann, R. Dingle, and A. Savage, Appl. Phys. Lett. 29, 323 (1976).

[36]For a review of deep-level theories, see, for example, Ref. 21 and references therein.

[37]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

[38]L. A. Hemstreet, Phys. Rev. B 15, 834 (1977); P. Vogl and J. Baranowski, Acta Phys. Polonica A67, 133 (1985); M. Scheffler, J. P. Vigneron, and G. B. Bachelet, Phys. Rev. Lett. 49, 1765 (1982); J. Bernholc, N. O. Lipari, S. T. Pantelides, and M. Scheffler, Phys. Rev. B 26, 5706 (1982); M. Jaros and S. Brand, *ibid.* 14, 4494 (1976); U. Lindefelt, J. Phys. C 12, L419 (1979); D. A. Papaconstantopoulos and E. N. Economou, Phys. Rev. B 22, 2903 (1980); G. A. Baraff and M. Schlüter, Phys. Rev. Lett. 41, 892 (1978); Phys. Rev. B 19, 4965 (1979); J. Bernholc and S. T. Pantelides, *ibid.* 18, 1780 (1978); J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. Lett. 41, 895 (1978); Phys. Rev. B 21, 3545 (1980); J. Bernholc, S. T. Pantelides, N. O. Lipari, and A. Baldereschi,

Solid State Commun. 37, 705 (1981).

[39]The deep-level energies are comparably independent of N_2.

[40]S. Y. Ren, J. D. Dow, and D. J. Wolford, Phys. Rev. B 25, 7661 (1982).

[41]D. J. Wolford, J. A. Bradley, K. Fry, and J. Thompson, in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, edited by D. J. Chadi and W. A. Harrison (Springer-Verlag, New York, 1984), p. 627.

[42]X. S. Zhao, G. H. Li, H. X. Han, Z. P. Wang, R. M. Tang, and R. Z. Che, Acta. Phys. Sin. 33, 588 (1984) [Chinese Phys. 5, 337 (1985)].

[43]M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969).

[44]M. S. Daw and D. L. Smith, Phys. Rev. B 20, 5150 (1979); Solid State Commun. 37, 205 (1981); J. Vac. Sci. Technol. 17, 1028 (1980); Appl. Phys. Lett. 36, 690 (1980); M. S. Daw, D. L. Smith, C. A. Swarts, and T. C. McGill, J. Vac. Sci. Technol. 19, 508 (1981).

[45]For $Al_{0.7}Ga_{0.3}Ga$ and GaAs, the T_2 cation-vacancy level is predicted to lie at 0.11 and -0.04 eV, respectively, with respect to the relevant valence-band maximum.

[46]See, for example, J. P. Buisson, R. E. Allen, and J. D. Dow, J. Phys. (Paris) 43, 181 (1982); E. S. Ho and J. D. Dow, Phys. Rev. B 27, 1115 (1983).

[47]The abscissae of Figs. 7 and 8 are, to within an additive constant, $\beta_s w(s)$ and $\beta_p w(p)$, where the atomic-orbital energies in the solid w are tabulated in Ref. 25, and we have $\beta_s=0.8$ and $\beta_p=0.6$.

[48]H. P. Hjalmarson (private communication).

[49]H. P. Hjalmarson, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1979 (unpublished).

[50]O. F. Sankey, H. P. Hjalmarson, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. Lett. 45, 1656 (1980); O. F. Sankey and J. D. Dow, Appl. Phys. Lett. 38, 685 (1981); J. Appl. Phys. 52, 5139 (1981); Phys. Rev. B 26, 3243 (1982).

[51]M.-F. Li and P. Y. Yu, Solid State Commun. 61, 13 (1987); M.-F. Li, P. Y. Yu, E. R. Weber, and W. Hansen, Appl. Phys. Lett. 51, 349 (1987).

[52]A casual examination of the theory, Fig. 7, indicates that Sn should have a deep level ~ 0.4 eV above the Si level (with B in between). However, Sn experimentally lies only 0.056 eV above Si in GaAs,[3] the difference being attributable within the current theory to lattice relaxation. Hence Sn doping rather than Si doping is unlikely to represent much of an improvement in small-period superlattices.

PHONONS IN $Al_xGa_{1-x}As$ ALLOYS

Akiko Kobayashi*
Department of Physics and Materials Research Laboratory
University of Illinois at Urbana-Champaign
1110 West Green Street, Urbana, Illinois 61801
and
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

and

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

and

Eoin P. O'Reilly**
Department of Physics and Materials Research Laboratory
University of Illinois at Urbana-Champaign
1110 West Green Street, Urbana, Illinois 61801

(Received 24 February 1985)

Theoretical densities of phonon states in random $Al_xGa_{1-x}As$ alloys are presented and compared with published Raman scattering data. The calculations employ the recursion method and a Born-von Karman rigid-ion model of the lattice dynamics; they permit the assignment of major spectral features either to parent modes in GaAs and AlAs or to vibrations of specific alloy clusters. The phonon spectral densities of states are shown to be nearly "persistent" — a linear superposition of GaAs and AlAs spectral densities. Deviations from the persistence limit are attributed to "alloy modes." Many assignments of lines in Raman spectra are confirmed, but some lines are reinterpreted.

1. Introduction

In this paper we report calculations of the densities of phonon states for $Al_xGa_{1-x}As$ substitutional alloys, and compare our results with Raman scattering data of Tsu, Kawamura, and Esaki [1], Kim and Spitzer [2], Saint-Cricq et al. [3], and Jusserand and Sapriel [4]. We show that the densities of states of these alloys are, more or less, "persistent" in the sense of Onodera and Toyozawa [5]: except in three spectral regions, the densities of states of the alloys $Al_xGa_{1-x}As$ are linear superpositions of the densities of states of GaAs and AlAs:

$$D(\Omega;Al_xGa_{1-x}As)=(1-x)D(\Omega;GaAs)+xD(\Omega;AlAs). \quad (1)$$

The deviations from the above are attributed to "alloy modes", namely characteristic vibrations of small clusters in alloys not present in either GaAs or AlAs. The main features of the Raman data are all accounted for, including disorder-activated Raman modes which arise from either persistent modes or alloy modes in the densities of states that are Raman-forbidden in perfect zincblende compounds.

We employ the recursion method [6-9] in our calculations, treat the ions as rigid, include first- and second-nearest-neighbor force constants, and neglect long-ranged Coulomb forces. To our knowledge, these are the first calculations of this type for III-V semiconductive alloys. Coherent potential approximation (CPA) calculations [10] have been reported previously. However, the CPA, in its usual form, breaks down whenever the local modes of clusters of minority atoms become important -- i.e., on the order of x^2 or $(1-x)^2$ and higher. Because of the persistent nature of the densities of states, Eq. (1), the CPA, or almost any alloy theory, will successfully predict the main features of the alloy spectra; but the

* Present Address: Department of Physics and Astronomy, University of Maryland, College Park, MD 20742.
** Present Address: Department of Physics, University of Surrey, Guildford Surrey GU2 5XH, U.K.

0749-6036/85/060471+09 $02.00/0

©1985 Academic Press Inc. (London) Limited

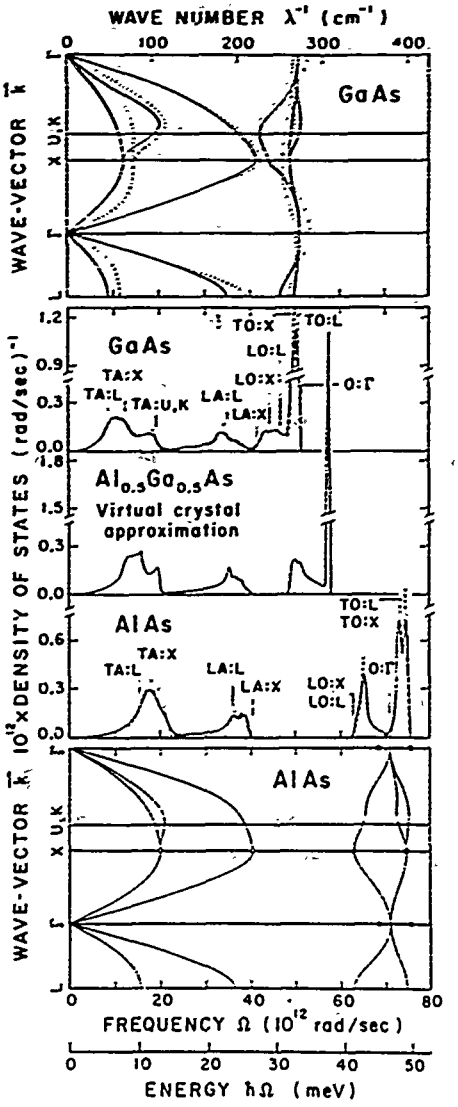


Fig. 1. Displayed from top to bottom are phonon dispersion curves for GaAs, densities of phonon states for GaAs, density of phonon states for $Al_{0.5}Ga_{0.5}As$, densities of phonon states for AlAs, and phonon dispersion curves for AlAs, respectively. The phonon dispersion curves for GaAs (solid lines) were obtained using the force-constant parameters of Ref. [11], and are compared with the neutron scattering data of Ref. [18] (dotted lines). The phonon dispersion curves for AlAs (solid lines) were obtained in the present model, and are compared with the infrared reflection data and the Raman scattering data of Ref. [13] at the Γ point (circles), and the optical absorption data of Ref. [14] at the X point (circles). (It was pointed out by Barker et al. [15] that the assignments for TO:X and LO:X, given in Ref. [14], should be reversed. Here we use the assignment of Barker et al.) The phonon densities of states for GaAs and AlAs were obtained by the recursion method (solid lines) and are compared with the densities of states obtained by the Lehmann-Taut method (dashed lines). The density of states for $Al_{0.5}Ga_{0.5}As$ was obtained by the Lehmann-Taut method using the virtual-crystal approximation for the force-constant parameters and the cation-site mass in this alloy. Phonons at specific wavevectors in the Brillouin zone, such as the optic mode at Γ, O:Γ, have those wavevectors indicated after a colon. The main symmetry points are $\Gamma=(0,0,0)$, $L=(2\pi/a_L)(1/2,1/2,1/2)$, $X=(2\pi/a_L)(1,0,0)$, $U=(2\pi/a_L)(1,1/4,1/4)$, and $K=(2\pi/a_L)(3/4,3/4,0)$.

deviations from the persistence limit — which are the non-trivial alloy [illegible] — are not reliably predicted by the VCA and most other alloy theories. In contrast, the recursion method is capable of including all orders of x, and predicts accurate alloy spectra provided that the proper ensemble averages over the various local environments are evaluated.

2. Outline of the calculation.

The first step of our method is to obtain force constants by fitting the phonon dispersion relations of GaAs and AlAs with a first- and second-neighbor rigid-ion Born - von Karman force constant model. For GaAs, the force constants of Banerjee and Varshni [11] were found to be excellent (here we set their long-ranged force constant "x" to zero); for AlAs the elastic constant data for c_{12} [12] and the following measured squared phonon frequencies were used to determine the six independent force constants of the model: $[\Omega^2(LO{:}\Gamma)+2\Omega^2(TO{:}\Gamma)]/3$, $\Omega^2(TO{:}X)$; $\Omega^2(LO{:}X)$; $\Omega^2(LA{:}X)$; and $\Omega^2(TA{:}X)$ [13-15]. (Here we denote optic and acoustic modes by O and A, longitudinal and transverse modes by L and T, and the wavevector $\vec{k}$ using standard symbols [16]. Hence LA:X refers to the longitudinal acoustic mode at the X-point of the Brillouin zone.) The force constants ν, μ, and λ for anions and cations in AlAs, in the notation of Banerjee and Varshni [11], were assumed to be related according to a central-force model [17]: ν=μ-λ. The resulting phonon dispersion curves are compared with data [13-15,18] in Fig. 1 (top and bottom), and the force-constant parameters obtained in the present model are tabulated in Table I.

No long-ranged forces are included in the fit to the phonon dispersion relations; hence the observed splitting at the Γ point of the Brillouin zone between the longitudinal optic (LO) and transverse optic (TO) modes cannot be reproduced by the model. However, this splitting is small, ~18 cm^{-1} in GaAs and ~38 cm^{-1} in AlAs, in comparison with optical mode energies of ~270 cm^{-1} and ~390 cm^{-1}, respectively. While explicitly ignoring the long-ranged Coulomb forces, we note that a model with first- and second-nearest-neighbor force constants can at least adequately represent some of the effects of the long-ranged Coulomb forces at the zone-boundary points (which dominate the density-of-states spectra). Moreover, the qualitative differences in the phonon dispersion relations due to the long-ranged Coulomb forces are major only in the long-wavelength limit (near the Γ point in a Brillouin zone); the direct effect on the global features of the densities of phonon states is rather small [19]. Caution should be taken, however, when vibrational properties in the long-wavelength limit are of concern, such as in Raman spectroscopy or infrared reflection spectroscopy.

To obtain densities of states we employ 1000-atom clusters and execute the recursion method to 51 levels of continued fractions; such parameters were found to be necessary to obtain good representations of the phonon densities of states of GaAs and AlAs. Details of the calculations can be found in Ref. [20]. In Fig. 1, we display the resulting phonon densities of states for GaAs and AlAs obtained by the recursion method (solid line) compared with those obtained by the Lehmann-Taut method [21] (dashed line). In the densities of states obtained by the recursion method, we note that the sharp van Hove singularities [22] are blurred. This is characteristic of the recursion method, which has a slow convergence to spectral singularities, as they are explicitly due to the infinite periodicity of the crystal lattice.

3. Results

In Fig. 1, we also display the density of states for the alloy $Al_{0.5}Ga_{0.5}As$, calculated using the virtual-crystal approximation in which the force constants and the cation-site masses are linearly interpolated between those of GaAs and AlAs. This spectrum exhibits only one "amalgamated" [5] optic band and thus the well-known "two-mode behavior" [13,23] of this alloy is not reproduced by the virtual crystal approximation.

In the recursion-method theory presented here, the force-constant parameters for an alloy of composition x are determined using those of the parent compounds (Table I). For $Al_xGa_{1-x}As$, the first-nearest-neighbor force constants, α, of pure GaAs and of pure AlAs are used for Ga-As pairs and Al-As pairs in the alloy, namely,

$$\alpha(Al_xGa_{1-x}As;Ga\text{-}As) = \alpha(GaAs)$$

$$\alpha(Al_xGa_{1-x}As;Al\text{-}As) = \alpha(AlAs)\ . \qquad (2)$$

We use the same relation for $\beta(Al_xGa_{1-x}As;Ga\text{-}As)$ and $\beta(Al_xGa_{1-x}As;Al\text{-}As)$. The second-nearest-neighbor force constants, λ, for four possible pairs of atoms in the alloy are determined by:

$$\lambda(Al_xGa_{1-x}As;Ga\text{-}Ga)=\lambda_c(GaAs)$$

$$\lambda(Al_xGa_{1-x}As;As\text{-}As)=(1-x)\lambda_a(GaAs)+x\lambda_a(AlAs)$$

$$\lambda(Al_xGa_{1-x}As;Al\text{-}Al)=\lambda_c(AlAs)$$

$$\lambda(Al_xGa_{1-x}As;Ga\text{-}Al)=(1-x)\lambda_c(GaAs)+x\lambda_c(AlAs), \qquad (3)$$

and we use the same relation for the other second-nearest-neighbor force constants μ and ν for these four possible pairs of atoms. The distribution of atomic masses at various zincblende lattice sites is determined for $Al_xGa_{1-x}As$ as follows: (i) all anion-site atoms are As atoms; (ii) at a specific cation-site, the mass is that of Al with probability x or the mass of Ga with probability 1-x (as determined by a random number generator).

TABLE I. Force-constant parameters (in units of 10^3 dyne/cm) and masses (in units of 10^{-24} g). Here the labels a and c refer to the anion and the cation, respectively.

	GaAs	AlAs
α	-39.525[a]	-41.546
β	-34.000[a]	-37.760
λ_a	4.500[a]	1.772
λ_c	4.500[a]	-4.246
μ_a	-3.697[a]	-2.217
μ_c	-4.467[a]	-0.693
ν_a	-3.697[a]	-3.989
ν_c	-4.467[a]	3.553
M_a	124.4102	124.4102
M_c	115.7722	44.8044

[a] Parameters for GaAs obtained from Ref. [11].

The recursion method has been described elsewhere [6-9], and is well-suited for calculating the local density of states at a specific site in a disordered system: $d(\vec{R};\Omega)$. Since $d(\vec{R};\Omega)$ depends sensitively on the local atomic environment, in order to obtain the total density of states, which is insensitive to specific local order, we sum over an ensemble of local densities of states as follows: (i) at the center of the 1000-atom cluster, a specific five-atom minicluster is generated, e.g., a central As atom surrounded by one Ga and three Al atoms; (ii) the probability p(x) of this cluster occurring in an alloy with composition x is determined (see Appendix); (iii) the remaining 995 atoms of the cluster are added, with As atoms on anion sites and either Ga (with probability 1-x) or Al (with probability x) atoms on cation sites; and (iv) the local density of states at the central site of the minicluster, embedded in its alloy environment of 995 atoms, is computed. The process is repeated for all possible miniclusters, each embedded in the 995 atom cluster. The anion-site density of states is then given by

$$D_{anion}(\Omega) = \Sigma_i \, p_i(x) \, d_i(\vec{R};\Omega) \, , \tag{4}$$

where the sum is over all possible miniclusters, and d_i and p_i are the local density of states and the probability of occurrence of the i^{th} minicluster respectively. The cation-site density of states is determined from the local densities of states of a Ga atom and an Al atom, each of which is surrounded by four As atoms, and it can be written as

$$D_{cation}(\Omega) = (1-x) \, d_{Ga}(\vec{R};\Omega) + x \, d_{Al}(\vec{R};\Omega) \, . \tag{5}$$

The total density of states is the sum of the anion-site and cation-site local densities of states.

The calculated total densities of phonon states for $Al_xGa_{1-x}As$, for various compositions x, are shown in Fig. 2. To a good approximation these spectra are "persistent", namely they appear to be linear superpositions of GaAs and AlAs spectra (See the dashed line of Fig. 2), as determined by Eq. (1). In three spectral regions, there are significant deviations from this "persistent" behavior: (i) near 75 cm^{-1}, in the acoustic band, (ii) near 250 cm^{-1}, in the GaAs-like optic band, and (iii) near 370 cm^{-1}, in the AlAs-like optic band. These are all "alloy modes". By examining the local densities of states at various sites for different alloy configurations, we are able to associate each of these features with vibrations of specific small clusters. We discuss this in the following.

In the midrange of alloy compositions, x=0.5, the transverse and longitudinal acoustic modes are primarily associated with the vibrations of As atoms that are surrounded by Ga and Al atoms at the nearest-neighbor sites. The deviation from the persistence spectrum at ~75 cm^{-1} is due to fluctuations in the alloy composition: it comes from the contributions of clusters containing Ga and Al atoms which are present in a ratio different from 50% each.

The persistence spectrum (dashed line of Fig. 2) near 270 cm^{-1} is composed mainly of the following two types of vibrational modes: (i) vibrations of a central Ga atom that has Ga atoms for its twelve second nearest neighbors, and (ii) vibrations of a central As atom that has Ga atoms for its four first nearest neighbors. The deviation from the persistence spectrum around this frequency region, namely the shoulder at ~250 cm^{-1}, is thus attributed to the following alloy modes: (i) vibrations of a central Ga atom with Al atoms replacing Ga atoms at the second-nearest-neighbor sites, and (ii) vibrations of a central As atom with Al atoms replacing Ga atoms at the first-nearest-neighbor sites.

The persistence spectrum (dashed line of Fig. 2) at ~345 cm^{-1} and ~390 cm^{-1} is due to vibrations of a central Al atom that has Al atoms at the second-nearest-neighbor sites. The deviations from the persistence limit are attributed to the alloy modes, which are vibrations of a central Al atom with Ga atoms replacing Al atoms at the second-nearest-neighbor sites.

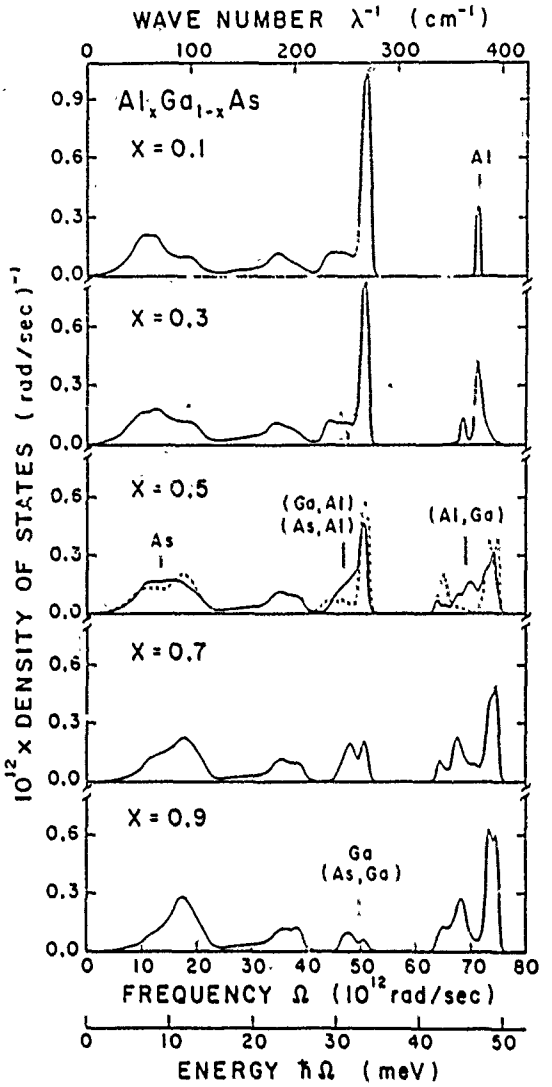


Fig. 2. Densities of phonon states D(Ω) for $Al_{1-x}Ga_xAs$ alloys obtained by the recursion method (solid lines) along with density of states obtained using the persistence approximation, Eq. (1), for x = 0.5 (dashed line). The assignments given to prominent peaks represent clusters or atoms that are responsible for the vibrations giving rise to the peaks. For example, As denotes clusters with a central As atom; (As,Al) denotes clusters of a central As atom with Al nearest-neighbors; and (Ga,Al) denotes clusters of a central Ga atom with Al second-nearest neighbors (or nearest cation neighbors).

The clusters responsible for the impurity modes at alloy composition x = 0.1 (GaAs:Al local mode) and x = 0.9 (AlAs:Ga gap mode) are as follows: (i) the sharp peak near 370 cm^{-1} of $Al_{0.1}Ga_{0.9}As$ is due to isolated Al atoms surrounded by a mostly-GaAs environment, and (ii) the peak near 260 cm^{-1} of $Al_{0.9}Ga_{0.1}As$ is due to vibrations of isolated Ga atoms surrounded by a mostly-AlAs environment, and due to the vibrations of As atoms bonded with these isolated Ga atoms.

The fact that the alloy spectra are persistent (and definitely not virtual-crystal-like or amalgamated -- see Fig. 1), especially for optic modes, means that the principal spectral features can be roughly labeled by the quantum numbers of the parent compounds GaAs and AlAs -- despite the fact that the wave-vector k and other symmetry labels for the quantum numbers are not strictly valid for the alloy. Therefore, any reasonable alloy theory, which is based on experimentally known optic frequencies of the parent compound semiconductors, should be able to explain the principal features of the data. This is why very simple models, such as the random-element-isodisplacement (REI) model [23-25], have been successful for the purpose of obtaining optic mode frequencies as a function of alloy composition. The features that are difficult for any model to predict are the alloy modes. In this regard, we note that the present method, despite its limitation to short-range forces, is clearly capable of including any alloy modes, whereas most other theories, including the coherent potential approximation (CPA) [10] and other similar effective-medium theories, are inappropriate for including the effects of clusters of minority atoms -- i.e., effects on the order of x^2 or $(1-x)^2$.

4. Comparison with data

In comparing the calculated densities of states with Raman or infrared data, one must make allowances for the deficiencies of the theory, including the facts that the calculated spectra do not include Raman or infrared matrix elements, and the model neglects the polarizability of the bonds and the long-ranged Coulomb forces. Thus, when viewing the theoretical predictions, it may be necessary to mentally shift peaks of order ≈20 cm^{-1}, split peaks, broaden them, and change their intensities in order to bring the theory into coincidence with the data. Nevertheless, even with these uncertainties, the theory can be very useful for interpreting data.

a. x=0.76

Fig. 3(a) shows Raman data for $Al_xGa_{1-x}As$ with alloy composition x = 0.76, obtained by Tsu et al. [1], and the best-resolved spectrum, at T = 350 K, among the resonant Raman-scattering data taken by Jusserand and Sapriel [4]. The experimental conditions for obtaining these data, in a perfect crystal, would yield only the LO:Γ modes: backscattering from the (100) surface with incident and scattered light polarizations parallel to the [011] crystal axis. However the disorder of the alloy breaks

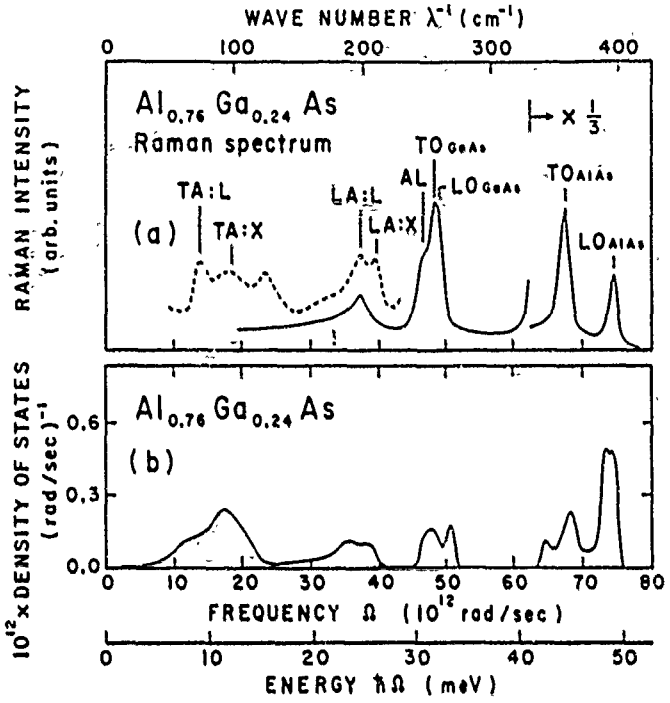


Fig. 1. (a) Raman spectrum for $Al_{0.76}Ga_{0.24}As$ after Ref. [1] (solid lines) and resonant Raman spectrum for $Al_{0.75}Ga_{0.25}As$ after Ref. [4] (dashed line); (b) calculated density of phonon states for $Al_{0.76}Ga_{0.24}As$. Various mode assignments of Refs. [1] and [4] are indicated. AL denotes an acoustic local mode.

the Raman selection rules, leading to the activation of other non-LO modes. Thus, while the theory (See Figs. 1 and 3(b)) confirms the general assignment of the LO_{AlAs} peak, it suggests that disorder-activated zone-boundary TO:L and TO:X modes, which produce a peak in the calculated density of states at ~390 cm^{-1} (Fig. 3(b)), contribute as well to this peak — giving it breadth and an asymmetric shape.

Concentrating on the fact that the selection rules forbid the TO modes (in the crystal), we speculate that the TO_{AlAs} assignment of the ~360 cm^{-1} peak should be revised: the theory (Fig. 3(b)) has an alloy mode at ~360 cm^{-1}, which results from (See Fig. 2) vibrations of Al atoms with some Ga atoms replacing Al atoms at the second-nearest-neighbor sites (typically three Ga atoms and nine Al atoms at the second-nearest-neighbor sites for x = 0.76).

We confirm the other major mode assignments of Ref. [1]: the peaks designated as GaAs-like LO and (disorder-activated) TO modes in the data correspond to the persistent part of the GaAs-like optic band at ~270 cm^{-1} in the density of states (See Fig. 1). The shoulder of the GaAs-like optic peak (AL: denoted Acoustic Local mode in the data) can be attributed to the alloy mode at ~250 cm^{-1} in the density of states (Fig. 2). This mode is due to vibrations of As atoms with Al atoms replacing Ga atoms at the nearest-neighbor sites (typically three Al atoms and one Ga atom at the nearest-neighbor sites for x = 0.76).

The two lowest bands observed in the resonant Raman-scattering data of Ref. [4] (Fig. 3(a)) are the disorder-activated longitudinal acoustic mode and the disorder-activated transverse acoustic mode respectively. The acoustic region of the density of states, for this alloy composition, retains most of the features of AlAs, and can be crudely assigned with the symmetry points of AlAs. Comparing each feature in the experimental data with the density of states (Fig. 1), we confirm in general the LA:X(AlAs), TA:X(AlAs), and LA:L(AlAs) peak assignments, while noting that, respectively, disorder-activated LA:U,K(AlAs) and TA:L(AlAs) modes should also contribute to these peaks. We tentatively reassign the peak labeled TA:L(AlAs) in the data, at least in part, to an alloy mode which results from vibrations of As surrounded by (typically three) Al atoms and (at least) one Ga atom.

b. x=0.54

Figure 4(a) shows the Raman data obtained by Kim and Spitzer [2] for alloy composition x = 0.54. The AlAs-like LO mode and the GaAs-like LO mode are very sharp, obeying the

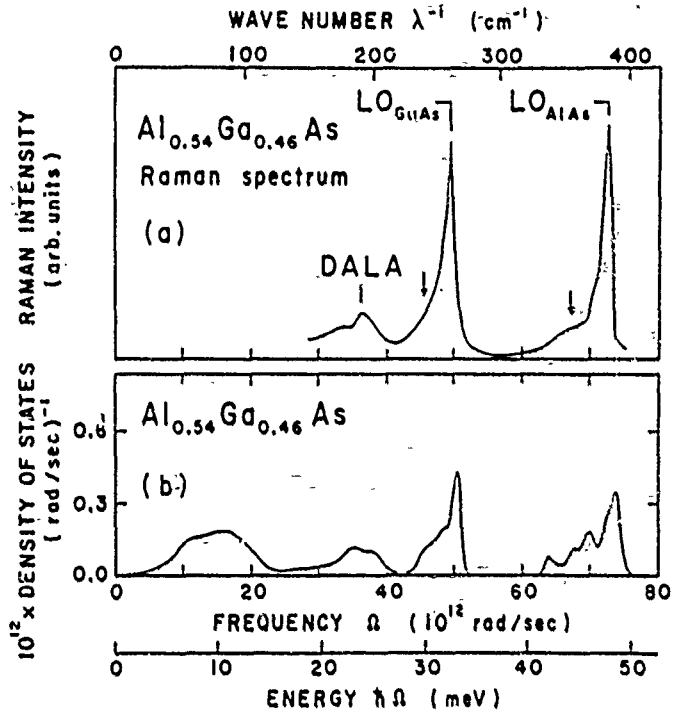


Fig. 4. (a) Raman spectrum for $Al_{0.54}Ga_{0.46}As$ after Ref. [2]; (b) calculated density of phonon states for the same alloy composition. The disorder-activated longitudinal acoustic mode is denoted DALA.

selection rules for this experimental geometry (backscattering from the (001) surface with light polarizations parallel to the [110] crystal axis). The theory confirms the DALA, LO_{GaAs}, and LO_{AlAs} assignments. Furthermore, comparing the data with the density of states, the shoulders of the two main peaks appear to have resulted from the alloy modes which are disorder activated. The clusters responsible for these modes are nearly the same as those in the case of x = 0.5, as described earlier (Fig. 2): The LO_{GaAs} mode has sidebands assigned to clusters with a central Ga atom and some Al atoms as second-neighbors and to clusters with As surrounded by some nearest-neighbor Al atoms. The LO_{AlAs} line has a low-energy tail due to vibrations of clusters with central Al atoms with some Ga second-nearest-neighbors.

iii. x=0.2

Unlike the data for x=0.76 and x=0.54, the Raman spectrum of Fig. 5(a) [3] was measured under conditions meant to forbid the LO:Γ and TO:Γ modes (in the case of a perfect zincblende crystal): scattering from the (100) surface with light polarizations parallel to the [001] crystal axis. Hence the main peaks are due to broken selection rules and are disorder-activated (DA). The peak assigned to two-phonon scattering by transverse acoustic modes (2TA) and the disorder-activated AlAs-like optic mode (DAO) near 370 cm^{-1} were removed from the spectrum by the authors of Ref. [3], yielding the processed spectrum of Fig. 5(b) -- which those authors argue is similar to the density of vibrational states for GaAs. Our calculations confirm their density of states argument and support their assignments. The somewhat sharper feature in the theory at ~280 cm^{-1} is an artifact of the theoretical model due to the omission of long-ranged forces: as a result the TO:Γ and LO:Γ modes are degenerate in the theory but split in the data.

5. Summary

The recursion calculations provide a good account of the phonons in $Al_xGa_{1-x}As$ alloys. In general, the spectra are well-described by the persistence limit, Eq. (1), which is why so many assignments of alloy spectral lines in terms of phonons in GaAs or AlAs have been correct. In particular the "two-mode" behavior of the $Al_xGa_{1-x}As$ optic mode is a consequence of the persistent nature of phonons in these alloys: thus one observes distinct AlAs-like and GaAs-like optical phonon modes.

The genuine alloy effects are the "alloy modes" that are not well-described by the

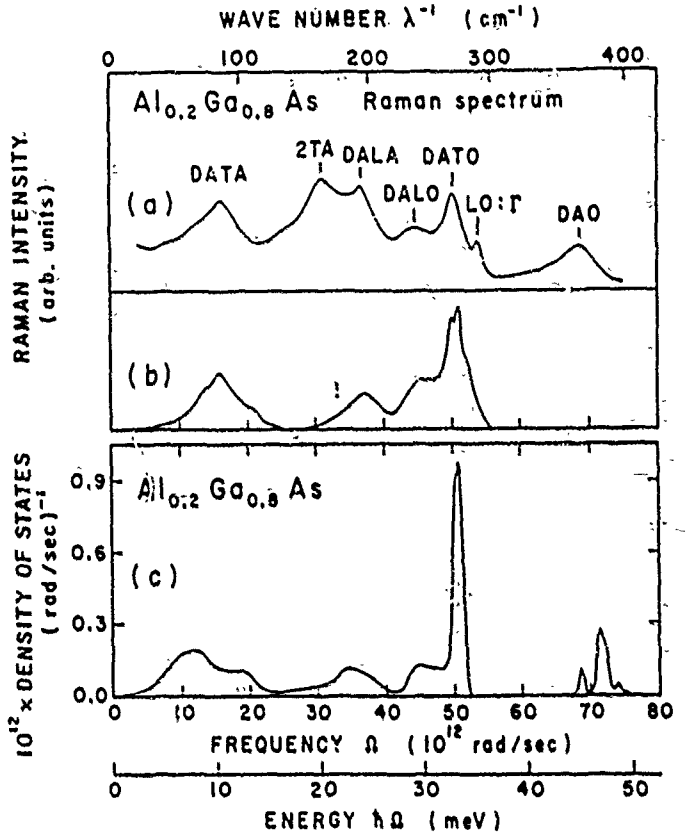


Fig. 5. (a) Raman spectrum for $Al_{0.2}Ga_{0.8}As$ after Ref. [3]; (b) spectrum obtained in Ref. [3] by removing the background, the two-phonon contribution, and the DAO (Disorder-Activated Optic mode); (c) calculated density of phonon states for the same alloy composition. DA indicates "disorder-activated." The mode identifications in (a) are those of Ref. [3].

persistence approximation, Eq. (1). The recursion-method theory produces these modes and identifies each of them with the vibrations of a specific atom or cluster of atoms. These identifications appear to provide improved assignments of some Raman spectral lines. The success of the theory in this regard, namely in describing the corrections to the persistence limit, is what makes the recursion method a satisfactory theory of alloys. However the present calculations should be regarded as only a first semi-quantitative attempt to describe phonons in III-V pseudo-binary alloys. The theory should be improved by including long-ranged forces, especially Coulombic forces and the effect of the polarizability of the alloy constituents.

Acknowledgments — We are grateful to the Department of Energy, Division of Materials Sciences (Contract No. DE-AC02-76-ER00198) for their support. One of us (J.D.D.) gratefully acknowledges the support of the Office of Naval Research (Contract No. N00014-84-K-0352).

APPENDIX: Probabilities

The probabilities of occurrence of five-atom miniclusters are calculated assuming a random pseudo-binary alloy in which all anion sites are occupied by As atoms and the Al and Ga atoms are randomly distributed on the cation sites. For a minicluster with a central Al or Ga atom, the four neighboring atoms are As atoms. For a central As atom, the probability of m Al and n Ga atoms on the neighboring sites is $x^m(1-x)^n$, and there are $4!/(n!m!)$ such clusters.

REFERENCES

[1] R. Tsu, H. Kawamura, and L. Esaki, Proc. International Conf. Phys. of Semiconductors, Warsaw (Elsevier, Amsterdam, 1972), Vol. 2, P. 1135. Also, H. Kawamura, R. Tsu, and L. Esaki, Phys. Rev. Lett. 29, 1397 (1972).

[2] O. K. Kim and W. G. Spitzer, J. Appl. Phys. 50, 4362 (1979).

[3] N. Saint-Cricq, R. Carles, J. B. Renucci, A. Zwick, and M. A. Renucci, Solid State Commun. 39, 1137 (1981). Also, R. Carles, A. Zwick, M. A. Renucci, and J. B. Renucci, Solid State Commun. 41, 557 (1982).
[4] B. Jusserand and J. Sapriel, Phys. Rev. B24, 7194 (1981).
[5] Y. Onodera and Y. Toyozawa, J. Phys. Soc. of Japan 24, 341 (1968).
[6] R. Haydock, in Solid State Physics, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), Vol. 35, p. 215 and references therein.
[7] C. M. M. Nex, J. Phys. A 11, 653 (1978); Computer Phys. Commun. 34, 101 (1984).
[8] M. J. Kelly, in Solid State Physics, ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), Vol. 35, p. 296.
[9] V. Heine, in Solid State Physics, ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), Vol. 35, p. 1.
[10] R. Bonneville, Phys. Rev. B29, 907 (1984).
[11] R. Banerjee and Y. P. Varshni, Can. J. Phys. 47, 451 (1969).
[12] C. Colvard, R. Merlin, M. V. Klein, and A. C. Gossard, Phys. Rev. Lett. 45, 298 (1980) and private communication.
[13] M. Ilegems and G. L. Pearson, Phys. Rev. B1, 1576 (1970).
[14] A. Onton and R. J. Chicotka, Phys. Rev. B10, 591 (1974).
[15] A. S. Barker, Jr., J. L. Merz, and A. C. Gossard, Phys. Rev. B17, 3181 (1978).
[16] The main symmetry points in k-space are $\Gamma=(0,0,0)$, $L=(2\pi/a_L)(1/2,1/2,1/2)$, $X=(2\pi/a_L)(1,0,0)$, $W=(2\pi/a_L)(1,1/4,1/4)$, and $K=(2\pi/a_L)(3/4,3/4,0)$, where a_L is the lattice constant.
[17] H. M. J. Smith, Phil. Trans. Roy. Soc. (London) A241, 105 (1948).
[18] J. L. T. Waugh and G. Dolling, Phys. Rev. 132, 2410 (1963).
[19] C. Herscovici and M. Fibich, J. Phys. C13, 1635 (1980).
[20] A. Kobayashi, Ph. D. thesis, University of Illinois, 1985.
[21] G. Lehmann and M. Taut, phys. stat. sol. (b)54, 469 (1972).
[22] L. van Hove, Phys. Rev. 89, 1189 (1953).
[23] I. F. Chang and S. S. Mitra, Phys. Rev. 172, 924 (1968); Phys. Rev. B2, 1215 (1970).
[24] Y. S. Chen, W. Shockley, and G. L. Pearson, Phys. Rev. 151, 648 (1966).
[25] I. F. Chang and S. S. Mitra, Adv. Phys. 20, 359 (1971).

Theory of charge-state splittings of deep levels associated with sulfur pairs in Si

Gyungock Kim* and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

Seongbok Lee
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
and Department of Physics, University of Texas at Austin, Austin, Texas 78712†

(Received 9 July 1987; revised manuscript received 12 January 1988)

We report self-consistent calculations for the deep levels of nearest-neighbor (S,S) substitutional defect pairs in Si, including (i) the charge-state splitting, and (ii) ratios of hyperfine tensor components. The good agreement with available data lends credence to the mesobonding model of paired-chalcogen deep levels.

I. INTRODUCTION

A controversy exists concerning the character of the deep-level wave functions of chalcogen nearest-neighbor substitutional pairs in Si. On the one hand, Sankey and Dow[1,2] and Hu *et al.*[3] claim that the "molecular" wave function for the (S,S) deep level in the gap of Si is *mesobonding*, that is, a σ-like a_1-symmetric[4] *bonding linear combination* of the two antibonding p-like T_2-symmetric isolated-S wave functions. [These T_2 states of isolated S are not deep in the gap of Si, but are resonant with the conduction band; the isolated-S states in the gap are known to be of A_1 symmetry and of antibonding character, but descend into the valence band for (S,S) pairs.[3,5–12]] On the other hand, Wörner and Schrimer[13] claim that the corresponding wave function for (Se,Se) is *totally antibonding*: an antibonding linear combination of antibonding isolated-Se s-like A_1-symmetric deep-level wave functions. In the Wörner model, the A_1-derived state of the molecule lies within the fundamental band gap.[14] [Both of the theories indicate that the deep levels of (S,S) and (Se,Se) should have similar wave functions.]

In this paper we report self-consistent calculations of the wave functions and energies for $(S,S)^+$ and $(S,S)^0$ deep levels in Si. Our results lend support to the mesobonding model and complement other self-consistent theories of extended defects in semiconductors.[15]

II. METHOD

Our approach is based on the Hjalmarson *et al.* theory of deep impurity levels,[16] and solves the secular equation

$$\det[1-G_0(E)V]=0\,, \tag{1}$$

where V is the defect potential operator and $G_0(E)$ is the bulk-Si Green's function

$$G_0(E)=(E-H_0)^{-1}=\sum_{\mathbf{k},\lambda}\frac{|\mathbf{k},\lambda\rangle\langle\mathbf{k},\lambda|}{E-E(\mathbf{k},\lambda)}\,. \tag{2}$$

The Green's function is evaluated using the sp^3s^* empirical tight-binding model of Vogl *et al.*,[17] by simply finding the eigenvalues $E(\mathbf{k},\lambda)$ and eigenvectors $|\mathbf{k},\lambda\rangle$ of the Hamiltonian.[18] The defect potential $V\equiv H-H_0$ is constructed along the general lines suggested by Hjalmarson *et al.*:[16] In a localized Löwdin-orbital basis of s, p_x, p_y, and p_z orbitals the defect-potential matrix is a 32×32 matrix involving only the two S sites and the six neighboring Si sites directly bonded to S atoms. Following established approximations,[16] we neglect the long-ranged non-central-cell Coulombic parts of the defect potential and ignore lattice relaxation around the impurities—approximations that introduce uncertainties of a few tenths of an eV into the predicted absolute deep level energies.[16,17] However, for studies of charge-state splittings, which are *differences* in deep-level energies, these theoretical uncertainties largely cancel, leading to a theory of the energy difference much more accurate than 0.1 eV.

The symmetry group of a nearest-neighbor substitutional (S,S) pair oriented along the (1,1,1) crystal axis is C_{3v} (Ref. 4) (versus T_d for isolated S). The irreducible representations are a_1 (σ-like), a_2 (rotation about the S—S bond), and e (π-like). Thus the 32×32 secular determinant factors into one 10×10, one 2×2, and one 20×20 determinant, each with the form

$$\det(1-G_0V)=0\,, \tag{3}$$

but involving matrix elments between the basis functions of a single irreducible representation only. Details of the calculation and explicit expressions for the basis functions may be found elsewhere.[19]

The defect potential is diagonal in a Löwdin-orbital basis for the perturbed crystal, with the values

$$\langle i,\sigma;\mathrm{R}\,|\,V\,|\,i,\sigma;\mathrm{R}\rangle=\beta_i[E(i,\sigma;\mathrm{R})-W(i,\sigma;\mathrm{R})] \tag{4}$$

where $\beta_i=0.8$ for $i=s$ and $\beta_i=0.6$ for $i=p_x$, p_y, or p_z.[16] Here R labels the atom and σ denotes spin; $W(i,s;\mathrm{R})$ is the value of $E(i,\sigma;\mathrm{R})$ for bulk Si.

The defect potential is determined self-consistently us-

 ©1989 The American Physical Society

TABLE I. Computed occupation numbers $n(i,\sigma;\mathbf{R})$ for the (S,S) and isolated S defects in Si.

Occupation number	$(S,S)^0$	$(S,S)^+$	$(S)^0$	$(S)^+$
$\sum_{\sigma} n(s,\sigma;\mathbf{R}_0)$	2.01	1.96	1.97	1.92
$\sum_{\mu,\sigma} n(p_\mu,\sigma;\mathbf{R}_0)$	3.86	3.86	3.84	3.87
$\sum_{\sigma} n(s,\sigma;\mathbf{R}_2)$	1.26	1.26	1.25	1.30
$\sum_{\mu,\sigma} n(p_\mu,\sigma;\mathbf{R}_2)$	2.65	2.60	2.64	2.57

ing the scheme of Haldane and Anderson.[20] The atomic-orbital energies are

$$E(s,\sigma)=E_s^0+U_{ss}\sum_{\sigma}{}'\, n(s,\sigma')+U_{sp}\sum_{\mu,\sigma'} n(p_\mu,\sigma') \tag{5a}$$

and

$$E(p_\nu,\sigma)=E_p^0+U_{pp}\sum_{\mu,\sigma}{}'\, n(p_\mu,\sigma')+U_{sp}\sum_{\sigma'} n(s,\sigma')\,, \tag{5b}$$

where ν and μ run over x, y, and z, and the prime on the summations means that terms with $\sigma=\sigma'$ and $\nu=\mu$ are excluded. The occupation numbers n are either unity or zero in atoms. The Coulomb integrals U_{ss}, U_{sp}, and U_{pp}, together with the orbital energies E_s^0 and E_p^0, have been determined for Si and S by fitting atomic spectra.[7]

In the solid, the occupation numbers at each site, $n(i,\sigma;\mathbf{R})$, are allowed to assume nonintegral values, because charge is more delocalized then in an atom. At zero temperature, they are expressed in terms of the defect-perturbed Green's function $G\equiv G_0(1-G_0V)^{-1}$:

$$n(i,\sigma;\mathbf{R})=(-1/\pi)\int_{-\infty}^{E_v}\mathrm{Im}\langle i,\sigma,\mathbf{R}\,|\,G(E)\,|\,i,\sigma,\mathbf{R}\rangle + \sum_j |\langle i,\sigma,\mathbf{R}\,|\,\Psi_j\rangle|^2 \tag{6}$$

where E_v is the valence-band maximum and the summation runs over the occupied bound states in the band gap. Here the bound states are normalized as follows:[21]

$$\langle\Psi_j\,|\,V(d/dE)[G_0(E)]V\,|\,\Psi_j\rangle=-1\,. \tag{7}$$

Since the occupation numbers n depend on the defect potential V (and vice versa), Eqs. (3), (4), and (6) are solved iteratively until self-consistency is obtained. This is done for the deep levels of both $(S,S)^+$ and $(S,S)^0$, the difference between the energies of the two deep levels being the charge-state splitting.

III. RESULTS

A. Occupation numbers

The computed occupation numbers for a S site ($\mathbf{R}_0$) and a Si site ($\mathbf{R}_2$) are given in Table I. They add to (almost) six electrons per S atom and four per Si atom, indicating that the atoms are all nearly neutral in their central cells, regardless of the charge state of the defect: the extra electron of $(S,S)^0$ is rather delocalized and spends very little time in the central cells. A similar situation held for isolated S in Si (Ref. 6) that has virtually the same occupation numbers $n(i,\sigma,\mathbf{R})$ as (S,S), despite the fact that the deep-level wave function is qualitatively different—being A_1 in character rather than T_2 derived. (See Table I.)

B. Comparison with data

The principal results of the calculation, as they relate to data, are displayed in Table II. These include the charge-state splittings $E[(S,S)^0]-E[(S,S)^+]$ and $E[(S)^0]-E[(S)^+]$, respectively, and various measured functions of the hyperfine tensor's components, $A_\parallel$ and $A_\perp$ for $(S,S)^+$ and $B_\parallel$ and $B_\perp$ for $(S)^+$, as computed in Refs. 1, 2, and 5. Note that all of these quantities are in excellent agreement with the data. The predicted T_2-derived deep level of $(S,S)^0$ lies a few tenths of an eV below the measured level, a result attributable to the s^*-p coupling of the simplified Vogl tight-binding model, which is well known to push the indirect conduction-band minima of Si down to their experimental energy, while simultaneously depressing the p-like T_2-derived

TABLE II. Comparison of theory with data for the charge-state splittings $E[(S,S)^0]-E[(S,S)^+]$ and $E[(S)^0]-E[(S)^+]$ for (S,S) and isolated-S deep levels in Si, absolute energy $E[(S,S)^0]$ of the neutral (S,S) deep level (with respect to the valence-band maximum), and various ratios of the principal values for the hyperfine tensor A of $(S,S)^+$ and B of $(S)^+$. The data for $A_\perp/A_\parallel$ and $(A_\parallel-A_\perp)/(A_\parallel+2A_\perp)$ are for $(Se,Se)^+$ because we were unable to find comparable data for $(S,S)^+$. This substitution of Se data for S data is justified because the S and Se deep levels in Si are known to be almost the same (Refs. 2 and 5): Both defects' wave functions are Si-dangling-bond-like (Ref. 16).

	Theory	Data
$E[(S,S)^0]-E[(S,S)^+]$	0.19 eV	0.18 eV[a]
$E[(S)^0]-E[(S)^+]$	0.27 eV	0.30 eV[a]
$E[(S,S)^0]$	0.44 eV	0.98 eV[a]
$A_\parallel/B_\parallel$	0.35	0.37[b]
$A_\perp/A_\parallel$	0.95	0.97 for $(Se,Se)^+$[b]
$(A_\parallel-A_\perp)/(A_\parallel+2A_\perp)$	0.016	0.01 for $(Se,Se)^+$[b]

[a]References 2 and 9.
[b]Reference 13.

deep levels a bit more than they should be.[22] This artifact of the model can be corrected by adjusting the strength of the s^*-p coupling until the theoretical isolated-$(S)^0$ deep level lies at the experimental value of 0.86 eV—in which case the other quantities of Table II are changed very little and the theoretical $(S,S)^0$ deep level is found to lie at the experimental energy of 0.98 eV to within 1%.

Our calculated charge distribution for the $(S,S)^+$ deep level is given in Table III and is essentially the same as that found by Sankey and Dow[1,2] for $(S,S)^+$ using a non-self-consistent theory which produced remarkable agreement with data. Likewise our self-consistent calculations for the deep level of isolated $(S)^+$ (Ref. 6) agree with an earlier non-self-consistent calculation by Ren *et al.*[5] Thus our theoretical predictions for $(S,S)^+$ and $(S)^+$ are in agreement with both previous theory and the data.[8,10–13]

TABLE III. Computed electronic charge distribution of the a_1-symmetric mesobonding deep-level state in the fundamental band gap of Si, associated with nearest-neighbor substitutional sulfur-pair impurities $(S,S)^+$. The index μ runs over x, y, and z; the index R_i runs over R_0 and R_1, that is, over the sulfur sites; and R_j runs over R_2, R_3, R_4, R_5, R_6, and R_7: the first-neighbor silicon atoms to each S impurity.

State	Electron charge (%)
$\sum \lvert\langle\Psi \lvert s,R_i(S)\rangle\rvert^2/2$	3.4
$\sum_{j,\mu} \lvert\langle\Psi \lvert p_\mu,R_j(S)\rangle\rvert^2/2$	1.9
$\sum_{i,\mu} \lvert\langle\Psi \lvert s,R_i(Si)\rangle\rvert^2/6$	1.3
$\sum_{j,\mu} \lvert\langle\Psi \lvert p_\mu,R_j(Si)\rangle\rvert^2/6$	7.7

IV. CONCLUSIONS

The quantitative agreement between the self-consistent theory and the data is a strong indication that the mesobonding model of paired-chalcogen deep levels in Si is correct. Electron-nuclear double resonance measurements of the charge distribution of $(S,S)^+$, if they produce similar distributions to those of Table III, will provide final resolution of the mesobonding versus totally antibonding controversy.

ACKNOWLEDGMENTS

The authors wish to thank E. Yamaguchi for helpful discussions. We are grateful to the U.S. Office of Naval Research for generous support (Contract No. N00014-84-0352), and one of us (J.D.D.) thanks the Minnesota Supercomputer Institute for support.

*Present address: HHMI, NSLS-X4, BNL, Upton, NY 11973.
†Present address.

[1]O. F. Sankey and J. D. Dow, Phys. Rev. B 26, 3243 (1982).
[2]O. F. Sankey and J. D. Dow, Solid State Commun. 51, 705 (1984).
[3]W. M. Hu, Y. M. Gu, and S. Y. Ren, Acta Phys. Sin. 35, 1583 (1986).
[4]We use the group-theoretical notation of M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964), with upper case letters for isolated S and lower case referring to (S,S) pairs.
[5]S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).
[6]S. Lee, J. D. Dow, and O. F. Sankey, Phys. Rev. B 31, 3910 (1985). The numerical results of the present work supercede those of this earlier work, although the essential physics is unchanged: here we find the A_1 deep level of $(S)^0$ at 0.28 eV above the valence-band maximum, with a charge-state splitting of 0.27 eV. Here we assumed $\beta_s=0.8$ and $\beta_p=0.6$, and a Si band gap of 1.2 eV. The computed occupation numbers n are given in Table I. The wave-function coefficients γ_s (Ref. 2) for $(S)^0$ and $(S)^+$ are 0.30 and 0.31, respectively, compared with 0.32 found by Ren *et al.* (Ref. 5). We have 9.6% of the $(S)^+$ deep level's charge in the central cell [versus the experimental finding of ~10% (Refs. 10 and 12)], and we find that each of the four nearest-neighbor Si hybrids has 1.7% of the charge in an s orbital and 13.8% in a p orbital. This means that the nearest-neighbor hybrid is predicted to be 89% p-like, 13.8/(1.7 + 13.8), whereas electron-spin-resonance and electron-nuclear double resonance experiments find it to be 91% p-like (Refs. 10 and 12).
[7]O. F. Sankey and J. D. Dow, Phys. Rev. B 27, 7641 (1983).
[8]G. W. Ludwig, Phys. Rev. 137, A1520 (1965).
[9]S. D. Brotherton, M. J. King, and G. J. Parker, J. Appl. Phys. 52, 4649 (1981).
[10]H. G. Grimmeiss, E. Janzen, H. Eunen, O. Shirmer, J. Schneider, R. Wörner, C. Holm, E. Sirrl, and P. Wagner, Phys. Rev. B 24, 4571 (1981); H. G. Grimmeiss, E. Janzen, and K. Larsson, *ibid.* 25, 2627 (1981).
[11]G. D. Watkins and J. W. Corbett, Phys. Rev. 134, A1359 (1964).
[12]J. R. Niklas and J. M. Spaeth, Solid State Commun. 46, 121 (1983).
[13]R. Wörner and O. F. Schirmer, Solid State Commun. 51, 665 (1984).
[14]This is in contrast to Refs. 1 and 2, which find the totally antibonding deep level of (S,S) in the valence band.
[15]R. Car, P. Kelly, A. Oshiyama, and S. Pantelides, Phys. Rev. Lett. 52, 1814 (1984); G. A. Baraff and M. Schlüter, Phys. Rev B 33, 7346 (1986).
[16]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980); see also, W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977) for a discussion of the concepts that form the foundation for this work.
[17]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983). In the present paper, we use the notation T_i instead of the original notation V for the off-diagonal matrix elements.
[18]J. D. Dow, in *Highlights of Condensed-Matter Theory*, Proceedings of the International School of Physics "Enrico Fermi," Course 89, Varenna, 1983, edited by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna, Italy, and North-Holland, Amsterdam, 1985), pp. 465–494.

[19]G. Kim, J. D. Dow, and S. Lee, Arabian J. Sci. Eng. 14, 4 (1989); G. Kim, Ph.D. thesis, University of Notre Dame, 1988 (unpublished).

[20]F. D. M. Haldane and P. W. Anderson, Phys. Rev B 13, 2553 (1976); P. W. Anderson, *ibid.* 124, 41 (1961); P. Vogl and J. Baranowski, Acta Phys. Pol. A 67, 133 (1985).

[21]G. A. Baraff and M. Schlüter, Phys. Rev. B 19, 4965 (1979); 30, 1853 (1984).

[22]H. P. Hjalmarson, Ph.D. thesis, University of Illinois, 1979 (unpublished).

Solid State Communications, Vol.56,No.7, pp.553-556, 1985.
Printed in Great Britain.

0038-1098/85 $3.00 + .00
Pergamon Press Ltd.

EXPLANATION OF THE ANOMALOUS RAMAN SPECTRA OF $(GaSb)_{1-x}Ge_{2x}$

Kathie E. Newman and John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556, USA

and

Akiko Kobayashi*
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556, USA
and
Department of Physics and Materials Research Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801, USA

and

Robert Beserman
Technion, Israel Institute of Technology
Technion City, Haifa 32 000, Israel

(Received 16 May 1985 by A. A. Maradudin)

The following anomalies in the $(GaSb)_{1-x}Ge_{2x}$ Raman data are simply explained using an alloy theory and an order-disorder phase-transition model: (i) the discontinuous dependence on x of the Ge-like LO mode, (ii) the anomalous LO-TO splitting, (iii) the discontinuous change in the derivative with respect to x of the GaSb-like LO line width, (iv) the maximum as a function of x of the Ge-like LO line width, and (v) the asymmetries of the LO lines.

In this communication we report calculations of the densities of vibrational states for metastable, substitutional, crystalline $(GaSb)_{1-x}Ge_{2x}$ alloys [1,2] and provide explanations of the following anomalies in the Raman data [3] for these alloys: (i) the discontinuous dependence on x of the Ge-like longitudinal optic (LO) mode frequency Ω, (ii) the existence of a discernible splitting between the long-wavelength ($\vec{k}=\vec{0}$) longitudinal optic and transverse optic (TO) modes for $x<0.3$ but not for $x>0.3$, (iii) the GaSb-like LO-mode width W' that increases with x and exhibits a discontinuous change in dW'/dx at $x=0.3$, (iv) the Ge LO-mode width W that, as a function of x, is maximum for $x=0.3$ and has a discontinuous derivative dW/dx near $x=0.3$, and (v) asymmetries of the LO modes that are not smooth functions of x.

The calculations are based on a Born von Karman rigid-ion [4] model of the lattice dynamics of GaSb and Ge. The alloy fluctuations are incorporated using the recursion method [5] (executed to 51 levels for 1000 atoms), with the distributions of atoms on the nominal anion and cation sites of a zincblende lattice being governed by the zincblende-diamond phase-transition model with a critical alloy composition $x_c=0.3$ [6,7], as evaluated in a mean-field approximation [8]. To facilitate the calculations, interatomic forces with ranges longer than the second-nearest-neighbor distance are omitted from the densities of states evaluation. The calculations (except for the phase-transition aspects) are similar to those reported for $Al_xGa_{1-x}As$ [9]; details of the calculational method and tables of the relevant force constants are available [10,11]. The output of the calculation is an ensemble-averaged approximation to the density of vibrational states per unit cell

$$D(\Omega) = N^{-1} \Sigma_s \delta(\Omega-\Omega_s),$$

where N is the number of unit cells and the sum is over all eigenmodes with eigenvalues Ω_s. The results for $(GaSb)_{1-x}Ge_{2x}$ are displayed in Fig. 1, for various values of the composition x and the order parameter M of the phase transition, which is proportional to the average electric dipole moment per unit cell. (Recall that in this model the LO and TO modes at $\vec{k}=\vec{0}$ are degenerate.) Many features of the alloy density of states spectra can be associated with the densities of states of either GaSb or Ge: for $x=0.5$ we compare the average of the GaSb and Ge spectra with the alloy density of states. In Fig. 1, we associate deviations from this average with "alloy modes" [9] and identify the major spectral features with the vibrations of specific bonds in the alloy.

Especially noteworthy is the fact that the Ge optic mode, which is due mostly to vibrations of Ge atoms and Ge-Ge bonds, has a low-energy Ga-Ge sideband that is discernible in Fig. 1 for $x<0.5$. The calculated splitting between the Ga-

* Current address: Department of Physics and Astronomy, University of Maryland, College Park, MD 20742.

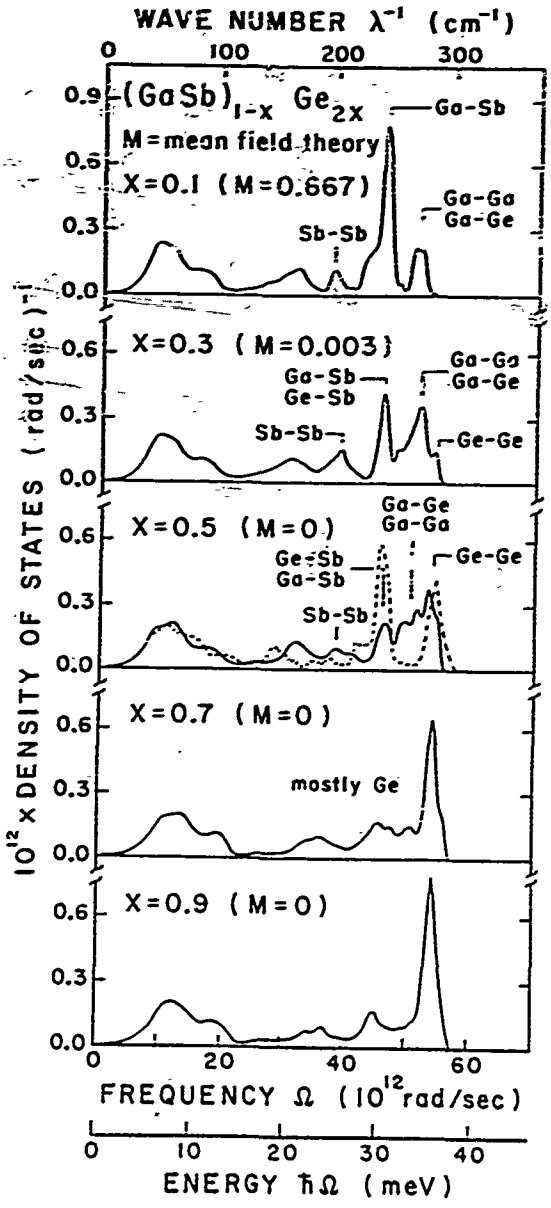


Fig. 1. Calculated densities of phonon states for $(GaSb)_{1-x}Ge_{2x}$ alloys, assuming the phase-transition model [7] with an order parameter M (solid lines). The dashed line is the average of the GaSb and Ge densities of states. Various peaks are associated with specific bonds.

Ge and Ge-Ge spectral features is $\approx 7 cm^{-1}$ [12]. Using the calculated splitting and the known LO $\vec{k}=\vec{0}$ mode frequencies in Ge and GaSb [13], we find an explanation of the anomalous discontinuity in the Ge-like LO-mode position (Fig. 2): Both the Ge-Ge and Ga-Ge vibrations coexist within the "Ge-like LO peak" at slightly different frequencies Ω; in Ge-rich material the Ge-Ge peak is prominent, but in GaSb-rich material the Ga-Ge feature dominates the "Ge-like LO mode." Hence the discontinuity in the plot of the peak position versus x is due to switching of the most prominent spectral feature from Ga-Ge to Ge-Ge.

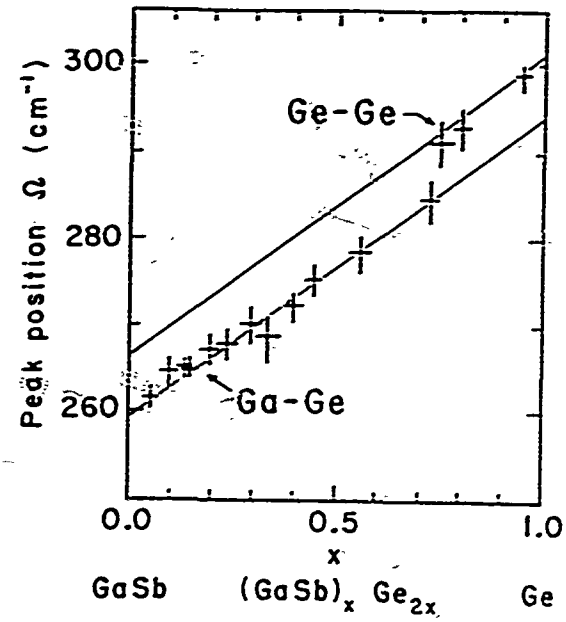


Fig. 2. Peak positions Ω (in cm^{-1}) of the Ge-like LO modes versus alloy composition x, after Ref. [3] (data points). The parallel lines through the data for the LO modes associated with the Ge-Ge and Ga-Ge bonds are separated by the calculated amount, $\approx 7 cm^{-1}$ [12], with the Ge-Ge and Ga-Ge end points fit to the observed LO-mode frequencies.

Another feature of the data [3] is the apparent absence of a discernible LO-TO splitting [14] of the $\vec{k}=\vec{0}$ GaSb-like modes for x>0.3, although the breadth of the principal spectral features may prevent observation of small splittings. Nevertheless, the phase-transition model [7] offers a simple explanation of this phenomenon: The net average dipole moment per unit cell vanishes at the critical composition $x_c=0.3$; as a result the LO-TO splitting, which in a perfect crystal is proportional to the square of the dipole moment [14,15], likewise vanishes [16]. In Fig. 3 we plot the theoretical prediction for the optic mode obtained from the alloy theory of Fig. 1 (with no long-range forces), shifted slightly [17], and split a posteriori by an amount proportional to the square of the order parameter M [18], to account approximately for the charge transfer among Ga, Sb, and Ge atoms. These simple ideas seem to provide a pleasing explanation of the data [13].

The total width W of the GaSb-like mode increases as a function of x, with dW/dx changing discontinuously at x=0.3 [3]. This is symptomatic of a phase transition with $x_c=0.3$: for $x>x_c$ the distinction between anion and cation sites is lost, which is reflected in the GaSb-like LO-phonon line width. The continued increase in the apparent width of this broad feature as x increases results from the GaSb mode merging with the Ge-Sb features of the spectrum, as seen in Fig. 1.

The Ge-like optic mode lies at high energy and is discernible for x>0; hence it does not merge with a broad continuum of other modes as the GaSb-like mode does. Its width is maximum at

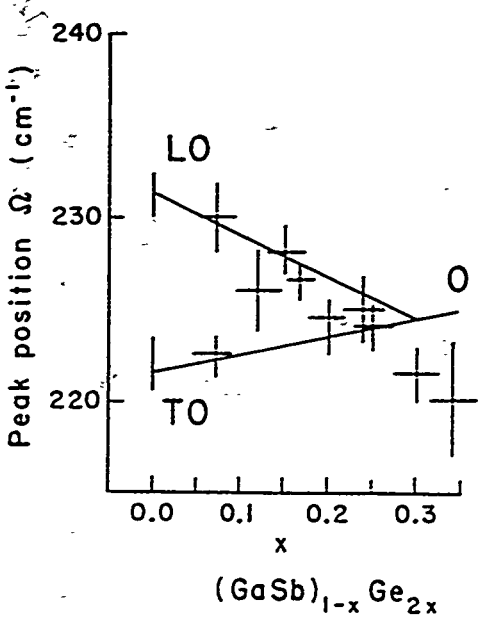


Fig. 3. The peak positions Ω (in cm^{-1}) of the Raman lines for the GaSb-like LO and TO modes versus alloy composition x. The solid line is the shifted [17] theory with a Lyddane-Sachs-Teller splitting [14] proportional to the square of the order parameter of the phase-transition theory [13].

x=0.3. The total width is due to many different modes (including, e.g., modes associated with Ge-Ge, Ga-Ge, and Ga-Ga bonds) and reflects the disorder in the alloy. Qualitatively this width should be a maximum (minimum) when the disorder is a maximum (minimum). Since the entropy is a measure of disorder, in Fig. 4 we compare the

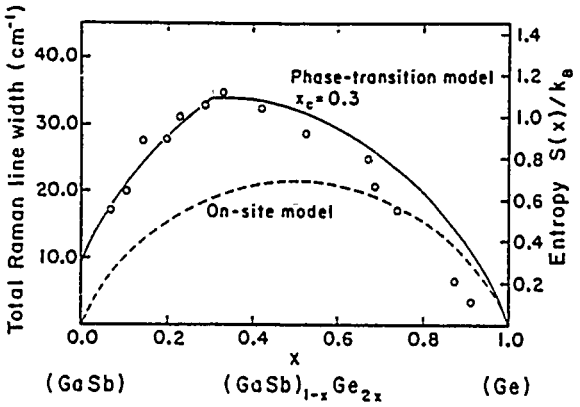


Fig. 4. Total Raman line width (in cm^{-1}) versus composition x for the Ge-like LO mode in $(GaSb)_{1-x}Ge_{2x}$. Data are indicated by circles. The solid line is the entropy per site $S(x)/k_B$ (divided by Boltzmann's constant) evaluated using the phase-transition theory [7] and the dashed line is the theory for the on-site model, which does not allow Sb-Sb or Ga-Ga bonds. The right-hand scale is for $S(x)/k_B$. Note that the maximum of S(x) coincides with the maximum line width in the phase-transition model, but not in the on-site model.

entropy of the phase-transition model (for x_c=0.3) [7] with the Raman line width. The entropy per site (in a mean-field approximation) is [7]:

$$S(x)/k_B = [(1-x+M)/2]\ln[(1-x+M)/2] + x\ln x + [(1-x-M)/2]\ln[(1-x-M)/2],$$

where M is the order parameter of the transition [7] and k_B is Boltzmann's constant. The agreement between the theory and the data is gratifying.

The peak of the entropy at x=0.3 is indicative of the order-disorder zincblende-diamond phase transition with critical composition x_c=0.3. A feature of this phase-transition model is that Sb atoms may occupy nominal cation sites, and vice versa — leading to Sb-Sb and Ga-Ga bonds. If the Sb atoms are constrained to occupy only anion sites and the Ga atoms reside only on cation sites (the "on-site model"), then the theory predicts an entropy that has a maximum at x=0.5 (See Fig. 4) [10]. Thus the data for the total line width are consistent with the phase-transition model and inconsistent with the on-site model.

This raises the issue of why the data do not exhibit prominent spectral features associated with Sb-Sb and Ga-Ga bonds (since the appreciable number of antisite Sb and Ga atoms predicted by the theory produce such bonds). In particular, a strong Sb-Sb peak had been expected near 193 cm^{-1} [3]. In fact, the theory (Fig. 1) shows that the Sb-Sb features are weak even in the mean-field approximation to the phase transition, and they are likely to be weaker yet in a theory that allows for atom-atom correlations (and hence has fewer Sb-Sb bonds [19]). Thus Fig. 1 suggests that the Sb-Sb bonds merely contribute to a relatively broad and featureless shoulder of the GaSb-like LO line, rather than produce a prominent peak in the spectrum.

Finally, we note that the asymmetries of the GaSb-like and Ge-like LO lines have been reported to have anomalous dependences on composition x, with discontinuous derivatives [3]. This is symptomatic of the fact that the lines are composed of features associated with several different modes: the different modes have changing importance in the overall line as x varies, and this manifests itself as an anomalous dependence of the overall line asymmetry on composition.

In summary, the main features of the Raman data for $(GaSb)_{1-x}Ge_{2x}$, once thought to be mysterious and anomalous, are naturally attributed to the combined effects of alloy fluctuations and an order-disorder zincblende-diamond phase transition [20].

Acknowledgment — We are grateful to the Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support. During part of this work one of us (A. K.) was supported by the Department of Energy, Division of Materials Sciences (Contract No. DE-AC02-76-ER01198). We are especially grateful to J. E. Greene, M. V. Klein, T. N. Krabach, and T. C. McGlinn for stimulating discussions about their measurements.

REFERENCES

[1] J. E. Greene, J. Vac. Sci. Technol. B1, 229 (1983); J. L. Zilko and J. E. Greene, J. Appl. Phys. 51, 1549 (1980); 51, 1560 (1980). S. A. Barnett, B. Kramer, L. T. Romano, S. I. Shah, M. A. Ray, S. Fang; J. E. Greene, Layered Structures, Epitaxy, and Interfaces, ed. J. M. Gibson and L. R. Dawson (North Holland Publishing Co., Amsterdam, 1984).

[2] Z. I. Alferov, R. S. Vartanyan, V. I. Korol'kov, I. I. Mokan, V. P. Ulin, B. S. Yavich, and A. A. Yakovenko, Fiz. Tekh. Poluprovodn. 16, 887 (1982) [English transl.: Sov. Phys. Semiconductors 16, 567 (1982)] have grown $(GaAs)_{1-x}Ge_{2x}$ alloys.

[3] R. Beserman, J. E. Greene, M. V. Klein, T. N. Krabach, T. C. McGlinn, L. T. Romano, and S. I. Shah, Proc. 17-th Internatl. Conf. Phys. Semiconductors, San Francisco (Springer-Verlag), 1985). See also T. N. Krabach, N. Wada, M. V. Klein, K. C. Cadien, and J. E. Greene, Solid State Commun. 45, 895 (1983).

[4] R. Banerjee and Y. P. Varshni, Can. J. Phys. 47, 451 (1969). The long-range force parameter "x" of this model is taken to be zero here, in order to facilitate the calculations for the alloy.

[5] R. Haydock, in Solid State Physics, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), Vol. 35, p. 215 and references therein; C. M. M. Nex, J. Phys. A 11, 653 (1978); Computer Phys. Commun. 34, 101 (1984); M. J. Kelly, in Solid State Physics, ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), Vol. 35, p. 296; V. Heine, in Solid State Physics, ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), Vol. 35, p. 1.

[6] K. E. Newman, A. Lastras-Martinez, B. Kramer, S. A. Barnett, M. A. Ray, J. D. Dow, J. E. Greene, and P. M. Raccah, Phys. Rev. Lett. 50, 1466 (1983).

[7] K. E. Newman and J. D. Dow, Phys. Rev. B27, 7495 (1983).

[8] The mean-field theory treats $\vec{k}=\vec{0}$ properties adequately. But the short-range order, to be treated properly, requires a more complete treatment of correlations that will reduce the number of Sb-Sb and Ga-Ga bonds below the mean-field-theory value.

[9] A. Kobayashi, J. D. Dow, and E. P. O'Reilly, Superlattices and Microstructures, to be published.

[10] A. Kobayasshi, K. E. Newman, and J. D. Dow, Phys. Rev. B, to be published.

[11] A. Kobayashi, Ph.D. thesis, University of Illinois, 1985.

[12] This splitting is difficult to estimate for many values of x, but appears to be reasonably independent of x in the present model. The value used here was obtained for x=0.5.

[13] The theory in its present form omits long-range forces and so does not accurately predict the slopes $d\Omega/dx$ of the phonon peak frequencies Ω.

[14] R. Lyddane, R. G. Sachs, and E. Teller, Phys. Rev. 59, 218 (1941).

[15] Since the LO-TO splitting is for $\vec{k}=\vec{0}$, we assume that a virtual-crystal approximation provides an adequate zero-order description of the effect of long-range forces in the alloy.

[16] Local fluctuations of the dipole moment per unit cell may nevertheless allow observation of separate $\vec{k}=\vec{0}$ LO and TO features for $x>x_c$.

[17] The theory has been shifted downward slightly, 4 cm^{-1}, to coincide with the data.

[18] The proportionality constant is fixed by fitting the observed LO-TO splitting in GaSb.

[19] E. A. Stern, F. Ellis, K. Kim, L. Romano, S. I. Shah, and J. E. Greene, Phys. Rev. Lett. 54, 905 (1985) have studied $(GaSb)_{1-x}Ge_{2x}$ alloys grown on amorphous substrates, using EXAFS. These materials have somewhat different growth conditions and hence different vibrational properties from those studied here. B. Koiller, M. A. Davidovitch, and R. Osório, to be published, have recently evaluated the theory of Ref. [7] in a Bethe approximation, and have found fewer As-As bonds for $(GaAs)_{1-x}Ge_{2x}$ than in the mean-field approximation.

[20] The discussion here is in terms of bonds. The same physics occurs in a band picture, but is less easily recognized because it is a consequence of three competing effects on the Ge $\vec{k}=\vec{0}$ LO mode as x decreases from unity: (i) The alloy disorder breaks the $\vec{k}=\vec{0}$ selection rule, permitting $\vec{k}\neq\vec{0}$ Ge LO modes (forbidden for x=1) to appear in the Raman data, initially broadening the line; (ii) The width of the disorder-activated Ge-like LO band decreases as the Ge atoms become increasingly isolated from other Ge atoms; and (iii) The alloy disorder itself gives rise to new sideband modes, such as the Ga-Ge modes, which broaden the main Raman line inhomogeneously. For a complete discussion, see Ref. [10].

Effects of a zinc-blende–diamond order-disorder transition on the crystal, electronic, and vibrational structures of metastable $(GaAs)_{1-x}(Ge_2)_x$ alloys

Kathie E. Newman, John D. Dow, and Bruce A. Bunker
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46555

L. L. Abels, P. M. Raccah, S. Ugur, and Da Zhong Xue
Department of Physics, University of Illinois at Chicago, Box 4348, Chicago, Illinois 60680

Akiko Kobayashi*
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
and Department of Physics, 1110 West Green Street, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801
(Received 14 October 1986; revised 5 October 88)

X-ray diffraction data, ellipsometry data, Raman-scattering data, and theory combine to provide strong evidence of a zinc-blende–diamond order-disorder transition with $x_c \simeq 0.3$ that affects the crystal, electronic, and vibrational structures of $(GaAs)_{1-x}(Ge_2)_x$ metastable, substitutional, crystalline alloys. It is argued, based on analyses of extended x-ray-absorption fine-structure data for $(GaSb)_{1-x}(Ge_2)_x$, that the number of anion-anion bonds in these $(A^{III}\text{-}B^{V})_{1-x}(C_2^{IV})_x$ alloys is significant, and requires a theory that goes beyond the mean-field approximation.

The proposal[1,2] that long-lived, metastable, substitutional crystalline alloys,[3–5] such as $(GaAs)_{1-x}(Ge_2)_x$ and $(GaSb)_{1-x}(Ge_2)_x$, exhibit an order-disorder zinc-blende–diamond phase transition as a function of alloy composition x has stimulated considerable interest and discussion[6–11] and some controversy.[9–11] The original proposal was motivated by data for the direct energy band gap $E_0(x)$ of $(GaAs)_{1-x}(Ge_2)_x$, which showed nonparabolic V-shaped bowing as a function of x, with the minimum of the V at the phase-transition composition x_c, near $x \simeq 0.3$ (Ref. 2). In this paper we report x-ray diffraction data,[4] ellipsometry data, Raman-scattering linewidth data, and theory which together demonstrate, conclusively, the existence of the predicted order-disorder transition and its role in determining the crystal, electronic, and vibrational structures of $(GaAs)_{1-x}(Ge_2)_x$ metastable alloys.

The idea of the phase-transition model[2] is that for small x, the alloy $(GaAs)_{1-x}(Ge_2)_x$ is GaAs-like (with Ge substitutional "impurities"), retaining the zinc-blende crystal structure; but for large x, the alloy is Ge-like (with Ga and As substitutional "impurities"), retaining the diamond structure.[12] The zinc-blende phase has well-defined anion sites occupied almost exclusively by As or Ge atoms and cation sites occupied predominantly by Ga or Ge. For Ge-rich alloys, however, the diamond-structure crystal has Ga atoms occurring on both nominal anion and cation sites since, in the limit of x near unity, there are too few Ga and As atoms to force the Ga atoms all to align on one "cation" sublattice. A phase transition must occur between these limiting extremes, i.e., there must exist an intermediate composition x_c such that for $x > x_c$, the distinction between anion and cation sites is lost, the average net dipole moment per unit cell vanishes, and the crystal structure of the alloy becomes diamond rather than zinc blende. The V-shaped bowing of the direct-gap data indicates that the critical composition is near $x_c \simeq 0.3$.[1,2]

The phase-transition model of the alloys, in its first form, used only the mean-field approximation, and thus the theory contained the well-known problems of that approximation. In particular, as shown by Koiller *et al.*,[6] mean-field theory, which predicts adequately the long-wavelength properties of the $(GaAs)_{1-x}(Ge_2)_x$ alloys, is less reliable at short wavelengths or short distances, and overestimates somewhat the number of antisite defects or As—As bonds to be expected in $(GaAs)_{1-x}(Ge_2)_x$. *Quantitatively*, the calculation of the order parameter $M(x)$, which is proportional to the average electric dipole moment per unit cell,[2] is not correct in either the original mean-field theory or the Koiller *et al.* correlated theory: these mean-field theories predict that $M(x)$ varies as $(x - x_c)^{1/2}$, rather than the currently expected form, $M(x) \sim (x - x_c)^{\beta}$, with β being the critical exponent ($\beta \simeq \frac{1}{3}$, assuming that x_c depends linearly on temperature).[13] In the limit $x \to 0$, $M(x)$ is closer to unity in a correlated theory[6] than in mean-field theory. Although these difficulties may affect some of the detailed quantitative predictions of the mean-field theory (as discussed below), the problems with a mean-field approximation do not affect either the qualitative prediction of the zinc-blende–diamond phase transition near $x_c \simeq 0.3$, or many of the semiquantitative predictions of the long-wavelength properties of these alloys.

One experimental signature of a zinc-blende crystal structure which differentiates it *qualitatively* from a diamond crystal structure is the (200) x-ray diffraction spot,

 ©1989 The American Physical Society

which must vanish abruptly at a zinc-blende–diamond phase transition, because then the distinction between cation and anion sites vanishes. Thus, if the phase-transition model is correct for these alloys, for $x < x_c$ a distinctive (200) x-ray spot should be observed, but for $x > x_c$ such a distinctive spot should be absent. To verify the model, samples of $(GaAs)_{1-x}(Ge_2)_x$ alloy were prepared on GaP substrates and measured by x-ray diffraction.[4]

$(GaAs)_{1-x}(Ge_2)_x$ films on undoped semi-insulating GaP substrates were grown[4] in a multitarget rf sputtering system, the general features of which have been described elsewhere.[14,15] The key to growing single-phase metastable alloys $(GaAs)_{1-x}(Ge_2)_x$ is the use of low-energy Ar-ion bombardment in order to collisionally mix the upper one or two atomic layers of the film during growth. (The substrates were kept at 450–550 °C during growth.) The detailed sample preparation and growth techniques are discussed in Ref. 16. Electron channeling measurements using a JEOL scanning electron microscope verified that the films were single crystals, and film compositions were determined to ±0.5 at. % by a JEOL electron microprobe referenced to Ge and GaAs wafers as standards, using the MAGICIV computer program[17] to make matrix corrections for fluorescence, absorption, and atomic number. The x-ray measurements employed a high-precision triple-crystal diffractometer and a Rigaku RU-200 rotating anode source operating at 55 kV and 100 mA, and emitting 0.7093-Å Mo $K\alpha_1$ radiation. All samples analyzed were (100)-oriented single crystals.

In Fig. 1 we show measurements of the x-ray (200) spot intensities $I_{(200)}$ (appropriately normalized) as functions of alloy composition x for $(GaAs)_{1-x}(Ge_2)_x$ metastable alloys (grown on GaP substrates) and for $(GaSb)_{1-x}(Ge_2)_x$ (grown on GaAs).[4,18] These intensities drop precipitously at the critical compositions x_c and vanish for $x > x_c$: the distinction between anion and cation sites vanishes and the net average electric dipole moment per unit cell (the order parameter of the phase transition) is zero. Similar behavior has been observed by other workers in $(GaAs)_{1-x}(Ge_2)_x$.[19] Clearly, the x-ray diffraction data alone provide unambiguous and compelling evidence that the zinc-blende character of the crystal structure and atomic geometry disappears near $x_c \simeq 0.4$—as implied by the zinc-blende–diamond order-disorder transition theory and previous interpretations of the direct-gap data.[1,2] (The slight difference between the value $x_c \simeq 0.4$ extracted from the x-ray data for samples grown on GaP substrates and $x_c \simeq 0.3$ for the other measurements of samples on GaAs substrates is not significant, because x_c depends on the growth conditions.) The experimental data are also in good qualitative (but not quantitative) agreement with the mean-field calculation of the order parameter, also shown in Fig. 1.[20]

Ellipsometry data show the effects of the order-disorder transition on the electronic structure. The measured energy of the E_1 edge[21] of $(GaAs)_{1-x}(Ge_2)_x$ alloys is displayed as a function of x and compared with the theory (evaluated in a mean-field generalized virtual-crystal approximation[2]) in Fig. 2. (This edge has been resolved from the $E_1 + \Delta$ edge[21] by using Lorentzian fitting, with an estimated experimental error of 50 meV.) The experimental results show a definite kink near $x_c \simeq 0.3$, in good semiquantitative agreement with the theory.[22] Such V-shaped bowing of a band gap as a function of x in an alloy is a general feature of a phase transition (Refs. 2 and 23), and has been observed for the direct band gaps of both $(GaAs)_{1-x}(Ge_2)_x$ (Ref. 1) and $(GaSb)_{1-x}(Ge_2)_x$ (Ref. 24), with minima near $x_c \simeq 0.3$. Both theory[2] and experiment agree that the amount of bowing found for $E_1(x)$ should be of the same order of magnitude but less than that found for the direct gap.[22] The data show the E_1 bowing to be about half of the E_0 bowing.

The phase transition also manifests itself in the vibrational structure of $(GaAs)_{1-x}(Ge_2)_x$, despite the fact that the masses of Ga, As, and Ge atoms are all nearly equal. The disorder of the alloy should influence the *width* of

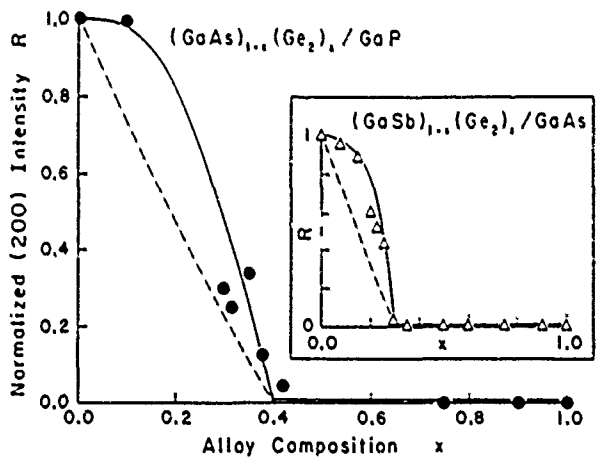


FIG. 1. The measured normalized diffracted x-ray beam intensity ratios of the (200) x-ray diffraction spots of $(GaAs)_{1-x}(Ge_2)_x$ grown on GaP substrates (circles) and $(GaSb)_{1-x}(Ge_2)_x$ grown on GaAs (100) substrates (triangles in inset) vs alloy composition x after Ref. 4. We have $R \equiv (I_{(200)}/I_{(400)})_{alloy}/(I_{(200)}/I_{(400)})_{III\text{-}V}$, where the alloy is $(GaAs)_{1-x}(Ge_2)_x$ or $(GaSb)_{1-x}(Ge_2)_x$ and III-V refers to GaAs or GaSb. In all cases, the intensities $I_{(200)}$ are measured relative to the (400) beam intensities, since only the (200) beam is expected to disappear when the zinc-blende structure vanishes. These data show that $(GaAs)_{1-x}(Ge_2)_x$ and $(GaSb)_{1-x}(Ge_2)_x$ undergo phase transitions at critical compositions x_c (with $x_c \simeq 0.4$ and $x_c \simeq 0.3$, respectively) in which the zinc-blende character—the distinction between anion and cation sites—is suddenly lost. The solid lines are smooth curves through the data, and the dashed lines are the mean-field predictions for the dependences of the ratios R on the alloy compositions x (assuming that x_c is 0.4 and 0.3, respectively): $R = |M(x)/[M(x=0)(1-g\cdot)]|^2$. (This formula for R uses atomic form factors for GaSb, GaAs, and Ge obtained from Ref. 18 and ignores such corrections as Debye-Waller factors. Hence we have $g = 0.257$ and $g = 0.051$, respectively.)

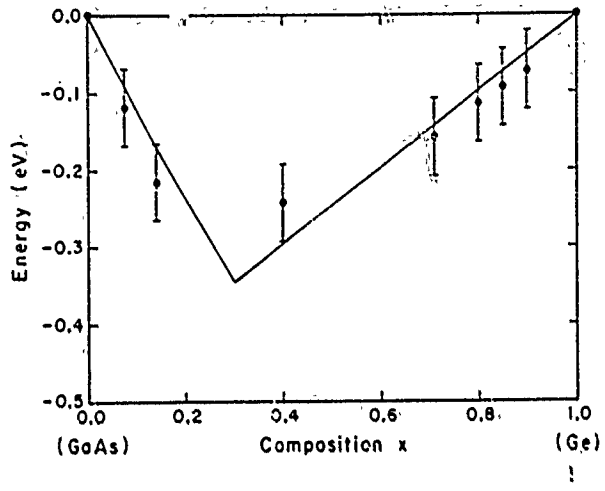


FIG. 2. Measured variation of the E_1 edge (Ref. 21) of $(GaAs)_{1-x}(Ge_2)_x$ (grown on GaAs substrates) versus alloy composition x, compared with the theory. To emphasize the "bowing" we have subtracted a straight line from each of the theory (solid line) and the data (points) in order to force the variation to be zero for $x=0$ and $x=1$.

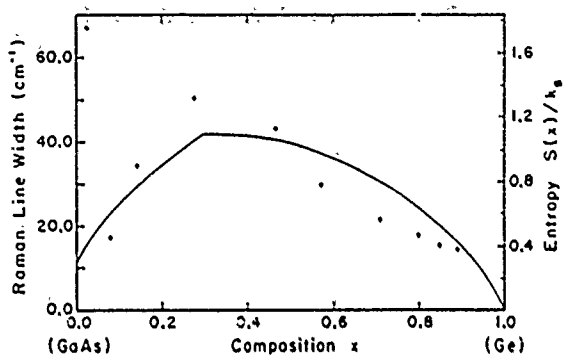


FIG. 3. Measured full width at half maximum (in cm^{-1}) of the long-wavelength Ge longitudinal optic Raman line of $(GaAs)_{1-x}(Ge_2)_x$ vs alloy composition (points, left-hand axis) in comparison with the entropy per site $S(x)$, divided by Boltzmann's constant, as evaluated in the phase-transition model (solid line, right-hand axis).

the Raman line, both because the Ge atoms will be polarized by Ga and As (activating "forbidden" Raman lines and shifting line frequencies) and because the many different cluster environments of Ge, Ga, and As should give rise to inhomogeneous broadening of the LO and TO modes in the alloy, causing them to overlap (they are separated by only $\simeq 15\ cm^{-1}$ in GaAs). Indeed, the Raman spectra show one prominent *broad* long-wavelength optic line which, for some compositions x, contains two discernible peaks (this line was fit using least-squares methods to a two-peak model in Ref. 16. Thus, the nature of the alloy disorder should manifest itself in the width of this line: when the disorder is a maximum, we expect the linewidth to be a maximum, due to the contributions from the disorder-activated modes and new alloy cluster modes. Thus, roughly speaking, the linewidth is a (semiquantitative) measure of the disorder or the *entropy* of the alloy. In Fig. 3 we display the Raman linewidth obtained from $(GaAs)_{1-x}(Ge_2)_x$ as a function of x. We also present, for comparison, the calculated entropy per site $S(x)$ of the zinc-blende–diamond phase-transition model.[2,25] We find that the entropy peaks near x_c ($\simeq 0.3$) and has a discontinuous derivative with respect to x at x_c, features also apparent in the data.[26] Thus the vibrational states of the alloy confirm both the electronic evidence gleaned from the ellipsometry data and the crystal-structure evidence provided by x-ray diffraction: $(GaAs)_{1-x}(Ge_2)_x$ undergoes a transition as a function of alloy composition x near $x_c \simeq 0.3$. Similar results were also obtained for the Ge linewidth as a function of composition in the related alloy $(GaSb)_{1-x}(Ge_2)_x$.[8,27]

The direct-gap data, the x-ray diffraction data, the ellipsometry data, the Raman linewidths, and the theory, when taken together, all provide strong evidence of a zinc–blende-to-diamond phase transition in $(GaAs)_{1-x}(Ge_2)_x$, and its effects on the crystal, electronic, and vibrational structures of these alloys. Available data for $(GaSb)_{1-x}(Ge_2)_x$, a similar metastable alloy, also lend support to the theory.[8,27]

Thus the present theory resolves some of the controversy surrounding the original phase-transition theory,[2] by demonstrating that there is a phase transition with $x_c \simeq 0.3$. However, the aspect of the original theory that has generated the most controversy[9,11] is the prediction that the alloys $(GaAs)_{1-x}(Ge_2)_x$ and $(GaSb)_{1-x}(Ge_2)_x$ should contain numerous antisite defects, that is, As—As or Sb—Sb bonds. References 9, 11, and 28 propose alternative models of these alloys that *do not include* the "wrong" bonds As—As or Sb—Sb. The model in Ref. 28 is a percolation model, with a zinc-blende–diamond transition composition x_c characteristic of a site-diluted diamond lattice, $x_c = 0.572$.[29] The Holloway-Davis model[9] is a modified percolation model, with a calculated x_c of $\simeq 0.75$. The Kim-Stern model[11] is a "kinetic" extended cellular-automata model, producing $x_c \simeq 0.26$ for growth in the [100] direction and $x_c < 0.18$ for spherical growth. We note that only one of these alternative models, that of Kim and Stern,[11] agrees even qualitatively with the x-ray diffraction data of Fig. 1.

Predictions based on the Kim-Stern model for either the ellipsometry data or the Raman linewidth are not presently available, but the entropy may be evaluated in an on-site model[25] that requires all Ga atoms and Sb atoms to remain on their natural sites. This on-site model forbids Ga—Ga or Sb—Sb bonds and is, in that respect, similar to the Kim-Stern model, but is different in that it has no critical composition. The entropy determined from the on-site model is in marked disagreement with the Raman linewidth data, while the entropy from the phase-transition model[8,25] has the same general alloy dependence as the observed linewidths.

Recent extended x-ray-absorption fine-structure (EXAFS) data[10] yield partial information about the atom-atom correlations in $(GaSb)_{1-x}(Ge_2)_x$ grown on glass. The interpretation of these data depends on a number of assumptions. We assume first that the alloys are *on*-stoichiometry (that is, there are the same number of Sb

atoms as Ga atoms); that the crystal is of good quality (that is, that Sb does not phase segregate on the surface or along grain boundaries); and that there are *no* vacancies.[30] We then extract from the data the probabilities of forming Ga—Sb, Ge—Sb, and Sb—Sb bonds (for details, see Appendix A) The results are shown in Fig. 4 and are compared with those determined from the phase-transition model using mean-field theory [for details, see Appendix A, and, in particular, Eqs. (A2)]. In making these comparisons, we must know the phase-transition composition x_c for $(GaSb)_{1-x}(Ge_2)_x$ grown on glass. Since zinc-blende–lattice spots were only seen for a sample with $x=0.1$ (not analyzed using EXAFS),[10] the EXAFS data used in generating Fig. 4 appear to be for samples that are all in the diamond phase. With that assumption, the extracted Sb—Sb bond probabilities

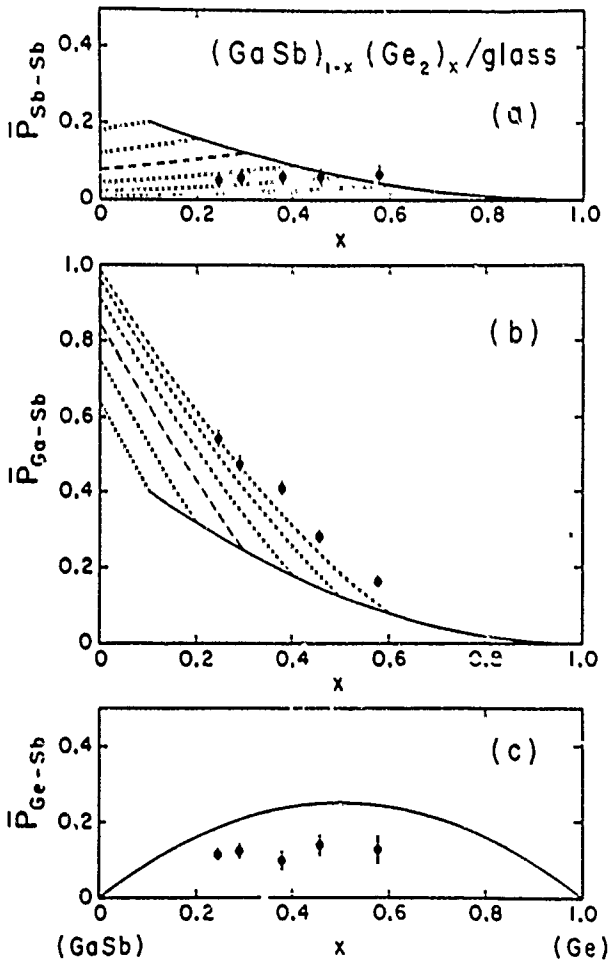


FIG. 4. Comparison of mean-field bond probabilities vs composition x with those extracted from EXAFS data (circles) (Ref. 10) for $(GaSb)_{1-x}(Ge_2)_x$ grown on glass: (a) Sb—Sb bond probability $\bar{P}_{Sb-Sb}$; (b) Ga—Sb bond probability $\bar{P}_{Ga-Sb}$; (c) Ge—Sb bond probability $\bar{P}_{Ge-Sb}$. Solid curves represent the theory for $x > x_c$ (diamond phase) while dashed curves show mean-field results for $x < x_c$ (zinc-blende phase, with x_c chosen to be 0.1, 0.2, 0.3 long dashes), 0.4, 0.5, and 0.6. Note that the extracted EXAFS Sb—Sb bond probabilities are meaningful only for those samples that are in the diamond phase. For details, see the text and Appendix A.

$\bar{P}_{Sb-Sb}$, shown as circles in Fig. 4(a), are all between 5% and 7%. This result is qualitatively consistent with the idea from the original phase-transition theory[2] that $\bar{P}_{Sb-Sb}$ is nonzero[31] and implicitly contradicts the central assumption of the Kim-Stern model, that $\bar{P}_{Sb-Sb}$ is zero. Now compare the Ga—Sb bond probabilities determined from the EXAFS data [denoted circles in Fig. 4(b)] with the results from the phase-transition model using mean-field theory[2] [solid line (diamond phase) and dashed lines (zinc-blende phase) in Fig. 4(b), with x_c chosen between 0.1 and 0.6]. Qualitative agreement with the theoretical result is obtained for x_c chosen to be larger than 0.1, and of the order of that observed in Fig. 1 for $(GaSb)_{1-x}(Ge_2)_x$ grown on GaAs: $x_c \simeq 0.3$ [long dashed lines in Fig. 4(b)]. If the phase-transition composition of $(GaSb)_{1-x}(Ge_2)_x$ is indeed $x_c \simeq 0.3$, then only the extracted bond probabilities for Sb—Sb with $x > 0.3$ are reliable in Fig. 4(a) (see Appendix A). In either case, it appears that Sb—Sb bonds do occur with non-negligible probabilities (5–7 %) for compositions x near 0.5.

Further EXAFS measurements of the Sb edge would clarify this issue by directly revealing the number of Sb—Sb bonds. It would be desirable if these measurements were made on samples for which x-ray diffraction data exist, of the quality of Fig. 1, e.g., the single-crystal samples of $(GaSb)_{1-x}(Ge_2)_x$ grown on GaAs.

Thus the phase-transition model, evaluated in a mean-field approximation, provides a rather satisfactory description of the long-wavelength properties of $(GaAs)_{1-x}(Ge_2)_x$ and $(GaSb)_{1-x}(Ge_2)_x$. As expected, the mean-field approximation does not provide a satisfactory quantitative description of short-wavelength atomic correlations, as determined by EXAFS; and a theory that correctly predicts that number of "wrong" anion-anion bonds is needed.

ACKNOWLEDGMENTS

We are grateful to the Office of Naval Research (Contract Nos. N00014-84-K-0352, N00014-79-C-0486, and N00014-85-K-0158) for their generous support of this research. One of us (K.E.N.) acknowledges useful discussions with B. L. Gu, P. A. Fedders, and E. A. Stern and partial support from the Eloise Gerry Foundation of Sigma Delta Epsilon, Graduate Women in Science. We are especially grateful to S. A. Barnett, J. E. Greene, M. A. Ray, B. Kramer, and I. Shah for allowing us to show their data prior to publication (Ref. 4) and to compare them with the theory in Fig. 1.

APPENDIX

Here we derive the formulas used in extracting bond probabilities from the EXAFS data.[10]

Denote the probabilities that atoms i and j occupy cation or anion sites to be P^c and P^a, respectively, where i,j take on the values Ga, Ge, and Sb. Similarly, let P_{ij}

be the *oriented-pair* probability that an atom i on a cation site is bonded to an atom j on an anion site. Note, $P_{ij}=P_i^c P_j^a+\gamma_{ij}$, where γ_{ij} is the difference between a mean-field theory ($\gamma_{ij}=0$) and a correlated theory ($\gamma_{ij}\neq 0$).[31] The oriented-pair probability P_{ij} is *different* than the probability $\bar{P}_{ij}$ of forming a bond between atoms i and j:

$$\bar{P}_{ij}=P_{ij}+P_{ji} \quad \text{if } i\neq j$$
$$=P_{ij} \quad \text{if } i=j \ . \tag{A1}$$

In mean-field theory we find

$$\bar{P}_{\mathrm{Ga-Sb}}=[(1-x)^2+M^2]/2 \quad \text{for } x<x_c$$
$$=(1-x)^2/2 \quad \text{for } x>x_c \ ,$$
$$\bar{P}_{\mathrm{Ga-Ge}}=\bar{P}_{\mathrm{Sb-Ge}}=x(1-x) \ , \tag{A2}$$
$$\bar{P}_{\mathrm{Sb-Sb}}=\bar{P}_{\mathrm{Ga-Ga}}=[(1-x)^2-M^2]/4 \quad \text{for } x<x_c$$
$$=(1-x)^2/4 \quad \text{for } x>x_c \ ,$$

and

$$\bar{P}_{\mathrm{Ge-Ge}}=x^2 \ .$$

Now, assuming that all sites are occupied (there are no vacancies), we have the following relationship between the oriented-pair probabilities and the on-site probabilities:

$$\sum_i P_{ij}=P_j^a \tag{A3}$$

and

$$\sum_j P_{ij}=P_i^c \ , \tag{A4}$$

where

$$\sum_i P_i^c=\sum_j P_j^a=1 \ . \tag{A5}$$

Similarly, the bond probabilities obey the normalization rule:

$$\bar{P}_{\mathrm{Ge-Ga}}+\bar{P}_{\mathrm{Ga-Sb}}+\bar{P}_{\mathrm{Ga-Ge}}+\bar{P}_{\mathrm{Sb-Sb}}+\bar{P}_{\mathrm{Sb-Ge}}+\bar{P}_{\mathrm{Ge-Ge}}=1 \ . \tag{A6}$$

The normalizations for the oriented-pair probabilities are thus more restrictive than those for the bond probabilities: Three probabilities enter Eqs. (A2) and (A3), versus six in Eq. (A5).

The EXAFS measurements[10] determine the numbers of Sb neighbors of Ga and Ge, which we denote here as the functions $A(x)$ and $B(x)$, respectively. These functions are related to the *bond* probabilities by

$$A(x)=4\bar{P}_{\mathrm{Ga-Sb}}/(1-x) \tag{A7}$$

and

$$B(x)=4\bar{P}_{\mathrm{Sb-Ge}}/2x \ . \tag{A8}$$

Since zinc-blende–lattice spots were not observed in x-ray diffraction[10] (with the exception of a sample with $x=0.1$ that was not analyzed using EXAFS), we assume for the moment, that all samples were in the diamond phase. This implies then that

$$P_{ij}=P_{ji} \quad (x>x_c) \ , \tag{A9}$$

allowing us to extract both *oriented-pair* probabilities as well as *pair* probabilities from the EXAFS measurements:

$$\bar{P}_{\mathrm{Ga-Sb}}=2P_{\mathrm{Ga-Sb}}=2P_{\mathrm{Sb-Ga}}=(1-x)A(x)/4 \ , \tag{A10}$$

$$\bar{P}_{\mathrm{Ge-Sb}}=2P_{\mathrm{Ge-Sb}}=2P_{\mathrm{Sb-Ge}}=xB(x)/2 \ , \tag{A11}$$

and, using Eq. (A3),

$$\bar{P}_{\mathrm{Sb-Sb}}=P_{\mathrm{Sb-Sb}}=(1-x)/2-P_{\mathrm{Ga-Sb}}-P_{\mathrm{Ge-Sb}} \ . \tag{A12}$$

These extracted probabilities are plotted in Fig. 4 along with the results from mean-field theory, Eqs. (A2). Note that $\bar{P}_{\mathrm{Sb-Sb}}$ is decidedly nonzero and that all the extracted bond probabilities are of the same order of magnitude as the mean-field results.

---

*Present address: University of Maryland, College Park, MD 20742.

[1]K. E. Newman, A. Lastras-Martinez, B. Kramer, S. A. Barnett, M. A. Ray, J. D. Dow, J. E. Greene, and P. M. Raccah, Phys. Rev. Lett. 50, 1466 (1983).

[2]K. E. Newman and J. D. Dow, Phys. Rev. B 27, 7495 (1983).

[3]J. E. Greene, J. Vac. Sci. Technol. B 1, 229 (1983); J. L. Zilko and J. E. Greene, J. Appl. Phys. 51, 1549 (1980); 51, 1560 (1980); K. C. Cadien, A. H. Eltoukhy, and J. E. Greene, Appl. Phys. Lett. 38, 773 (1981).

[4]S. A. Barnett, B. Kramer, L. T. Romano, S. I. Shah, M. A. Ray, S. Fang, and J. E. Greene, in *Layered Structures, Epitaxy, and Interfaces,* edited by J. M. Gibson and L. R. Dawson (North-Holland, Amsterdam, 1984); S. I. Shah and J. E. Greene (unpublished); M. A. Ray and J. E. Greene (unpublished).

[5]Z. I. Alferov, R. S. Vartanyan, V. I. Korol'kov, I. I. Mokan, V. P. Ulin, B. S. Yavich, and A. A. Yakovenko, Fiz. Tekh. Poluprovodn. 16, 887 (1982) [Sov.—Phys. Semicond. 16, 567 (1982)].

[6]B. Koiller, M. A. Davidovich, and R. Osório, Solid State Commun. 55, 861 (1985).

[7]R. Beserman, J. E. Greene, M. V. Klein, T. N. Krabach, T. C. McGlinn, L. T. Romano, and S. I. Shah, in *Proceedings of the 17th International Conference on Physics of Semiconductors, San Francisco, 1984,* edited by D. J. Chadi (Springer-Verlag, New York, 1985), p. 961. For an updated version of the conclusions of this paper, see Ref. 8.

[8]K. E. Newman, J. D. Dow, A. Kobayashi, and R. Beserman, Solid State Commun. 56, 553 (1985).

[9]H. Holloway and L. C. Davis, Phys. Rev. Lett. 53, 1510 (1984).

[10]E. A. Stern, F. Ellis, K. Kim, L. Romano, S. I. Shah, and J. E. Greene, Phys. Rev. Lett. 54, 905 (1985).

[11]K. Kim and E. A. Stern, Phys. Rev. B 32, 1019 (1985).

[12]Figures illustrating the different crystal structures of the alloy may be found in Refs. 1 and 2. In the zinc-blende phase, the majority of anions (cations) occupy nominal anion (cation) sites, whereas in the diamond phase there is no site preference.

[13]See, for example, S.-K. Ma, *Modern Theory of Critical Phenoman* (Benjamin, Reading, MA, 1976).

[14]J. E. Greene, C. E. Wickersham, and J. L. Zilko, J. Appl. Phys. 47, 2289 (1976).

[15]J. E. Greene and C. E. Wickersham, J. Appl. Phys. 47, 3630 (1976).

[16]S. A. Barnett, M. A. Ray, Lastras, B. Kramer, J. E. Greene, P. M. Raccah, and L. L. Abels, Electron. Lett. 18, 891 (1982). Typographical errors in the plot of direct band gap versus alloy composition have been corrected in Ref. 1.

[17]J. W. Colby (unpublished).

[18]B. D. Cullity, *Elements of X-ray Diffraction* (Addison-Wesley, Menlo Park, CA, 1978); B. E. Warren, *X-ray Diffraction* (Addison-Wesley, New York, 1969).

[19]I. Banerjee, D. W. Chung, and H. Kroemer [Appl. Phys. Lett. 46, 494 (1985)] report that the (200) reflections in molecular-beam-epitaxy grown $(GaAs)_{1-x}(Ge_2)_x$ are extremely weak for $x > 0.31$.

[20]The normalized (200) intensity should be proportional to the square of the magnitude of the order parameter, $|M(x)|^2$, which, in a mean-field theory, is linear in $x - x_c$.

[21]The direct band gap at Γ ($\mathbf{k}=0$) is termed E_0. E_1 is the band gap at the L point in the Brillouin zone, and $E_1+\Delta$ is the slightly higher gap that is the spin-orbit-split partner of E_1.

[22]Note that the theoretical prediction of the E_1 transition is expected to be less accurate than that of the E_0 transition: Ref. 2 used an empirical tight-binding theory [P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 353 (1983); K. E. Newman and J. D. Dow, Phys. Rev. B 30, 1929 (1984)] that accurately fit the conduction and valence bandstructure energies at the Γ point [$\mathbf{k}=(0,0,0)$] and the X point [$\mathbf{k}=2\pi(1,0,0)/a$], but only the conduction-band edge at L [$\mathbf{K}=\pi(1,1,1)/a$].

[23]We note that the *rounding* of the V-shaped curve is due to correlations within the alloy. For more details, see the discussion in Ref. 6.

[24]P. M. Raccah, S. A. Barnett, J. Greene, and I. Shah unpublished).

[25]The entropy per site in mean-field approximation for the phase-transition model is

$$S(x)/k_B=[(1-x+M)/2]\ln[(1-x+M)/2]$$
$$+[(1-x-M)/2]\ln[(1-x-M)/2]+x\ln(x)\,,$$

where k_B is Boltzmann's constant. The entropy for the on-site model is $S(x)/k_B=(1-x)\ln(1-x)+x\ln(x)$.

[26]The location of the maximum of the entropy curve $S(x)$ depends on the value of x_c. If $x_c < \frac{1}{3}$, then $S(x)$ has a discontinuous derivative at x_c and a maximum at $x=\frac{1}{3}$. If $x_c > \frac{1}{3}$, then $S(x)$ both has a maximum and a discontinuous derivative at x_c. The calculation shown in Fig. 3 is for $x_c \simeq 0.3$, as found in Ref. 2.

[27]A. Kobayashi, K. E. Newman, and J. D. Dow, Phys. Rev. B 32, 5312 (1985).

[28]M. I. D'yakonov and M. E. Raikh, Fiz. Tekh. Poluprovodn. 16, 890 (1982) [Sov.—Phys. Semicond. 16, 570 (1982)].

[29]For a discussion of percolation, see, for example, R. Zallen, *The Physics of Amorphous Solids* (Wiley, New York, 1983).

[30]As an alternative explanation, Ref. 10 suggests that 30% of the Ge atoms are along grain boundary defects.

[31]For fits of the bond probabilities $\bar{P}_{Sb-Sb}$ to the EXAFS data using a correlated theory, see B. L. Gu, K. E. Newman, and P. A. Fedders, Phys. Rev. B 35, 9135 (1987).

Pseudofunction theory of the electronic structure of InN

M.-H. Tsai, David W. Jenkins, and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

R. V. Kasowski
E. I. du Pont de Nemours and Company, Central Research and Development Department, Experimental Station, Wilmington, Delaware 19898

(Received 22 June 1987)

The energy-band structure of InN is predicted using the pseudofunction method (a first-principles, self-consistent local-density scheme). Some significant differences exist between this electronic structure and extrapolated empirical tight-binding theory for InN.

I. INTRODUCTION

InN has recently been observed to have both a high mobility (4×10^3 cm^2/V sec, Ref. 1) and a band gap in the visible region[2] (orange, about 2 eV), leading to the suggestion that $In_{1-x}Ga_xN$ alloys may be fabricated into detectors, lasers, and light emitters in the visible and ultraviolet portion of the electromagnetic spectrum.[3,4] While the band gap of GaN lies in the ultraviolet region and its energy-band structure has been thoroughly studied,[5,6] the band structure of InN is not well understood—having been investigated only recently, with conflicting results.[1,2,7,8]

Calculations of the InN band structure based on empirical pseudopotential theory[1,2] have produced a fundamental band gap that, by construction, equals the observed value of 2 eV. To our knowledge, the only electronic structure theory for InN with any proven ability to actually predict the band gap is an extrapolated empirical tight binding theory.[7] Tight-binding parameters determined by fitting the known band gaps of other III-V semiconductors have been used to extrapolate the corresponding parameters for InN and to successfully predict its 2-eV band gap. Tight-binding theory, while predicting a good band gap, produces valence bands that are quite narrow in comparison with the valence bands of empirical pseudopotential theory. These conflicting features of the published band structures suggest that more theoretical and experimental work on InN is needed. Thus a genuine first-principles theory of the band structure of InN is called for.

In this paper we present the band structure of InN, as calculated using the pseudofunction method[9]—a first-principles, self-consistent scheme based on the local-density approximation. Our results differ in detail both from previous empirical pseudo-potential theory[1,2] and from tight-binding theory,[7] and indicate the need for measurements such as photoemission to resolve remaining controversies concerning the electronic structure of InN.

II. METHOD

We employ the pseudofunction method, which is a scheme for solving the one-electron Schrödinger equation self-consistently in a local-density approximation. Exchange and correlation are simulated by the Hedin-Lundquist potential.[10] Unlike many implementations of local-density theory, the pseudofunction method uses a real-space scheme for describing local bonding and is especially well suited for studying non-free-electron systems. For low-atomic-number atoms such as N, it should be more rapidly convergent than schemes that rely on the plane-wave character of the solid-state wave functions because this character generally results from the orthogonality of the wave functions to many core wave functions—and the low-atomic-number atoms have too few core states to effectively produce the free-electron character necessary for rapid convergence in k space.[11]

The pseudofunction method has elements in common with the linearized muffin-tin orbital (LMTO) method[12] and the extended muffin-tin orbital (EMTO) scheme,[13] and can be regarded as a local-density-approximation scheme involving a more general basis set than either the LMTO or EMTO bases. As such, it is, at least in principle, an improvement over those implementations of local-density theory. In principle, the LMTO method is an improvement of the LMTO method, because it involves a more extensive basis set. Similarly the improvement of the pseudofunction scheme over the EMTO method lies in the fact that its basis orbitals have not been subjected to muffin-tin boundary conditions (the radial logarithmic derivatives of localized muffin-tin orbitals are, by construction, always negative at the muffin-tin radius).

The pseudofunction method produces (as expected) rather good electron energy bands and bond lengths—results comparable with those of the best local-density methods. For example, the pseudofunction valence bands of bulk Si agree with those computed using the linearized augmented-plane-wave (LAPW) scheme[14] to

 ©1988 The American Physical Society

within typically a few tenths of an eV. The computed Si—Si bond length is 2.34 Å at low temperature, in excellent agreement with the data. The bulk electronic structure of AlN (a material similar to InN) predicted by the pseudofunction method agrees with photoemission data.[15] The Si(111)2×1 surface states of pseudofunction theory[16] also agree with those obtained using a norm-conserving pseudopotential.[17] Pseudofunction calculations for a potassium overlayer on the Si(001)2×1 surface found a bond length of 3.3 Å,[18] compared with the experimental value of 3.14±0.1 Å (Ref. 19) and the value 2.59 Å obtained using a norm-conserving pseudopotential scheme.[20] Moreover, the pseudofunction method has been shown to work well even for problems involving adsorbates of atoms such as C and O,[21] which occupy the same row of the Periodic Table as N. The method is also computationally fast compared with most others, and has been used to compute some very complicated electronic structures (using local-density theory): for example, the electronic structure of solids such as pyrochlore, with 22 atoms per unit cell.[22] Details of the method are available in the literature.[9]

III. RESULTS

Our predicted band structure is given in Fig. 1, and should be compared with tight-binding theory[7] (Fig. 2). An empirical pseudopotential band structure of InN has also been reported,[1,2] but will not be treated here because, as discussed previously,[4,23] the tight-binding theory is definitely superior. Our calculations assume the wurtzite crystal structure.[24]

The major qualitative features of our bands are similar to those of tight-binding theory; there are only a few significant quantitative differences. (i) We find small splittings at the zone center that are absent from tight-binding theory due to the fact that we used the observed

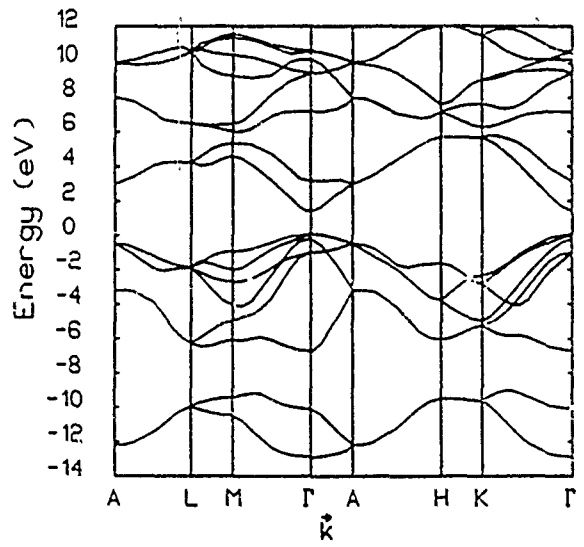


FIG. 1. Predicted band structure of wurtzite InN using the pseudofunction method. The symmetry points of the Brillouin zone are $A=(2\pi/c)(0,0,1/2)$, $L=(2\pi/a)(1/\sqrt{3},0,a/2c)$, $M=(2\pi/a)(1/\sqrt{3},0,0)$, $\Gamma=(0,0,0)$, $H=(2\pi/a)(1/\sqrt{3},1/3,a/2c)$, and $K=(2\pi/a)(1/\sqrt{3},1/3,0)$.

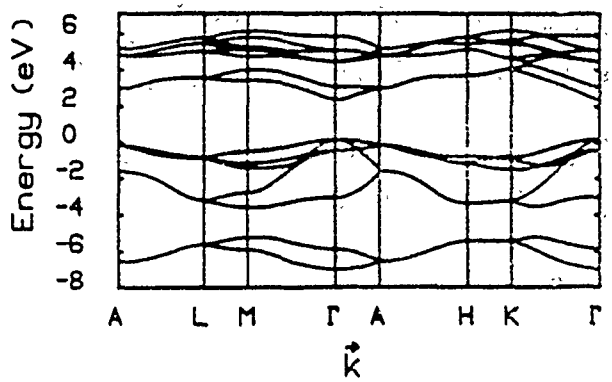


FIG. 2. Predicted band structure of wurtzite InN using the tight-binding method, after Ref. 7.

lattice constants c and a which have $c/a=1.6114$, whereas the empirical tight-binding theory assumed an ideal c/a ratio of $(\frac{8}{3})^{1/2}=1.6330$. (ii) We find a direct fundamental band gap of 1.3 eV, compared with the 2-eV result of tight-binding theory[7] and experiment.[2] This is a well-known limitation of local-density theory: it produces gaps that are too small[25]—a problem that can be circumvented, but only with great difficulty.[26] (iii) Our upper valence-band width is considerably wider than the tight-binding width: 6.8 eV versus 3.5 eV. Here there is no unambiguous theoretical criterion determining which theory is better: practitioners of both local-density theory and empirical tight-binding theory would each argue that their preferred type of theory is more likely to produce realistic electronic structures. In the case of AlN, the local-density[15,27] and tight-binding[28] theories gave bandwidths comparable with each other and with data.[27,29] This makes the discrepancy between the theories for InN particularly interesting. The issue of the correct valence-band width must be determined experimentally. Nevertheless, we suspect that the tight-binding bands are narrower than the local-density bands because the off-diagonal tight-binding matrix elements may be unduly small. These matrix elements are extrapolated from the corresponding matrix elements of other zinc-blende semiconductors, assuming that they scale inversely as the square of the bond length, Harrison's rule.[30] However, Harrison's rule is best applied to interpolate rather than extrapolate; it is known to be crude; and extrapolations based on it may not be trustworthy, especially for InN, with its large-radius cation and small anion. (iv) The valence-band effective masses are 1.6 and 1.7 free-electron masses (in the directions perpendicular to and parallel to the c axis, respectively) in the present model, compared with 2.7 and 2.7 for tight-binding theory. The corresponding conduction-band masses are 0.34 and 0.37 for the present work and 0.59 and 0.59 for tight-binding theory. The tight-binding masses are heavier by about 70%, reflecting the flatter bands. This is just another manifestation of the (likely) flat-band problem of the tight-binding theory. Assuming these masses and a static dielectric constant of 8.3,[31] we obtain shallow donor and acceptor binding energies[32] of 0.067 and 0.316 eV in a hydrogenic (Rydberg effective-mass) model versus 0.118

and 0.533 eV in tight-binding theory—which often obtains masses 100% in error). The observed binding energies of neutral substitutional donors and acceptors in InN are ~0.05 eV (Ref. 33) and ~0.20 eV,[4,34] respectively—generally consistent with either theory, but generally more supportive of the present local-density theory.

IV. SUMMARY

A major difference between the tight-binding and pseudofunction band structures is the width of the upper valence band, which is large in local-density theory but small in tight-binding theory. Which, if either, theory is correct can be determined by photoemission experiments.

ACKNOWLEDGMENTS

We gratefully acknowledge the generous support of the Office of Naval Research (Contract No. N00014-84-K-0352). We also thank N. E. Christensen for sending us a preliminary version of his local-density calculations, which give results apparently significantly different from ours. One of us (J.D.D.) gratefully acknowledges the support of the Minnesota Supercomputer Institute, and stimulating conversations with B. Goldenberg.

[1]T. L. Tansley and C. P. Foley, in *Proceedings of the Third International Conference on Semi-Insulating III-V Materials, Warm Springs, Oregon, 1984*, edited by D. C. Look and J. S. Blakemore (Shiva, London, 1984), pp. 497–500.
[2]C. P. Foley and T. L. Transley, Phys. Rev. B 33, 1430 (1986).
[3]K. Osamura, S. Naka, and Y. Murakami, J. Appl. Phys. 46, 3432 (1975).
[4]D. W. Jenkins and J. D. Dow (unpublished).
[5]S. Bloom, J. Phys. Chem. Solids 32, 2027 (1971).
[6]S. Bloom, G. Harbeke, E. Meier, and I. B. Ortenburger, Phys. Status Solidi B 66, 161 (1974).
[7]D. W. Jenkins, R.-D. Hong, and J. D. Dow, Superlatt. Microstruct. 3, 365 (1987).
[8]N. E. Christensen (unpublished).
[9]R. V. Kasowski, M.-H. Tsai, T. N. Rhodin, and D. D. Chambliss, Phys. Rev. B 34, 2656 (1986).
[10]L. Hedin and B. I. Lundquist, J. Phys. C 4, 2064 (1971).
[11]B. J. Austin, V. Heine, and L. J. Sham, Phys. Rev. 127, 276 (1962).
[12]O. K. Anderson, Phys. Rev. B 12, 3060 (1975).
[13]R. V. Kasowski, Phys. Rev. B 25, 4189 (1982).
[14]D. R. Hamann, Phys. Rev. Lett. 42, 662 (1979).
[15]R. V. Kasowski and F. S. Ohuchi, Phys. Rev. B 35, 9311 (1987).
[16]M.-H. Tsai and R. V. Kasowski, and T. N. Rhodin, Surf. Sci. 179, 143 (1987).
[17]J. E. Northrup and M. L. Cohen, Phys. Rev. Lett. 49, 1349 (1982); K. C. Pandey, *ibid.* 49, 223 (1982).
[18]M.-H. Tsai and R. V. Kasowski, Bull. Am. Phys. Soc. 32, 865 (1987); and R. V. Kasowski and M.-H. Tsai, Phys. Rev. Lett. 60, 546 (1988).
[19]T. Kendelewicz, P. Soukiassian, R. S. List, J. C. Woicik, P. Pianetta, I. Lindau, and W. E. Spicer, Phys. Rev. B 37, 7115 (1988).
[20]S. Ciraci and I. P. Batra, Phys. Rev. Lett. 56, 877 (1986).
[21]R. V. Kasowski, M.-H. Tsai, T. N. Rhodin, and D. D. Chambliss, Solid State Commun. 59, 57 (1986).
[22]R. V. Kasowski, W. Y. Hsu, A. W. Sleight, A. L. Wachs, A. P. Shapiro, T.-C. Chiang, and M.-H. Tsai, *TMC Ternary and Multinary Compounds, Materials Research Society Proceedings Series*, edited by S. Deb and A. Zunger (MRS, Pittsburgh, 1987), p. 533.
[23]See footnote 14 of Ref. 7.
[24]R. W. Wyckoff, *Crystal Structures* (Interscience, New York, 1960), Vol. I, Suppl. V. InN has a wurtzite structure with lattice constants $a=3.533$ Å and $c=5.693$ Å.
[25]M. S. Hybertsen and S. G. Louie, Phys. Rev. B 34, 5390 (1986).
[26]C. S. Wang, M. R. Norman, W. E. Pickett, and H. Krakauer, Bull. Am. Phys. Soc. 32, 820 (1987).
[27]W. Y. Ching and B. N. Harmon, Phys. Rev. B 34, 5305 (1986).
[28]A. Kobayashi, O. F. Sankey, S. M. Volz, and John D. Dow, Phys. Rev. B 28, 935 (1983).
[29]C. G. Olsen, J. H. Sexton, D. W. Lynch, A. J. Bevolo, H. R. Shanks, B. N. Harmon, W. Y. Ching, and D. M. Weiliczka, Solid State Commun. 56, 35 (1985).
[30]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), p. 47.
[31]The static dielectric constant is approximated by squaring the long-wavelength index of refraction (Ref. 2) to obtain $\epsilon=8.3$. We expect the splitting of the longitudinal- and transverse-optical phonons at the Brillouin-zone center to be small, causing the Lyddane-Sachs-Teller relationship [see, e.g., C. Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976), p. 306] to be nearly unity. Thus we approximate the static dielectric constant by the long-wavelength optical dielectric constant.
[32]W. Kohn, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5, p. 257.
[33]T. L. Tansley and C. P. Foley, J. Appl. Phys. 60, 2092 (1986).
[34]It is not generally recognized that the 0.2-eV feature in the data is in fact a shallow acceptor level (because the level is not "shallow" by the old definition—lying within 0.1 eV of the band edge). See Ref. 4.

PHYSICS OF DEEP LEVELS

JOHN D. DOW
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556 U.S.A.

ABSTRACT

The physics of deep levels is reviewed, with emphasis on the qualitative physics that has been elucidated as a result of the ideas of Lannoo, Lenglart, Hjalmarson, Vogl, Wolford, Hsu, Sankey, Allen, and others.

INTRODUCTION

In this paper we outline the physics of deep impurity levels in semiconductors, and show that the simple Hjalmarson et al. model of deep levels [1] can quantitatively predict the wavefunctions of substitutional sp^3-bonded deep impurities and explore the chemical trends of deep levels. We show that the Hjalmarson theory, when compared with data, (i) can determine the site (anion or cation) and symmetry (s-like A_1 or p-like T_2) of a defect, (ii) can eliminate typically 80% or more of the candidates for forming that defect, (iii) can successfully predict how the energy of the deep level varies from one impurity to another and from one host to another, and (iv) can predict unexpected phenomena caused by the deep level. We point out that it is very difficult to guess from theory and measured energy levels alone if the defects responsible for the observed levels are point defects or complexes. We emphasize that all of the best contemporary theories of deep levels are uncertain by a few tenths of an eV, and therefore it is not practical to determine which impurity is causing a given level by aligning the theoretical and experimental levels; in our opinion this theoretical uncertainty is not likely to be greatly reduced soon. Finally we show that the main differences between the Hjalmarson theory and various other theories [2,3] are (i) how the host band gaps are determined and (ii) computational costs: The Hjalmarson theory is based on the Vogl et al. empirical tight-binding Hamiltonian [4], and so fits the band gaps to data, while self-consistent pseudopotential and other theories calculate the host band gaps *a priori*, and then perform some adjustment. As a result, the self-consistent pseudopotential theories require orders of magnitude more computational effort to obtain comparably accurate deep levels; but these theories can predict total energies better than the Hjalmarson theory. Therefore, to understand the physics of deep levels, especially for complex defects, the simple Hjalmarson theory is preferable; while the self-consistent pseudopotential theories are better equipped to handle questions related to the total energies of defects, such as diffusion.

To begin the discussion of deep levels, we first make two important (and apparently self-contradictory) points: (i) Most deep levels are not deep; and (ii) Most deep impurity wavefunctions do not depend much on the impurity.

MOST DEEP LEVELS ARE NOT DEEP

Several years ago the definition of a deep level was based on its energy: a level in the fundamental band gap more than 0.1 eV distant from the nearest host band edge. In our discussions, we follow Hjalmarson [1] and use the term "deep" to describe localization: a level that is produced by the central-cell potential of the defect. Therefore, a "deep level" can lie above the

Mat. Res. Soc. Symp. Proc. Vol. 46. 1985 Materials Research Society

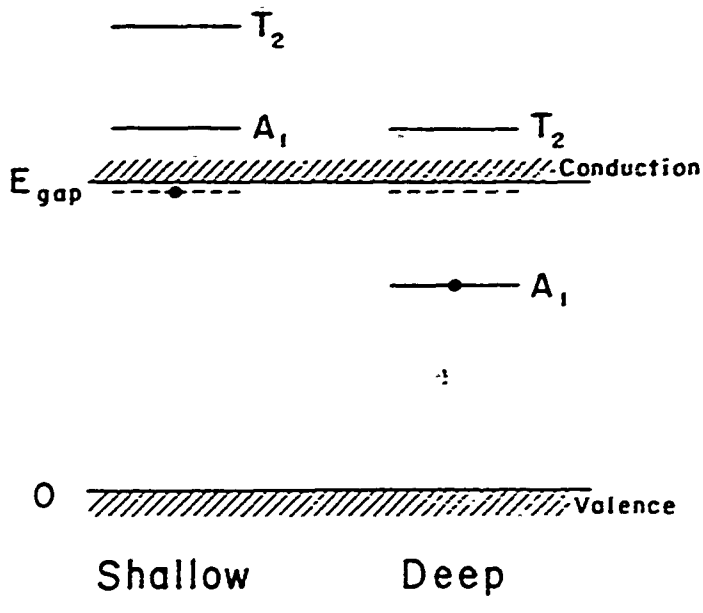


Fig. (1). Illustrating the difference between shallow and deep impurities. Note that both have deep levels, but that these levels lie outside the gap for shallow impurities.

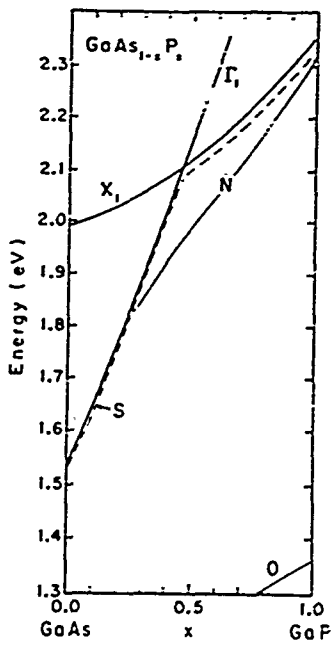


Fig. (2). X_1 and Γ_1 band edges, S shallow level, and N and O deep levels versus alloy composition x in $GaAs_{1-x}P_x$, as determined in Ref. [5], after [23].

conduction band edge, and not be "deep" by the old definition. Stated precisely, most defect levels that are "deep" by the new localization definition are not in the band gap, and so have "negative binding energies" and are not "deep" by the energy criterion.

An essential point is that every heterovalent substitutional sp^3-bonded defect must have <u>both</u> deep levels (due to the central-cell potential) and shallow levels (due to the long-ranged Coulomb potential associated with the valence difference $Z = -Ze^2/\varepsilon r$). Many readers will be astounded to learn that common shallow impurities, such as S on a P site in GaP have enormous central-cell defect potentials, = -5 eV deep, and yet these potentials appear to produce effects on only the 10 meV scale of typical shallow donor levels. (A good order of magnitude estimate of the central-cell potential is given by the difference in the atomic energies of S and P [4].) Clearly these large potentials must have effects on the eV scale — and they do: The local sp^3 bonding of the impurity is perturbed by the defect, producing different energies for the four tetrahedral bonds, namely four deep levels near the fundamental band gap. Three are p-like (of T_2 symmetry and degenerate in a tetrahedral environment) and one is s-like (A_1 symmetry). These deep levels due to the central-cell potential always exist near the fundamental band gap <u>in addition</u> to any shallow levels. (See Fig. [1].)

If one or more of these deep levels falls well within the fundamental band gap, then the impurity is termed a "deep impurity" and the level is "deep" by the energy definition (Fig. [1]). If none of these deep levels of a heterovalent impurity lies within the gap, the impurity is termed a "shallow impurity" because only its shallow levels are in the gap. However, the deep levels still exist, although they are resonant with and broadened by the host bands (Fig. [1]). For most semiconductors, which have small (=1 eV) band gaps on the =20 eV scale over which the spectral strength of the sp^3 bond is distributed, the majority of defects produce more resonant deep levels than deep levels in the gap. Thus most deep levels (by the localization definition) are not deep energetically.

Evidence supporting this picture is provided by the work of Wolford et al. [5] and Hsu et al. [6] on N, O, and S impurities on a P-site in $GaAs_{1-x}P_x$ alloys (Fig. [2]). In GaP, the N impurity produce a level energetically shallower than the shallow level of S. As a function of decreasing alloy composition x, however, the N level becomes deep and behaves like the deep O level, but not like the shallow S level. In fact, the N level becomes energetically deep for x=0.5, only to enter the conduction band for x=0.2. This behavior led Wolford et al. to conclude that N is in fact a deep level in a localization sense for all x, but lies outside the fundamental band gap, above the conduction band edge in GaAs.

Further confirmation of this picture has been supplied for P in Si_xGe_{1-x} alloys, where Bunker et al. [7,8] have provided experimental evidence that the deep level associated with P descends into the fundamental band gap near x=0.15. (See Fig. [3].)

WAVEFUNCTIONS INDEPENDENT OF IMPURITY

A rather surprising feature of many deep level wavefunctions is that for impurities at the same site (anion or cation) with the same symmetry in the same semiconductor, most defects with deep levels in the gap produce virtually the same deep level wavefunction. This is demonstrated in Fig. [4], where the A_1-symmetric wavefunctions of S, Se, and Te deep levels in Si, as determined by magnetic resonance [9] and by theory [10] are plotted versus the associated deep level energy E in the gap. The data alone show that the central-cell wavefunctions of S, Se, and Te are virtually the same; and the theory shows that the central-cell and first-neighbor wavefunctions are in agreement with the data and almost independent of either the deep level energy or the defect. Thus the deep-level wavefunctions are host-like, not impurity-like [1].

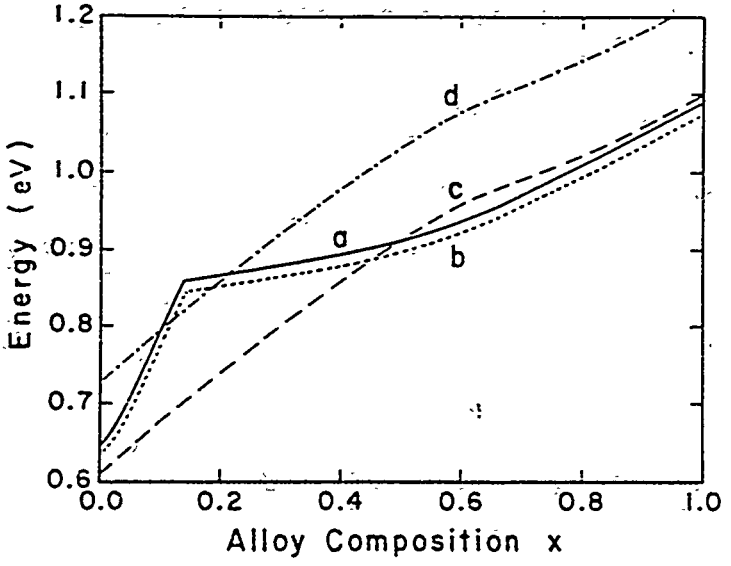


Fig. [3]. Energies versus alloy composition x in Si_xGe_{1-x} of (a) the conduction band edge, (b) the shallow P level, (c) the deep P level predicted in Ref. [8], and (d) the deep P level as determined by x-ray absorption measurements of [7], after [7]. Note that the deep level descends into the gap for x=0.15.

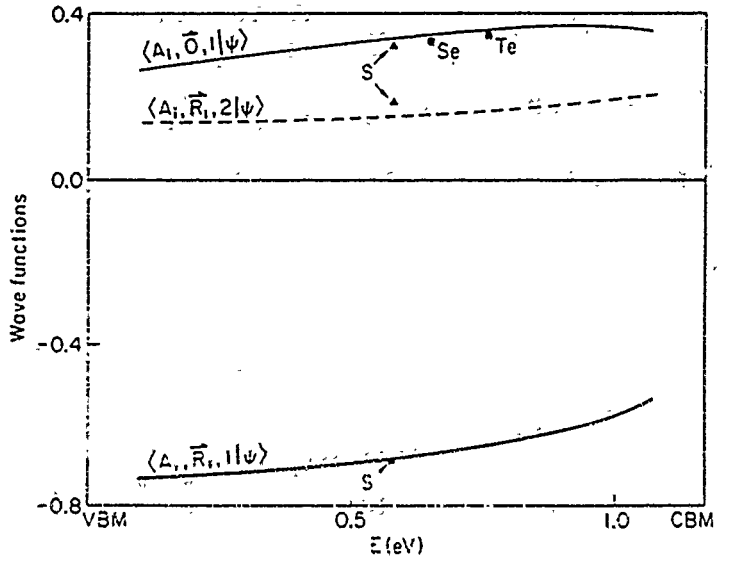


Fig. [4]. A_1-symmetric wavefunctions at the central site $\bar{0}$ and at the first-neighbor shell $\bar{R}_1$ for substitutional defects in Si, compared with data, after [10].

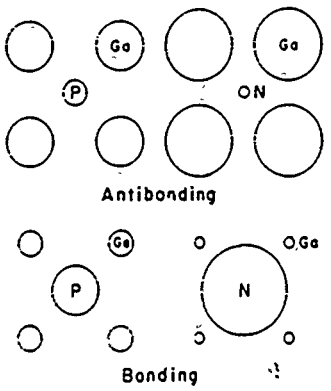


Fig. (5). Schematic wavefunctions of GaP and a N impurity on the P site in GaP, after [23].

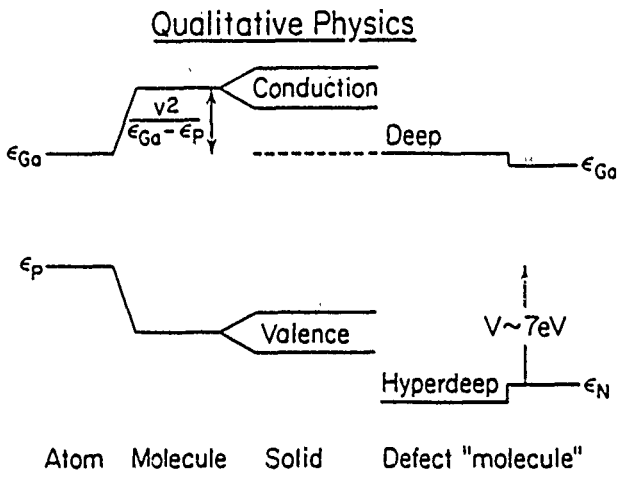


Fig. (6). Schematic energy level diagram for GaP and for a N defect in GaP, after [23].

The host-like character of deep-level wavefunctions could have been deduced from the experimental fact that impurities with atomic energies different by eV all produce deep levels in the gap that differ by an order of magnitude less, ≈0.1 eV. Clearly the deep levels in the gap are not impurity-like.

A schematic illustration of the physics of deep levels, for the case of N on a P site in GaP, is given in Figs. (5) and (6). N is more electronegative than P and so attracts more electronic charge to its bonding level (Fig. (5)), causing the antibonding level, by orthogonality, to have very little charge on the N-site and a host-like Ga-like wavefunction. In terms of energy levels (Fig. (6)), the host Ga and P atomic levels exhibit a bonding-antibonding splitting of order $v^2/(\varepsilon_{Ga}-\varepsilon_P)$, where v is a Ga-P transfer matrix element. N coupled to surrounding Ga exhibits a smaller bonding-antibonding splitting, $v^2/(\varepsilon_{Ga}-\varepsilon_N)$, because the energy denominator is ≈7 eV larger (v is almost the same for Ga-N and Ga-P [11]). The N-like level is the bonding "hyperdeep" level that lies below the valence band, is electrically inactive, and is not observed. The deep level in the gap is host-like (Ga-like) and antibonding in character. If we change the defect from P (V=0) to S (V= -5 eV) to N (V= -7 eV) to O (V= -15 eV) to a vacancy (V= -∞[2]), the deep level does not change its energy much (although the hyperdeep level does). As a result, a plot of deep level energy E versus central-cell defect strength V (or versus impurity atomic energy) looks like a parabola with the asymptote being the Ga dangling bond energy or pinning energy of Hjalmarson [1]. (See Fig. (7).)

For different sites (anion or cation) and symmetries (A_1 or T_2) the asymptotes lie at different energies, but several different impurities when on the same site all produce deep levels of a given symmetry within a few tenths of an eV of the pinning energy. Since the theoretical accuracy is also a few tenths of an eV, theory can determine the site and symmetry of an observed level, and can eliminate from consideration those impurities that lie more than 0.5 eV from the observed level, but it cannot definitively assign a level to a specific impurity. Such an approach identified oxygen and the antisite defect As_{Ga} as candidates for the defect EL2 [12]. Moreover, because of the host-like nature of deep level wavefunctions, most electronic probes cannot be of much assistance in defect identification -- probes that couple to the nucleus, such as ENDOR [9], or to the core electrons, such as EXAFS [13], are needed for unambiguous identification of defects.

COMPLEXES OF DEFECTS

An interesting and very useful feature of the theory of deep levels as applied to defect complexes is that the level spectrum of a complex is very nearly the sum of the spectra of the complex's constituents. This rule of thumb obtained by Sankey [14] is not rigorous but is sufficiently valid that in many cases (i) it is virtually impossible to determine from energy levels alone whether the responsible defects are isolated or complexed, and (ii) when trying to identify defects responsible for specific deep levels it is often an adequate approximation to consider only point defects, while recalling that the point defects so identified might be complexed.

To see why complexing does not alter deep levels much, consider two nearest-neighbor sp^3-bonded impurities and their deep levels. The resulting "diatomic molecule" (Fig. (8)) will have (i) the σ-like (a_1) states formed from the s-like (A_1) states of the individual "atoms" and from the p-like (T_2) states polarized along the spine of the molecule; and (ii) the π-like (e-symmetric) states derived from the T_2 levels that are polarized perpendicular to the spine of the molecule. The π-like states do not "feel" the spine of the molecule and so have energies virtually identical to those of the constituent defects. The σ-like states (to a good approximation) occur at energies that interlace the A_1 and T_2 levels of the constituent defects. Because these levels are often within several tenths of an eV of one another,

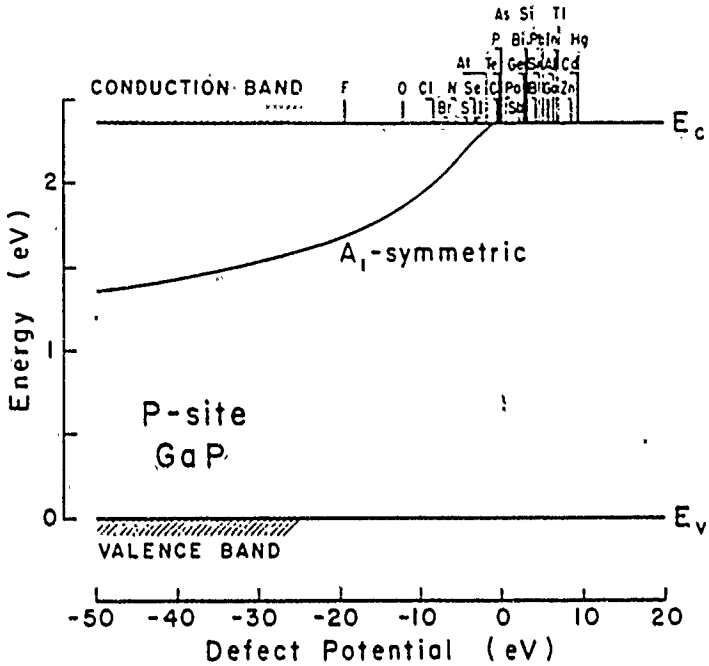


Fig. [7]. Deep energy levels predicted in [1] versus defect potential for the impurties at the top of the figure for A_1-symmetric deep levels of substitutional defects on the P-site of GaP. Note that the curve approaches an asymptote.

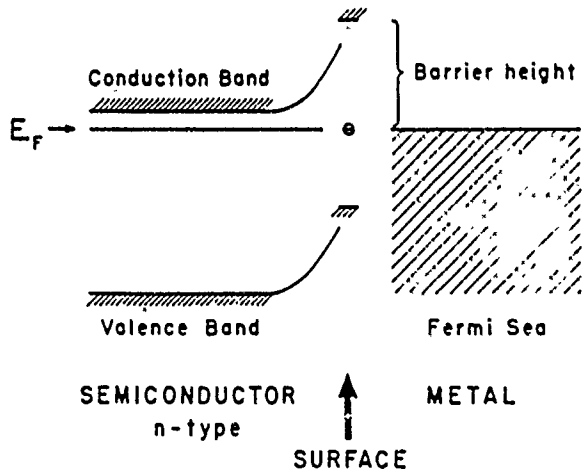


Fig. [8]. Illustrating how a surface defect level determines the Fermi level and leads to band-bending and Schottky barrier formation at a semiconductor/metal contact, as discussed by Bardeen [17].

the resulting s-like level structure will be similar to that of the constituent defects to within a few tenths of an eV.

SCHOTTKY BARRIERS

These ideas about complexing are especially useful for discussing Schottky barrier formation due to native defects in semiconductors, because they greatly reduce the number of native defects that must be considered. A defect at a semiconductor surface will have four deep levels. Near the fundamental band gap the p-like T_2 levels will not be degenerate, but will be split by the reduced symmetry of the surface. The surface-induced shift of the bulk A_1 level and splitting of the T_2 level will be of order 0.5 eV, and can cause a resonant level to descend into the gap. Thus an impurity that is shallow in the bulk may be deep at the surface, with levels in the gap.

In general, surface defect deep level spectra may be several tenths of an eV different from bulk deep levels of the same impurity. They will be similar to the spectra of bulk defect-vacancy pairs, however (because the surface is effectively a sheet of "vacancies": an impurity at a surface is surrounded by three host atoms and one "vacancy," with the other "vacancies" more distant). In any case, Allen [15] has extended the ideas of Hjalmarson, Vogl, and Sankey to treat deep levels of surface defects, following pioneering work by Daw, Smith, Swarts, and McGill [16].

Bardeen [17] and Spicer [18] recognized (i) that the lowest empty deep level of a neutral surface defect is effectively the Fermi energy for electrons at the surface, and (ii) that, in electronic equilibrium, the Fermi energies of the bulk semiconductor, bulk metal, and semiconductor surface must all align for a semiconductor/metal contact. As a result the bands bend, forming a Schottky barrier, in order to accommodate this alignment. The Schottky barrier height is the energy of the conduction band edge with respect to the lowest empty surface deep level (Fig. (9)).

In Fig. (10) we show the resulting predictions of Schottky barrier height versus alloy composition for III-V semiconductor alloys, assuming that the surface defect responsible for Schottky barrier formation is the cation on anion-site native antisite defect [19]. The agreement with the data is better than the accuracy of the theory.

This theory not only explains Schottky barrier formation in III-V alloys, it also explains the apparent dependence of barrier height on metal reactivity [20,21], the Schottky barriers formed when transition metals are deposited on Si [22], Ge, and diamond, and numerous once-puzzling features of the Schottky barrier data [20].

THE THEORY

The theory of deep levels [1] is a one-electron theory, for a host Hamiltonian operator H_0 with a defect perturbation operator V. The secular equation is

$$\det \{1 - (E-H_0)^{-1} V\} = 0 \qquad (1)$$

or

$$\det \{1 - \int dE' \, (E-E')^{-1} \, \delta(E'-H_0) V\} = 0 \qquad (2)$$

All current theories solve such an equation for the deep level E. Different

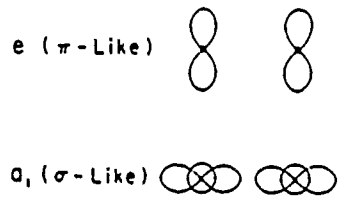


Fig. (9). Illustrating the π-like and σ-like orbitals of a complex of neighboring defects.

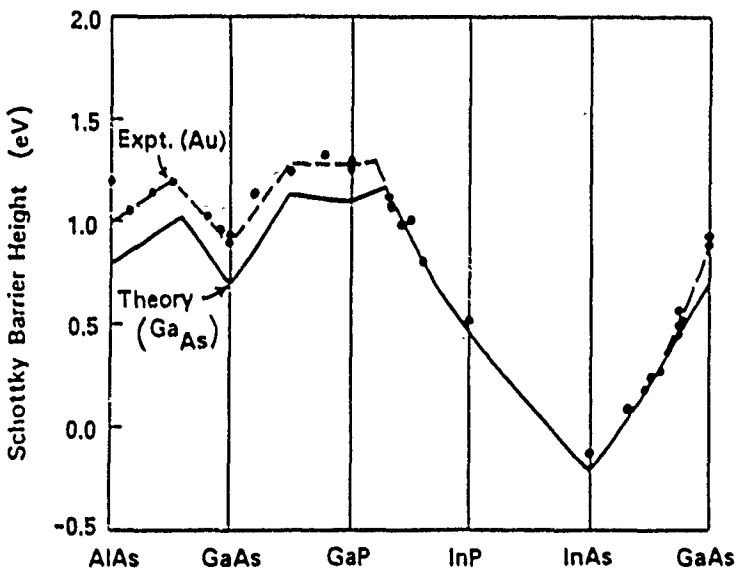


Fig. (10). Predicted Schottky barrier heights versus alloy composition for III-V alloys, compared with data involving Au contacts, after [19] and [23].

theories treat the spectral density operator $\delta(E-H_0)$ and the defect potential operator V differently. The sign of the contribution of the spectral density to the integral in Eq. (2) is different if E is in the conduction band or valence band: the conduction band states push the deep level down and the valence band states repel it upward — the balance between these opposing forces determines E. For a theory of deep levels to be successful this balance must be correct. All theories represent valence band states well; thus the issue is the position of the host band gap and spectral distribution of the host conduction band states. Hjalmarson solved this problem by using the empirical tight-binding model of Vogl et al. [4] for H_0. This model, by construction, has the correct band gap and a spectral distribution that gives good deep levels. Local-density theories calculate the band gap, invariably find it to be in error by typically 50%, and then adjust either the gap or the pseudopotential in an ad hoc manner. The resulting theories have comparable validity to the Hjalmarson theory for deep level spectra, but if the pseudopotential theory is based on a local density approximation, it should have superior predictions for total energies. Thus, when we compute deep level spectra, especially for complex defects, we use the Hjalmarson theory; but for diffusion and total energy calculations we recommend a local density approach.

The Hjalmarson theory is easy to evaluate, and pedagogical discussions of how to do so are available [23].

We are grateful to the Office of Naval Research for their continuing support (Contract No. N00014-84-K-0352) and to our many colleagues who have provided the ideas discussed here.

References

[1] H. P. Hjalmarson, P. Vogl, D. J. Wolford and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980), to be published.
[2] M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969).
[3] L. A. Hemstreet, Phys. Rev. B 15, 834 (1977); P. Vogl, to be published; M. Scheffler, J. P. Vigneron and G. B. Bachelet, Phys. Rev. Lett. 49, 1765 (1982); J. Bernholc, N. O. Lipari, S. T. Pantelides and M. Scheffler, Phys. Rev. B 26, 5706 (1982); M. Jaros and S. Brand, Phys. Rev. B 14, 4494 (1976); U. Lindefelt, J. Phys. C 12, L419 (1979); D. A. Papaconstantopoulos and E. N. Economou, Phys. Rev. B 22, 2903 (1980); G. A. Baraff and M. Schlüter, Phys. Rev. Lett. 41, 892 (1978); Phys. Rev. B 19, 4965 (1979); J. Bernholc and S. T. Pantelides, Phys. Rev. B 18, 1780 (1978); J. Bernholc, N. O. Lipari and S. T. Pantelides, Phys. Rev. Lett. 41, 895 (1978); Phys. Rev. B 21, 3545 (1980); J. Bernholc, S. T. Pantelides, N. O. Lipari and A. Baldereschi, Solid State Commun. 37, 705 (1981).
[4] P. Vogl, H. P. Hjalmarson and J. D. Dow, J. Phys. Chem. Solids, 44, 365 (1983).
[5] D. J. Wolford, W. Y. Hsu, J. D. Dow, and B. G. Streetman, J. Lumin. 18/19, 863 (1979).
[6] W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).
[7] B. A. Bunker, S. L. Hulbert, J. P. Stott, and F. C. Brown, Phys. Rev. Lett. 53, 2157 (1984).
[8] K. E. Newman and J. D. Dow, Phys. Rev. B 30, 1929 (1984).
[9] G. W. Ludwig, Phys. Rev. 137, A 1520 (1965).
[10] S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).

[11] W. A. Harrison, Electronic Structure and the Properties of Solids (San Francisco, CA, 1980); W. A. Harrison, in Festkörperprobleme, Vol. 17, edited by J. Treusch (Wiesbaden, 1977), p. 135. 2409 (1969).
[12] S. Y. Ren, J. D. Dow, and D. J. Wolford, Phys. Rev. B 25, 7661 (1982).
[13] B. A. Bunker, S. M. Heald, and J. Tranquada, Proc. 3rd International EXAFS Conference, Stanford, CA, 1984 (Springer-Verlag).
[14] O. F. Sankey, H. P. Hjalmarson, J. D. Dow, D. J. Wolford and B. G. Streetman, Phys. Rev. Lett. 45, 1656 (1980); O. F. Sankey and J. D. Dow, Appl. Phys. Lett. 38, 685 (1981); J. Appl. Phys. 52, 5139 (1981); Phys. Rev. B 26, 3243 (1982). See also J. D. Dow and O. F. Sankey, Am. Inst. Phys. Conf. Proc., Vol. 73, edited by R. A. Street, D. K. Beigelsen and J. C. Knights (New York, NY, 1981), p. 141.
[15] R. E. Allen and J. D. Dow, Phys. Rev. B 24, 911 (1981); R. E. Allen and J. D. Dow, Phys. Rev. B 25, 1423 (1982 .
[16] M. S. Daw and D. L. Smith, Phys. Rev. B 20, 5150 (1979); J. Vac. Sci. Technol. 17, 1028 (1980); Appl. Phys. Lett. 36, 690 (1980); Solid State Commun. 37, 205 (1981); M. S. Daw, D. L. Smith, C. A. Swarts and T. C. McGill, J. Vac. Sci. Technol. 19, 508 (1981).
[17] J. Bardeen, Phys. Rev. 71, 717 (1947).
[18] W. E. Spicer, P. W. Chye, P. R. Skeath, C. Y. Su, and I. Lindau, J. Vac. Sci. Technol. 16, 1422 (1979), and references therein; W. E. Spicer, I. Lindau, P. R. Skeath and C. Y. Su, J. Vac. Sci. Technol. 17, 1019 (1980), and references therein.
[19] R. E. Allen, T. J. Humphreys, and J. D. Dow, J. Vac. Sci. Technol. B 2, 449 (1984).
[20] J. D. Dow, R. E. Allen, and O. F. Sankey, Chem. and Phys. of Solid Surf. V, 483 (1984), ed. R. Vanselow and R. Howe, Springer Series in Chem. Phys. 35.
[21] J. D. Dow and R. E. Allen, J. Vac. Sci. Technol. 20, 659 (1982).
[22] O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1984).
[23] J. D. Dow, 1983 Varenna Summer School Lectures, Italian Physical Society, in press, and unpublished results.

Surface Science 168 (1986) 376–385
North-Holland, Amsterdam

THEORETICAL INTERPRETATION OF SCHOTTKY BARRIERS AND OHMIC CONTACTS

Roland E. ALLEN
Department of Physics, Texas A&M University, College Station, Texas 77843, USA

Otto F. SANKEY
Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

and

John D. DOW
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA

Received 10 June 1985; accepted for publication 23 July 1985

We review a theory of Schottky barriers that explains the following experimental findings: (i) observed barrier heights and Fermi-level pinning positions for GaAs, InP, GaSb, AlAs, GaP, InAs, and other III–V semiconductors; (ii) switching of the observed barrier heights and Fermi-level pinning positions for III–V semiconductors as a function of surface treatment or reactivity of the metal; (iii) alloy dependence of Schottky barrier heights for the ternaries $Al_{1-x}Ga_xAs$, $GaAs_{1-z}P_z$, $Ga_{1-y}In_yP$, $InP_{1-z}As_z$, and $In_{1-x}Ga_xAs$; (iv) different slopes dE/dx for different metal contacts to AlGaAs, and an apparent cusp in the slope for Al contacts as a function of alloy composition; (v) observed Schottky barriers for a wide variety of Si/transition-metal-silicide interfaces; (vi) observed barriers for Ge, diamond, and amorphous Si; (vii) observation that Fermi-level pinning for p-GaAs disappears at the annealing temperature of the antisite defect As_{Ga}. The theory provides a microscopic realization of the phenomenological defect model of Spicer, Lindau and coworkers. We find that most Schottky barriers are explained by dangling bonds – *intrinsic dangling bonds* for group IV semiconductors and *antisite* (as well as intrinsic) *dangling bonds* for III–V semiconductors. Ohmic contacts are explained in the present picture by shallow levels, which are also predicted by the theory.

Ohm's law usually fails to hold at a semiconductor/metal interface. Instead, the current density J is observed to depend exponentially on the applied voltage V. Observations of this phenomenon extend back more than a century, but the underlying microscopic causes are still controversial. Proposed theoretical interpretations include the following:
(1) Schottky's original model, in which equilibration of chemical potentials of semiconductor and metal requires charge transfer, resulting in an electro-

0039-6028/86/$03.50 © Elsevier Science Publishers B.V.
(North-Holland Physics Publishing Division)

static dipole layer at the interface and a barrier to the motion of electrons.
(2) Bardeen's Fermi-level pinning model [1] in which "surface states" of some kind result in a barrier that is nearly independent of the metal for a given semiconductor. Fermi-level pinning can be produced by intrinsic surface states, metal-induced gap states [2], and extrinsic states associated with various kinds of defects. (One expects, of course, that the barrier height will vary with the metal in the case of metal-induced gap states or metal-atom impurity states.)
(3) The defect model of Spicer, Lindau and coworkers [3–5], in which the Fermi-level pinning is produced by native defects of some kind, associated with missing anion or cation atoms in III–V semiconductors.

Here we review a theory of Schottky barriers and Ohmic contacts that is based on Fermi-level pinning and the defect model, in which observed barrier heights and Fermi-level pinning positions are assigned to particular defects [6–15].

The electron Schottky barrier height is the difference between the conduction band edge and the Fermi energy at the semiconductor surface:

$$\phi_B^n = E_C^S - E_F^S.$$

Within the context of the defect model, therefore, a microscopic theory of Schottky barriers is a theory of the relevant defect levels that "pin" the Fermi energy at the surface. For a sufficiently high concentration of defects, the surface Fermi energy will lie near the lowest acceptor level for an n-type semiconductor, or the highest donor level for a p-type semiconductor. (Further discussion is given in, e.g., refs. [6–14].)

In fig. 1, we show our calculated results for the defect levels associated with antisite defects and vacancies at the (110) surfaces of nine III–V semiconductors. Many of these levels have been previously reported [6–14]. They were calculated using the same scheme applied to bulk defects by Vogl, Hjalmarson and Dow [16,17], employing the measured surface relaxation [18,19]. In the following, we will compare the theoretical results of fig. 1 with the available observations for Schottky barriers, Fermi-level pinning, and Ohmic contacts on III–V semiconductors.

It should be emphasized that a variety of arguments is used in making our microscopic assignments of the defects responsible. For example, in the case of III–V semiconductors, we regard antisite defects as ordinarily more likely than vacancies for reasons that have been given elsewhere [6]. Also, as mentioned below, the variation in barrier height with alloy composition x can act as a signature of the defect type. Various experiments provide other valuable information concerning the identity of the defects; one example is the observation of an annealing temperature for the Fermi-level pinning on p-GaAs [21], mentioned below. In short, our assignments are based on qualitative arguments, chemical trends, and experimental information of various kinds, in

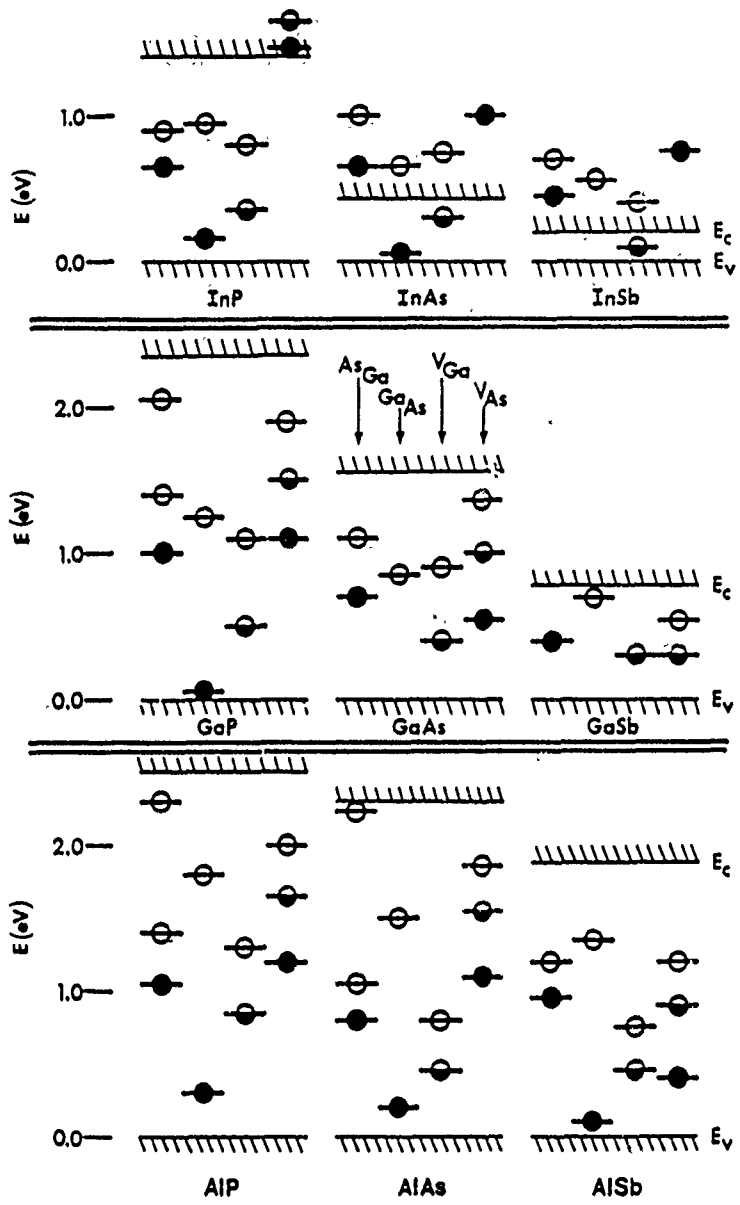


Fig. 1. Deep levels calculated for antisite defects and vacancies at (110) surfaces of III–V semiconductors. As indicated for GaAs, the levels are, left to right, for anion-on-cation-site, cation-on-anion-site, cation vacancy, and anion vacancy. The occupancy of the levels for the neutral charge state is shown; a full circle indicates the level contains two electrons (spin up and down), a half-full circle one electron, and an open circle no electrons. (Charge-state splittings are neglected.) For the In–V materials at the top, several resonances above the conduction band edge are also shown. E_V and E_C are the valence and conduction band edges. Some of these levels have been previously reported [6–14].

addition to detailed calculations of the defect levels.

Fig. 2 shows the experimental levels for Ge on GaAs(110) obtained by Mönch and Gant [20–22], which are approximately the same as those obtained by Spicer, Lindau and coworkers for various metals and oxygen on the same surface [3–5]. It appears that the acceptor level for the "cleavage-related" defect, and the acceptor and donor levels for the "chemisorption-related" defect, are quite satisfactorily explained by the corresponding levels for the antisite defects, Ga_{As} and As_{Ga}. The theoretical acceptor levels for both defects are associated with dangling bond orbitals – *antisite Ga and As dangling bonds*. The donor level is derived from a *bulk* As_{Ga} *donor level*, only shifted in energy somewhat at the surface.

Fig. 3 provides a similar comparison for n-type InP [26]. In this case, we again invoke *antisite In or P dangling bonds* to explain the experimental data for the noble metals, but *intrinsic Ga dangling bonds* (associated with a P vacancy) to explain the data for the more reactive metals. Various surface treatments are also interpreted as producing intrinsic Ga dangling bonds, or possibly surface impurities. The shallow donor level for V_P is explained by the "deep resonance" for this defect in fig. 1: Two electrons that would lie in this conduction band resonance spill out into the conduction band. These electrons are then bound to the doubly-charged defect site in shallow donor states, ≈0.1 eV below the conduction band edge. According to ref. [23] however, reactive metals can also produce the deeper Fermi energy ≈0.5 eV below the conduction band edge.

In fig. 4, we compare the experimental Schottky barrier heights for Au contacts to various semiconducting alloys with the theoretical barrier heights predicted for antisite cation dangling bonds. (Recall that $\phi_B^n = E_C^S - E_F^S$, where $E_F^S \approx$ defect level.) The agreement is more than satisfactory.

In fig. 5, we compare experimental barrier heights for Al and In contacts to $Al_{1-x}Ga_xAs$ with the theoretical heights predicted for antisite cation and anion dangling bonds. (The surface Fermi energy E_F^S, inferred from $\phi_B^n = E_C^S - E_F^S$, is plotted.) There appears to be a switching of defects as the alloy composition x is varied, with antisite cation dangling bonds providing a good description of the data for small x (large slope in dE/dx) and antisite anion dangling bonds doing equally well for large x (small dE/dx). The slope dE/dx thus appears to act as a signature of the defect type.

Other experimental observations also support the defect model and the present theory [3–5,20–33]. For example, fig. 6 shows further experimental evidence that the Schottky barriers for GaAs are associated with antisite defects: Mönch and coworkers [21] find that the Fermi-level pinning position for Ge on GaAs(110) anneals out at the annealing temperature of bulk As_{Ga}. As mentioned above, the pinning position for Ge on GaAs(110) is essentially the same as that for various metals on this surface.

Other proposed mechanisms, such as metal-induced gap states [2,34–42]

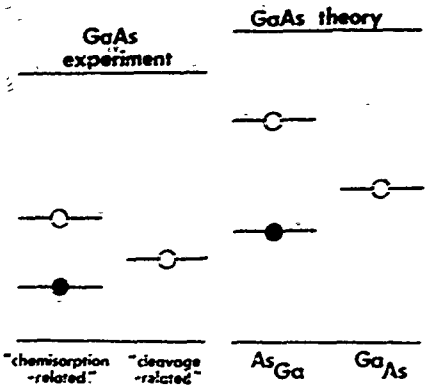


Fig. 2. Theoretical levels for surface antisite defects. As on Ga site (As_{Ga}) and Ga on As site (Ga_{As}), compared with experimental levels for "chemisorption-related" defects and "cleavage-related" defects [5, 20]. Open circles indicate theoretical acceptor levels or experimental Fermi-level pinning positions on n-type GaAs; solid circles indicate theoretical donor levels or experimental Fermi-level pinning positions on p-type GaAs. After refs. [6, 9].

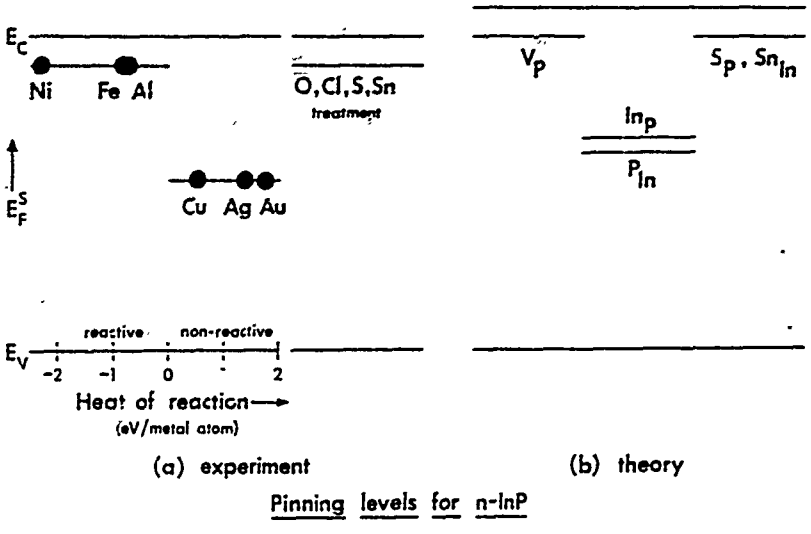


Fig. 3. Theoretical levels for surface antisite defects, In on P site and P on In site (coincidentally close in energy), compared with experimental Fermi-level pinning positions E_F^S for the noble metals Cu, Ag, and Au on n-type InP [26]. As can be seen in fig. 1, a surface P vacancy (V_P) produces a shallow donor level that can explain the quite different experimental Fermi-level pinning position for the reactive metals Ni, Fe, and Al [26]. Surface treatments with O, Cl, S, and Sn also are observed to produce Fermi-level pinning just beneath the conduction band edge [26]; such treatments may produce surface vacancies (V_P) or surface impurities (S_P, Sn_{In}). In terms of dangling bonds, we attribute E_F^S for the noble metals to *antisite dangling bonds* and E_F^S for the more reactive metals to *intrinsic dangling bonds*. After refs. [8, 9].

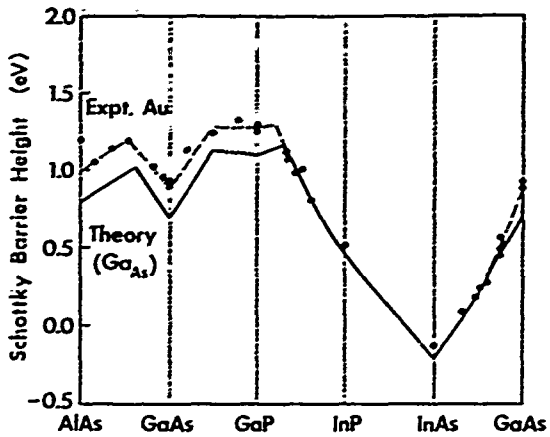


Fig. 4. Theoretical Schottky barrier height for cation-on-anion antisite defects (e.g. Ga_{As}) compared with experimental Schottky barrier height for Au contacts to the alloys $Al_{1-x}Ga_xAs$, $GaAs_{1-x}P_x$, $Ga_{1-x}In_xP$, $InP_{1-x}As_x$, and $In_{1-x}Ga_xAs$. The sources of the experimental data (dots, fitted in ref. [21] by the dashed line), are cited in ref. [10]. The agreement between theory and experiment, particularly on the right, is fortuitously good since the theory contains errors of several tenths of an eV. It is clear, however, that the theory does a quite satisfactory job of describing the behavior of the measured Schottky barrier as a function of alloy composition. After ref. [10].

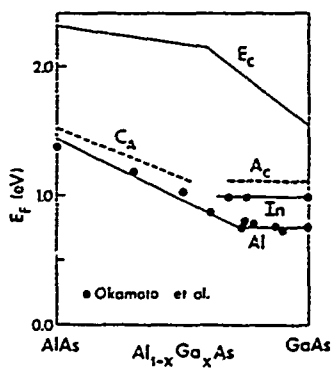


Fig. 5. Theoretical deep levels for cation-on-anion-site (C_A) and anion-on-cation-site (A_C) compared with experimental Fermi-level pinning position (E_F) inferred from Schottky barrier measurements of Okamoto et al. [53]. Here E_F^x is defined by $\varphi_B^n = E_C^x - E_F^x$, where φ_B^n is the measured Schottky barrier height and E_C is the conduction band edge. There appears to be a switching of defects as a function of alloy composition, with anion-on-cation-site explaining the data for In and Al contacts to Ga-rich alloys and cation-on-anion-site explaining the data for Al-rich alloys. Notice that the slope of the defect level as a function of the alloy composition x is a signature of the defect type. This is an example of the fact that the present theory [6–14] provides a microscopic explanation of the observed chemical trends of the Schottky barrier height. After ref. [14].

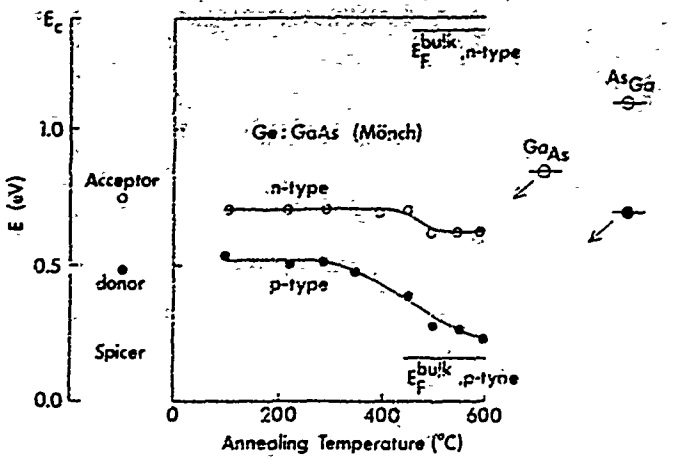


Fig. 6. Annealing out of Fermi-level pinning for Ge on p-type GaAs at the annealing temperature of the As_{Ga} antisite defect. After ref. [21]. In the present theory, As_{Ga} is annealed out, but the acceptor level for Ga_{As} remains. The acceptor and donor levels previously found by Spicer et al. [3–5] are indicated by the open and solid circles on the left.

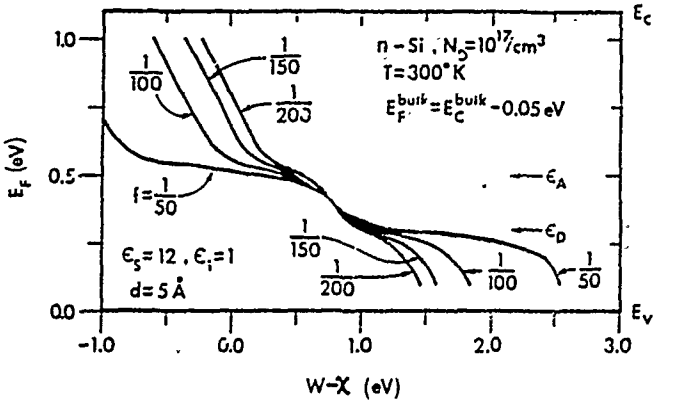


Fig. 7. Surface energy E_F^S at a semiconductor/metal interface as a function of $W - \chi$, where W is the metal's work function and χ is the semiconductor's electron affinity. Notice that one defect of every 50 or even every 100 semiconductor surface atoms produces good Fermi-level pinning for a substantial range of metals. This figure is for n-type Si, with 10^{17} bulk donors/cm^3 and room temperature. One surface acceptor and one surface donor level were assumed, as well as an interfacial transition region separating the metal and semiconductor of width $d = 5$ Å and dielectric constant $\epsilon_i = 1$. The present results are appropriate to surface defects, whereas those of ref. [44] are appropriate to bulk defects.

have great difficulty explaining results like those of fig. 6. Metal-induced gap states cannot explain Fermi-level pinning at submonolayer coverages [3–5, 20–23, 26, 31], or real world Schottky barriers with an oxide layer between the semiconductor and the metal. A more extensive case against metal-induced gap states as the mechanism for observed Schottky barriers is presented in ref. [12b].

We mention that the present theory assumed a concentration of relevant defects of the order of one defect per 100 surface atoms, in order for the Schottky barrier to be approximately the same for different metals. (A smaller defect concentration is required to explain Fermi-level pinning for submonolayer coverages [21].) In fig. 7, we illustrate that a 1/100 concentration can be quite sufficient for *surface* defects. (The calculations of fig. 7 are for $\varepsilon_i = 1$ and $d = 5$ Å, where ε_i is the effective dielectric constant and d the effective width of the dipole layer at the interface; the true values of these quantities are unknown, of course, but we believe these are not unreasonable estimates.) Similar results were obtained by Spicer et al. [43] and Bardeen [1]. On the other hand, Zur et al. [44] find that an order of magnitude higher concentration is required for *bulk* defects. The present theory is, of course, based on dangling bonds at the surface.

The results of figs. 1–6 are for III–V semiconductors. In the case of group IV semiconductors [45–52] – Si, Ge, diamond, and amorphous Si – we inter-

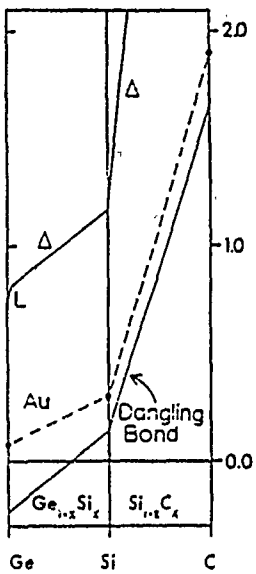


Fig. 8. Theoretical intrinsic dangling bond levels for Ge, Si, and C (diamond) compared with experimental surface Fermi-level positions inferred from Schottky barrier heights for Au contacts on these materials. Energies are measured relative to the valence band maximum E_V; the conduction band minimum at the L or Δ point in the Brillouin zone is shown at the top. After ref. [12b].

pret the observed Schottky barriers in terms of *intrinsic dangling bonds*. Representative results are shown in fig. 8, for Au contacts. We find the same level of agreement with the observed Schottky barrier heights for numerous Si/transition-metal-silicide interfaces [11], for W and Ni germanides interfaced with Ge and for Al and Ba on diamond [12a], and for various transition metals on amorphous Si [13].

In summary, the present theory explains a wide variety of experimental observations for Schottky barriers, Fermi-level pinning, and Ohmic contacts. For III–V semiconductors, the primary pinning agents appear to be antisite dangling bonds. For group IV semiconductors, the data are well explained by intrinsic dangling bonds. In some cases, other mechanisms appear to be involved. For example, the original Schottky mechanism appears to apply for noble metals on GaSe [26], where the concentration of defects is presumably low. Also, the prominent acceptor level for GaSb [3–5] may be due to the complex (Ga_{Sb}, V_{Ga}) [3, 7]. Finally, as described above, we attribute Fermi-level pinning on p-GaAs to a donor level that is bulk derived and only somewhat shifted in energy at the surface. However, the present theory mainly attributes Schottky barriers, Fermi-level pinning, and nearly Ohmic contacts to interfacial dangling bonds of one kind or another.

We thank the US Office of Naval Research for their support, which made this work possible (N00014-82-K-0447).

References

[1] J. Bardeen, Phys. Rev. 71 (1947) 717.
[2] V. Heine, Phys. Rev. A138 (1965) 1689.
[3] W.E. Spicer, P.W. Chye, P.R. Skeath, C.Y. Su and I. Lindau, J. Vacuum Sci. Technol. 16 (1979) 1422, and references therein.
[4] W.E. Spicer, I. Lindau, P.R. Skeath and C.Y. Su, J. Vacuum Sci. Technol. 17 (1980) 1019, and references therein.
[5] W.E. Spicer, I. Lindau, P.R. Skeath, C.Y. Su and P.W. Chye, Phys. Rev. Letters 44 (1980) 520.
[6] R.E. Allen and J.D. Dow, Phys. Rev. B25 (1982) 1423.
[7] R.E. Allen, H.P. Hjalmarson and J.D. Dow, Solid State Commun. 41 (1982) 419.
[8] J.D. Dow and R.E. Allen, J. Vacuum Sci. Technol. 20 (1982) 659.
[9] R.E. Allen, R.P. Beres and J.D. Dow, J. Vacuum Sci. Technol. B1 (1983) 401.
[10] R.E. Allen, T.J. Humphreys, J.D. Dow and O.F. Sankey, J. Vacuum Sci. Technol. B2 (1984) 449.
[11] O.F. Sankey, R.E. Allen and J.D. Dow, Solid State Commun. 49 (1984) 1.
[12] (a) O.F. Sankey, R.E. Allen and J.D. Dow, J. Vacuum Sci. Technol. B2 (1984) 491; (b) O.F. Sankey, R.E. Allen, S.F. Ren and J.D. Dow, J. Vacuum Sci. Technol., in press.
[13] O.F. Sankey, R.E. Allen and J.D. Dow, in: Proc. 17th Intern. Conf. on the Physics of Semiconductors, Eds. D.J. Chadi and W.A. Harrison (Springer, New York, 1985).
[14] J.D. Dow, O.F. Sankey and R.E. Allen, to be published.

[15] See also M.S. Daw and D.L. Smith, Phys. Rev. B20 (1979) 5150.
[16] P. Vogl, H.P. Hjalmarson and J.D. Dow, J. Phys. Chem. Solids 44 (1983) 44.
[17] H.P. Hjalmarson, P. Vogl, D.J. Wolford and J.D. Dow, Phys. Rev. Letters 44 (1980) 810.
[18] A. Kahn, E. So, P. Mark and C.B. Duke, J. Vacuum Sci. Technol. 15 (1978) 580.
[19] S.Y. Tong, A.R. Lubinsky, B.J. Mrstik and M.A. Van Hove, Phys. Rev. B17 (1978) 3303.
[20] W. Mönch and H. Gant, Phys. Rev. Letters 48 (1982) 512.
[21] W. Mönch, Surface Sci. 132 (1983) 92. This paper, and refs. [26, 29], also appear in Surfaces and Interfaces: Physics and Electronics, Ed. R.S. Bauer (North-Holland, Amsterdam, 1983).
[22] W. Mönch, in: Chemistry and Physics of Solid Surfaces V, Eds. R. Vanselow and R. Hove (Springer, New York, 1984) p. 501.
[23] T. Kendelewicz, N. Newman, R.S. List, I. Lindau and W.E. Spicer, J. Vacuum Sci. Technol. B3 (1985) 1206.
[24] F. Chekir and C. Barret, Surface Sci. 168 (1986) 838.
[25] P. Muret and A. Deneuville, Surface Sci., in press; 168 (1986) 830.
[26] R.H. Williams, Surface Sci. 132 (1983) 122, and references therein.
[27] H.H. Wieder, Inst. Phys. Conf. Ser. 50 (1980) 234.
[28] H.H. Wieder, Appl. Phys. Letters 38 (1981) 179.
[29] H.H. Wieder, Surface Sci. 132 (1983) 390, and references therein.
[30] L.J. Brillson, Surface Sci. Rept. 2 (1982) 123, and references therein.
[31] A.D. Katnani, G. Margaritondo, R.E. Allen and J.D. Dow, Solid State Commun. 44 (1982) 1231.
[32] R.W. Grant, J.R. Waldrop, S.P. Kowalczyk and E.A. Kraut, J. Vacuum Sci. Technol. 19 (1981) 477.
[33] Many other interesting results are reported in the papers of the present proceedings.
[34] J.C. Inkson, J. Phys. C6 (1973) 1350.
[35] S.G. Louie and M.L. Cohen, Phys. Rev. Letters 35 (1975) 866.
[36] E. Louis, F. Yndurain and F. Flores, Phys. Rev. B13 (1976) 4408.
[37] S.G. Louie, J.R. Chelikowsky and M.L. Cohen, Phys. Rev. B15 (1977) 2154.
[38] C. Tejedor, F. Flores and E. Louis, J. Phys. C10 (1977) 2163.
[39] F. Guinea, J. Sanchez-Dehesa and F. Flores, Phys. Rev. B16 (1983) 6499.
[40] J. Sanchez-Dehesa and F. Flores, Solid State Commun. 50 (1984) 29.
[41] J. Tersoff, Phys. Rev. Letters 52 (1984) 465; Surface Sci. 168 (1986) 275.
[42] G. Platero, J.A. Vergés and F. Flores, Surface Sci. 168 (1986) 100.
[43] W.E. Spicer, S. Pan, D. Mo, N. Newman, P. Mahowald, T. Kendelewicz and S. Eglash, J. Vacuum Sci. Technol. A2 (1984) 476.
[44] A. Zur, T.C. McGill and D.L. Smith, J. Vacuum Sci. Technol. B1 (1983) 608.
[45] See, for example, Thin Films and Interfaces, Eds. P.S. Ho and K.N. Tu (North-Holland, New York, 1982). Reprinted from Thin Solid Films 93 (1983).
[46] P.S. Ho and G.W. Rubloff, Thin Solid Films 89 (1982) 433.
[47] P.S. Ho, M. Liehr, P.E. Schmid, F.K. Legoues, E.S. Yang, H.L. Evans and X. Wu, Surface Sci. 168 (1986) 184.
[48] E.H. Rhoderick, Metal-Semiconductor Contacts (Clarendon, Oxford, 1978).
[49] J.M. Andrews and J.C. Philips, Phys. Rev. Letters 35 (1975) 56.
[50] G. Ottaviani, K.N. Tu and J.W. Mayer, Phys. Rev. B24 (1981) 3354.
[51] C.A. Mead and W.G. Spitzer, Phys. Rev. A134 (1964) 713.
[52] D. Cherns, G.R. Anstis, J.L. Hutchinson and J.C.H. Spence, Phil. Mag. A46 (1982) 849.
[53] K. Okamoto, C.E.C. Wood and L.F. Eastman, Appl. Phys. Letters 38 (1981) 636.

SUPERCOMPUTERS IN SOLID STATE PHYSICS

John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556

I. Introduction

The role of supercomputers in theoretical solid state physics has been revolutionized by the development of high-speed computers and software that facilitates their use. In less than a generation, the computer available to the average physicist has changed from a cumbersome mechanical calculator that could barely divide to an electronic wizard that can compute and display million-pixel pictures continuously. These advances in numerical computation have changed the character of theoretical physics from a primarily analytic discipline to a field now dominated by numerical calculations, simulations, and pattern recognition.

Atomic physics was the first area to feel the impact of computers: codes for solving the Hartree-Fock equations with configuration interaction almost rendered approximate analytic theories of atomic structure obsolete by the early 1960's. Solid state physics was also dramatically affected by the computer revolution, but the complexity of many-atom solids appeared, until recently, to be an insurmountable obstacle to a fully numerical approach to the structures of solids. As a result, three separate types of solid state theory have developed: (1) analytic theory, of which the theory of superconductivity [1] is a prime example; (2) numerical theory, such as energy band calculations [2]; and (3) simulations, such as Monte Carlo calculations of electron transport in semiconductors [3]. High-speed computers have created situations such that the analytic theories can be evaluated for increasingly realistic models of solids, the numerical theories can be extended to increasingly complex solids, and realistic simulations are now possible. In fact, the theory of electronic structure of simple solids is now becoming genuinely predictive in the sense that it is capable of providing valence energy band structures, based on either a priori theories or well-established semi-empirical prescriptions. High-speed computers will certainly cause this predictive capability to be extended in the near future to more complex solids. Moreover, increasingly complex numerical theories that use the full capabilities of supercomputers are likely to be developed -- extending the predictive capabilities of electronic structure theory to conduction band structures and to excited states of solids. Thus the advent of supercomputers guarantees a revolution in solid state theory.

The purpose of this paper is to speculate how the supercomputer revolution in its early days will affect solid state theory and to sketch a picture of how the supercomputer and its less agile cousin, the laboratory computer [4], will become tools of the average theorist.

II. Optimal Use of Supercomputers

Supercomputers are extremely fast and somewhat-complicated machines that are most often situated at remote locations. Software development for these machines is in a primitive state, and, as a rule, the machines are decidedly not user-friendly. For most users, supercomputers are at distant sites and have to be accessed over phone lines -- this restriction means that they are useful primarily for compute-limited rather than output-limited jobs. (A fully-utilized supercomputer will compute so many numbers that the output can be digested by humans only in graphical form -- and presently available communications are not well-suited to long distance transmission of large amounts of graphical output.)

The software and communications limitations imply that researchers interested in using supercomputers today should be prepared to spend a few months at the supercomputer site to learn how to "vectorize" their codes, or they should station a student at the site to interact directly with the supercomputer (and its high-speed graphics), or they should wait until the plans to develop user-friendly software and high-speed communications are further along. Hence, in the near future, the primary use of supercomputers by solid state physicists will be for large-scale production jobs on compute-limited programs. While it is always cost-effective in a narrow sense to use a supercomputer, because supercomputers offer the most megaflops per dollar, a proper cost-benefit analysis includes the considerable personnel time spent dealing with software and communications problems. In our experience, the benefits outweigh the costs when supercomputer usage approaches fifty hours of central processor time or when stationing a student at the supercomputer site becomes economically feasible. Therefore, today supercomputing is primarily for researchers who either (i) spend nearly full-time computing, (ii) work near a supercomputer site, (iii) have research budgets so small that they are forced to deal with the software and communications problems in order to obtain the lower computer costs, or (iv) have problems so large that laboratory or institutional computers cannot accommodate them. Users with lesser needs will find that personal computers with their superior software (e.g., a Compaq) or laboratory computers with their user-friendly and well-integrated software and hardware (e.g., a VAX, possibly with an array processor attached) offer cost-effective research computing with a minimum of frustration.

There are major efforts in progress to improve the user-friendliness of supercomputer software and to develop high-speed communications between supercomputers and remote installations. If these efforts fulfill their advertised promise, supercomputers will become much more useful and accessible for casual computations. Nevertheless, in the near future, most solid state physicists will perform most of their program development and debugging on personal or laboratory computers -- and will transfer large-scale production computation to supercomputers.

In the next decade we expect institutions such as research universities to develop balanced, integrated computer networks. The large universities will own supercomputers, while the smaller ones will have high-speed data links to supercomputer centers. Universities will continue to have computer and graphics centers for handling moderately heavy computation and for shifting the jobs that demand extremely fast computation to supercomputers. Individual researchers and research groups will continue to have laboratory computers dedicated to specific tasks and will use these computers to debug codes that are shipped to supercomputers for rapid execution. These laboratory computers will be replaced every five years or so, with the old computers being incorporated into university-wide networks designed primarily for student use. All lower levels of computer usage for research, such as personal computer usage, will also be incorporated into the same networks.

This picture of the near-term evolution of research computing into a stratified system, with different strata having different but complementary functions, is the basis for our projections of the needs for supercomputers in solid state physics.

III. Solid State Physics and Supercomputing

Solid state physics is an extremely broad field and there are many problems in the field that will be solved using supercomputers. Here we consider only two typical general problems: understanding the electronic and vibrational structures of semiconductors. We argue that some aspects of these problems are most efficiently solved using laboratory computers, especially when one accounts for the accessibility of laboratory computers in comparison with most supercomputers. Nevertheless, there are many interesting research questions that can be answered only with the help of a supercomputer -- and some that will require future-generation supercomputers. Therefore we strongly endorse efforts to upgrade the country's supercomputing capability, and to make supercomputers widely accessible to materials scientists and solid state physicists.

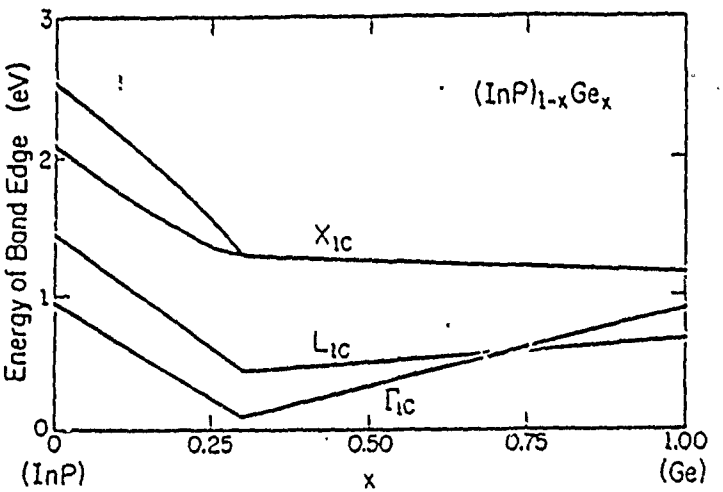


Fig. 1. Predicted electronic structure of the metastable subsitutional alloy $(InP)_{1-x}Ge_{2x}$, after Ref. [5].

a) Electronic Structure

Currently theorists can predict the electronic structures of solids with a high degree of confidence if they first know the sites of the constituent atoms. The calculations involved can be performed on a laboratory computer if there is some regularity or periodicity to the positions of the atoms and if the atomic geometries are not too complicated. For example, the band gap of a metastable substitutional crystalline alloy such as $(InP)_{1-x}Ge_{2x}$ can be calculated rather well [5]. Fig. 1, because this semiconductive alloy is highly covalent and similar to well-understood $(GaAs)_{1-x}Ge_{2x}$ [6]. The alloy effects can be treated in the virtual-crystal or "ama-gamated" approximation [7]: the interactions between atoms overpower small differences in the energies of different atoms, making the electronic structure insensitive to the details of alloy configurations or any slight lattice relaxation.

This pleasant situation fails to hold, however, (i) once charge-transfer between atoms becomes important (i.e., for ionic solids), or (ii) once atoms in an alloy cluster together, or (iii) in cases such that fluctuations in composition beyond mean-field theory become important, or (iv) when the fundamental unit cell of the solid is too large. For such problems, supercomputers will be needed. Furthermore, attempts to develop a fundamental thermodynamics of solids and to study electronically driven phase transitions or defect stabilized phases using actual electronic wavefunctions will require massive calculations -- on supercomputers.

b) Vibrational Structure

Predictions of the vibrational structure of solids, particularly in the spectral neighborhood of the optical phonons are very sensitive to details of short-range order and hence will require more computing power than corresponding electronic structure problems. In fact, the phonon spectrum of a simple random alloy, such as $(GaSb)_{1-x}Ge_{2x}$, is a superposition of lines associated with the different bonds (Fig. 2) [8], with the energy of each line depending on as many as hundreds of atoms nearest to the bond. To merely solve Newton's equations of motion for such a single cluster of, say, 1000 atoms requires diagonalization of a 3000 by 3000 matrix -- in the simplest model (short-ranged forces, rigid-ions).

There are simply too many different alloy configurations to include them all If one allows for any one of: anharmonic forces, long-ranged interactions, atomic polarization, off-site atomic distortions, finite temperatures, or clustered atomic configurations, the problem of calculating the phonon spectrum of this simple alloy for even a few alloy configurations becomes intractable even with today's supercomputers. This problem awaits future generations of supercomputers for its solution.

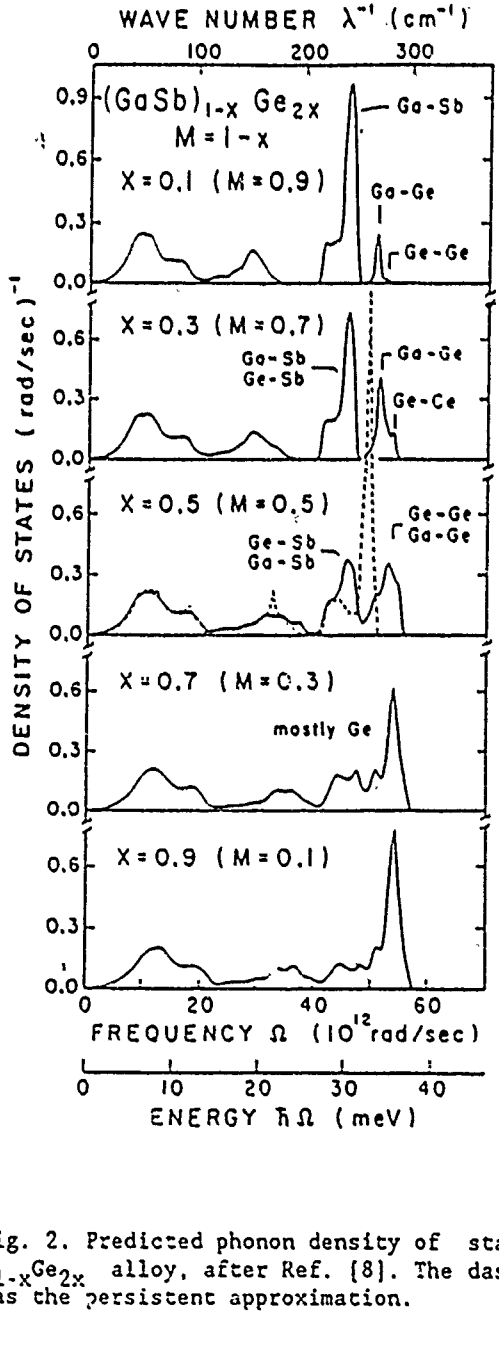


Fig. 2. Predicted phonon density of states of metastable $(GaSb)_{1-x}Ge_{2x}$ alloy, after Ref. [8]. The dashed line is what is known as the persistent approximation.

IV. Summary

Clearly supercomputers will have a profound effect on the level of theoretical understanding of solids, especially complex solids such as superlattices or microstructures. With increasing supercomputer speed, the solid state theorists will be able to treat increasingly complex structures in electronic materials -- and, before long, will be able to predict the properties of electronic devices. As the computing elements become smaller, faster, more complex, and governed by quantum rather than classical laws of physics, the laws of quantum mechanics implemented with supercomputers will become increasingly capable of predicting the properties of newer, smaller, and yet more powerful circuit elements. One of the great dreams of theoretical physics will be realized: Theory will be able to predict new man-made electronic device structures that, when fabricated, will have a profound effect on modern culture.

Acknowledgment -- We are grateful to the Office of Naval Research, Contract No. N00014-84-K-0352, for their support of the research on which this paper is based.

REFERENCES

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. 108, 1175 (1957).

[2] See, for example, Computational Methods in Band Theory, P. M. Marcus, J. F. Janak, and A. R. Williams, eds. (Plenum Press, New York, 1971).

[3] See, for example, D. Widiger, I. C. Kizilyalli, K. Hess, and J. J. Coleman, Superlattices and Microstructures 1, 465 (1985).

[4] Our definition of a laboratory computer includes any computer that can be the host for an array processor, including host-processor pairs.

[5] D. W. Jenkins, K. E. Newman, and J. D. Dow, J. Appl. Phys. 55, 3871 (1984).

[6] K. E. Newman and J. D. Dow, Phys. Rev. B 27, 7495 (1983).

[7] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968).

[8] A. Kobayashi, K. E. Newman, and J. D. Dow, Phys. Rev. B 32, 5312 (1985).

Effects of the band offset on interfacial deep levels

Richard P. Beres[a] and Roland E. Allen
Department of Physics, Texas A & M University, College Station, Texas 77843

John D. Dow
Department of Physics, University of Notre Dame,[b] Notre Dame, Indiana 46556 and Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

(Received 29 April 1987; accepted 11 September 1987)

The energy levels of antisite defects at a GaAs/Ge (110) interface are calculated and shown to be essentially unaltered with respect to the GaAs valence band maximum by different choices of the valence band offset.

I. INTRODUCTION

The physics of impurities and defects at interfaces needs considerable development as a result of the increasing scientific and technological importance of semiconductor superlattices and the ability to fabricate superlattices with very thin layers (and hence large interface-to-volume ratios). Impurities at interfaces are known to have different deep impurity levels from those in the bulk[1–4] because of the reduced point group symmetry: C_{1v} [5,6] for a GaAs/Ge (110) interface with its $(1\bar{1}0)$ reflection plane, instead of T_d.[6] In the bulk, an s- and p-bonded impurity would normally produce four deep levels in the vicinity of the fundamental bandgap[7]: one s-like A_1 level and a triply degenerate p-like T_2 level.[6] The interface shifts the A_1 level and splits the T_2 level into two nearly degenerate π-like a'- and a''-symmetric levels[5] and a σ-like a'[5] or a_1 level.[6] (See Figs. 1 and 2.) The splitting of the π-like levels is due to second- and more-distant neighbors.

A number of papers have dealt with the physics of deep impurity levels at interfaces,[8,9] although many more have considered deep levels at surfaces.[10,11] One feature of interfacial deep levels that has not been adequately explored is how the energies of deep impurities depend on valence-band offsets, few of which are known accurately. In this article we present theoretical results bearing on this question, taking as our prototypical defect As on a Ga interfacial site of a GaAs/Ge (110) perfect heterojunction.

II. CALCULATIONS

Our calculations are based on an empirical tight-binding model of electronic structure, and they use a Hamiltonian whose matrix elements reproduce the known (low-temperature) bandgaps of GaAs and Ge.[12] The defect potential matrix V is constructed using the rules of Hjalmarson *et al.*[13] The Hamiltonian for the (110) interface between GaAs and Ge is constructed, and the interfacial Green's function is computed, as discussed in Refs. 2 and 4. Then the interfacial deep levels are computed using the theory of Hjalmarson *et al.*[13] Details of the calculational procedure are available.[14]

In constructing the Hamiltonian for the interface, we must know the valence-band offset, because the electronic structure theory of Vogl *et al.*[12] assumes that the zero of energy for each semiconductor lies at the va-

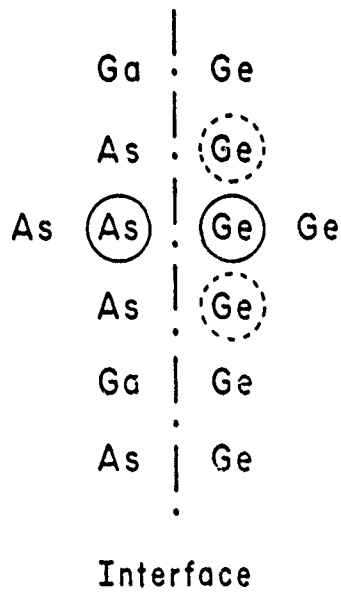


FIG. 1. Schematic illustration of an As_{Ga} defect at a GaAs/Ge (110) interface. The defect's nearest-neighbor environment is the same as for As_{Ga} in bulk GaAs, except for one Ge atom (in solid circle). This one Ge atom will shift the As_{Ga} s-like A_1 deep level and split the As_{Ga} p-like T_2 level into a σ-like a' level polarized along the As_{Ga}–Ge axis, plus a twofold degenerate π-like level polarized perpendicularly to the As_{Ga}–Ge axis. The σ-like level moves up in energy because Ge is more electropositive than the As atom of bulk GaAs that it replaces. The π-like level lies very near the bulk T_2 level and is split into a' (higher energy) and a'' levels by the second-neighbor Ge atoms (in dashed circles). The interface between GaAs and Ge is denoted by a chained line.

[a] Present address, Texas Instruments, Richardson, Texas.
[b] Permanent address.

 0003-6951/88/010164-03$01.75 © 1988 Materials Research Society

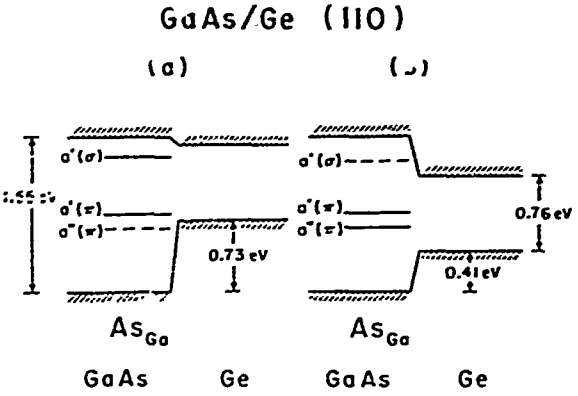


FIG. 2. Predicted energy levels near the fundamental bandgap for the As_{Ga} defect in GaAs at a GaAs/Ge interface. We assumed band offsets of (a) 0.73 eV[15] and (b) 0.41 eV.[16] Dashed levels are resonant with the bands of either GaAs or Ge.

lence-band maximum. For GaAs/Ge, which is thought to be a type I heterojunction, the valence-band maximum of Ge lies at higher energy by ΔE_v than that of GaAs (and the conduction band minimum lies at lower energy). The value of the band-offset ΔE_v is controversial, however. Therefore we repeat our calculations for each of two commonly cited theories of band offsets: (i) the Frensley–Kroemer theory, which predicts $\Delta E_v = 0.73$ eV[15] and (ii) the Harrison theory, which predicts $\Delta E_v = 0.41$ eV.[16] There have been numerous other theories of band offsets, such as the electron-affinity rule,[17] including some very recent ones[18]; but none can be considered quantitatively reliable, since the many-electron and lattice-relaxation effects omitted from all of these theories typically account for theoretical uncertainties of several tenths of an eV. Nevertheless, the values $\Delta E_v = 0.73$ eV and 0.41 eV span a reasonable range of band offsets and will serve the purpose of displaying any significant dependences of deep levels on ΔE_v.

We calculate the deep levels associated with interfacial Ga and As antisite defects in GaAs at a GaAs/Ge (110) interface. For all choices of the band-offset ΔE_v, we find that the deep levels of Ga_{As} lie distant from the bandgap and that the As defect on the interfacial Ga site, As_{Ga}, produces three p-like deep levels in or near the common fundamental bandgap of GaAs/Ge: two lower-energy, nearly degenerate π-like levels that are almost at the same energy as the bulk T_2 level and a higher-energy σ-like a' level. The upper π-like level is also of a' symmetry[5] and has positive parity with respect to the $(1\bar{1}0)$ reflection plane, whereas the lower level is a'' symmetric, with negative parity. (The σ-like a' level lies at higher energy than the π-like levels because Ge is more electropositive than the As atom of bulk GaAs that the Ge replaces when the interface is formed.) Our calculations show that, to a very good approximation, the energies of these three levels are fixed relative to the valence-band maximum of GaAs and are independent of the band offset. (See Fig. 2.)

This result can be understood rather simply: The primary effect of replacing half of a GaAs crystal with Ge, from the point of view of an interfacial As_{Ga}, results from the replacement of the defect's neighboring As atom by a Ge atom. The Ge atom perturbs the bulk As_{Ga} deep level with a perturbation potential whose size can be estimated as roughly the difference in Ge and As s-atomic orbital energies,[12] of order 3 eV. The resulting level shift is considerably smaller (by almost an order of magnitude), however, because only a fraction of the As_{Ga} deep-level wavefunction overlaps the Ge perturbation.[19,20] (Recall the first-order perturbation expression for energy shifts.[21]) This fraction is even smaller for the π-like a' and a'' states than for the σ-like a' state. Therefore, although the perturbation associated with replacing As by Ge is of order ~ 3 eV, the resulting level shifts are about an order of magnitude smaller: 0.1 eV for the π-like states and 0.6 eV for the σ-like state. Furthermore, changes in band offset of order 0.3 eV should be viewed as changes in the perturbation by Ge of order 10% and therefore cause $\sim 10\%$ changes of the already small level shifts: only 0.01 to 0.06 eV—a negligible few hundredths of an eV.

III. CONCLUSIONS

Therefore we find that the deep levels of interfacial defects are different from the deep levels of the same defects in the bulk, and that, when measured relative to the valence-band maximum of the material in which they reside, are (to a good approximation) independent of the energy of the valence band edge of the other material and the band offset of the heterostructure. These results, although obtained for an As_{Ga} defect at a GaAs/Ge (110) interface, are more universally valid and can be expected to hold in general for substitutional interfacial defects at perfect semiconductor interfaces.

ACKNOWLEDGMENT

One of us (J.D.D.) thanks the Department of Chemical Engineering and Materials Science of the University of Minnesota for their hospitality during the period in which this manuscript was prepared. And we thank S. Y. Ren for enlightening discussions.

We are grateful to the U. S. Office of Naval Research (Contract Nos. N00014-84-K-0352 and N00014-82-K-0447) and the Robert A. Welch Foundation for their support.

REFERENCES

[1] H. P. Hjalmarson, R. E. Allen, H. Buttner, and J. D. Dow, J. Vac. Sci. Technol. 17, 993 (1980)

[2]R. E. Allen, J. P. Buisson, and J. D. Dow, Appl. Phys. Lett. 39, 975 (1981).
[3]R. E. Allen, R. P. Beres, and J. D. Dow, J. Vac. Sci. Technol. B1, 401 (1983).
[4]J. D. Dow, R. E. Allen, O. F. Sankey, J. P. Buisson, and H. P. Hjalmarson, J. Vac. Sci. Technol. 18, 502 (1981).
[5]For this notation, see F. A. Cotton, *Chemical Applications of Group Theory* (Wiley-Interscience, New York, 1971), 2nd ed.
[6]For this group theoretical notation, see M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).
[7]J. D. Dow in *Highlights of Condensed-Matter Theory*, edited by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna, Italy, and North-Holland, Amsterdam, 1985), pp. 465–494.
[8]A. Madhukar and S. Das Sarma, J. Vac. Sci. Technol. 17, 1120 (1980).
[9]J. Pollmann, in *Festkörperprobleme*, edited by J. Treusch (Vieweg, Braunstein, 1980), Vol. XX, p. 117.
[10]M. S. Daw and D. L. Smith, Phys. Rev. B 20, 5150 (1979); J. Vac. Sci. Technol. 17, 1028 (1980); Appl. Phys. Lett. 36, 690 (1980); Solid State Commun. 37, 205 (1981); M. S. Daw, D. L. Smith, C. A. Swarts, and T. C. McGill, J. Vac. Sci. Technol. 19, 508 (1981).
[11]R. E. Allen and J. D. Dow, Phys. Rev. B 25, 1423 (1982); R. E. Allen and J. D. Dow, J. Vac. Sci. Technol. 19, 383 (1981); R. E. Allen and J. D. Dow, Applications Surf. Sci. 11/12, 362 (1982); J. D. Dow and R. E. Allen, J. Vac. Sci. Technol. 20, 659 (1982); R. E. Allen, H. P. Hjalmarson, and J. D. Dow, Solid State Commun. 41, 419 (1982); O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1984); M. A. Bowen, R. E. Allen, and J. D. Dow, Phys. Rev. B 30, 4617 (1984); R. E. Allen, T. J. Humphreys, J. D. Dow, and O. F. Sankey, J. Vac. Sci. Technol. B 2, 449 (1984); O. F. Sankey, R. E. Allen, and J. D. Dow, J. Ultramicroscopy 14, 127 (1984); O. F. Sankey, R. E. Allen, and J. D. Dow, J. Vac. Sci. Technol. B 2, 491 (1984); O. F. Sankey, R. E. Allen, S. F. Ren, and J. D. Dow, J. Vac. Sci. Technol. B 3, 1162 (1985); O. F. Sankey, R. E. Allen, and J. D. Dow, *Proceedings of the 17th International Conference on Physics of Semiconductors* edited by D. J. Chadi and W. A. Harrison (Springer-Verlag, New York, 1985), pp. 189; J. D. Dow, O. F. Sankey, and R. E. Allen, Appl. Surf. Sci. 22/23, 937 (1985); J. D. Dow, O. F. Sankey, and R. E. Allen, Mater. Sci. Forum 4, 39 (1985); R. E. Allen, O. F. Sankey, and J. D. Dow, Surf. Sci. 168, 376 (1986).
[12]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[13]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980); See also, W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).
[14]R. P. Beres, Ph.D. thesis, Texas A & M University, 1982 (unpublished).
[15]W. R. Frensley and H. Kroemer, J. Vac. Sci. Technol. 13, 810 (1976).
[16]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).
[17]R. L. Anderson, Hung. Acad. Sci. 2, 55 (1970); *Proceedings of the International Conference on Semiconductor Physics*, Prague, 1960 (Academic, New York, 1961), p. 563.
[18]See, for example, the series of articles in the *Proceedings of the 18th International Conference on the Physics of Semiconductors*, Stockholm, Sweden, 11–15 August, 1986, edited by Olof Engström (World Scientific, Singapore, 1987), pp. 155 ff.
[19]S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).
[20]S. Y. Ren, Sci. Sin. 27, 443 (1984).
[21]See, for example, L. I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1968), 3rd ed., p. 246.

BAND STRUCTURE OF InN

David W. Jenkins, Run-Di Hong, and John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46656, U.S.A.

(Received 15 May 1987)

The energy band-structure of wurtzite-structure semiconductive InN is predicted using empirical nearest-neighbor tight-binding theory. The tight-binding parameters are extrapolated from those of zincblende InP, InAs, and InSb by using empirical rules for the dependences of the parameters on bond length and on row of the Periodic Table. The predicted band gap is direct and agrees well with the data for this potential orange light-emitter. It is suggested that zincblende InN, if it can be grown, also will have a band-gap near 2 eV.

1. Introduction

The recent successful growth [1-3] of high-mobility, polycrystalline InN with a direct fundamental band gap in the orange part of the spectrum (2.05 eV [2]) suggests that this new material might some day be fashioned into large band-gap solid state lasers. It is widely believed that III-V materials are more easily doped both n- and p-type than the large band-gap II-VI semiconductors, and so a 2 eV-gap III-V semiconductor such as InN may find exciting applications as a visible light-emitting diode and laser. To stimulate more research on the optical properties of InN, we present here predictions of its energy band structure.

The predicted band structure is based on a nearest-neighbor empirical tight-binding model of wurtzite semiconductors, with an sp^3 basis at each atomic site [4-6]. This model has successfully described the electronic structures and deep impurity levels of a number of other wurtzite semiconductors with large, indirect fundamental band gaps [6,7]. The basis set consists of one s and three p orbitals centered on each atomic site. Hence the Hamiltonian matrix, for fixed wavevector $\vec{k}$, is 16x16. This matrix is displayed explicitly in Ref. [6].

2. Band structure of wurtzite InN

To determine the band structure of InN, we must first determine the tight-binding parameters, namely the diagonal matrix elements E and the off-diagonal elements V [6], and then diagonalize the Hamiltonian matrix. We obtain the tight-binding parameters for wurtzite InN from those of zincblende InP, InAs, and InSb, because the zincblende parameters are known (and exhibit chemical trends) [8] and the nearest-neighbor environments of the two crystal structures are virtually the same [9].

The on-site matrix elements E for InN were determined as follows: The three matrix elements of InP (Row 3 of the Periodic Table), InAs (Row 4), and InSb (Row 5) defined a parabolic function of r, where r is the Row of the Periodic Table. We took the value of this function for r = 2 to be the on-site matrix element for InN. As expected, the resulting matrix elements for the N-site were only weakly dependent on r, and the In-site matrix elements reflected the same chemical trends as the atomic orbital energies w_z [8,10]. (For a tabulation of those energies, see Ref. [8].)

The off-diagonal matrix elements V were assumed to follow Harrison's rule [11]: They are proportional to the inverse square of the bond length [9], and were obtained by multiplying the

TABLE I. Tight binding parameters (in eV) for InN in the notation of Ref. [6]. The parameters are the same for wurtzite and zincblende structures; for the zincblende structure there are additional parameters associated with the s* orbital [7].

E(s,a)	= -7.0721
E(p,a)	= 0.7531
E(s,c)	= -0.6766
E(p,c)	= 4.0306
V(s,s)	= -5.1158
V(x,x)	= 1.7511
V(x,y)	= 3.8027
V(sa,pc)	= 1.8009
V(pa,sc)	= 5.3898

0749-6036/87/040365+05 $02.00/0

© 1987 Academic Press Limited

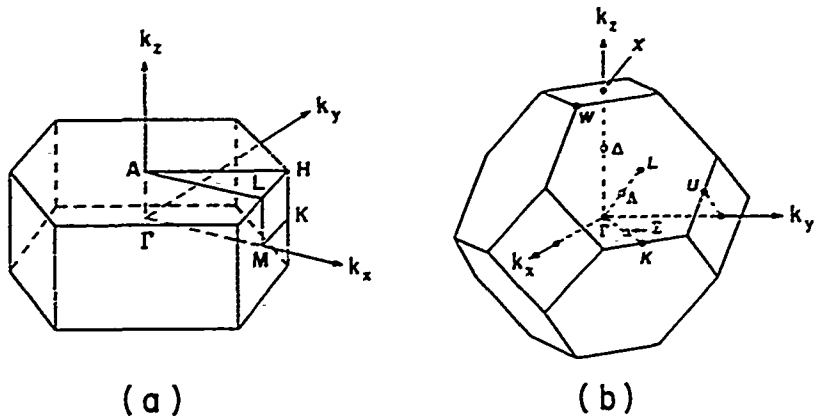


Fig. 1. Brillouin zones of the (a) wurtzite and (b) zincblende crystal structures.

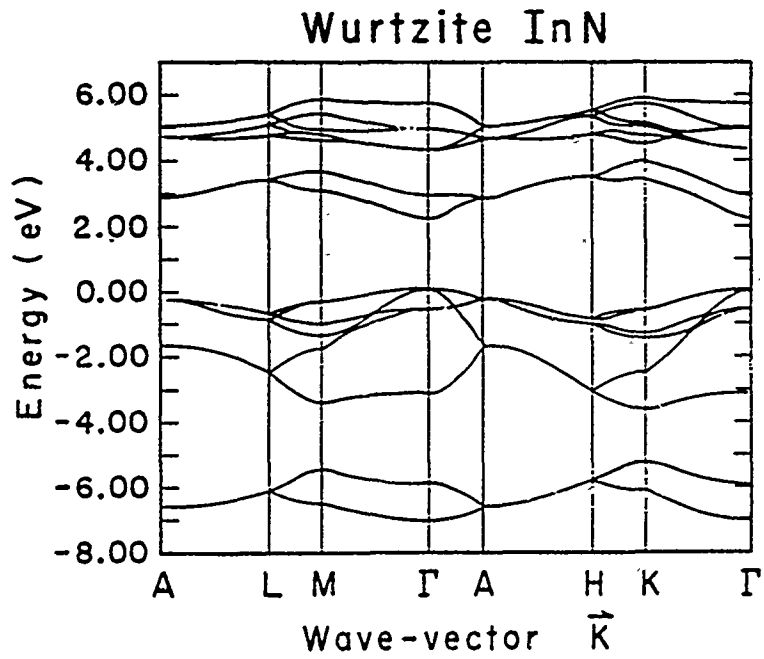


Fig. 2. Predicted band structure of wurtzite InN.

InP values by the square of the ratio of bond lengths for InP and InN.

The resulting parameters are presented in Table I; the Brillouin zone of the wurtzite structure is in Fig. 1a [12], and the predicted band structure of InN is displayed in Fig. 2.

Our band structure is in good agreement with the principal experimental fact [1]: we find a direct fundamental band gap of 2.2 eV (at low temperature), compared with the experimental value of 1.9 eV at room temperature [2]. Considering that the theory obtains a 2 eV gap by scaling parameters for InP, InAs, and InSb, all of which have considerably smaller band gaps (1.41, 0.43, and 0.23 eV, respectively), the agreement with the data is remarkable.

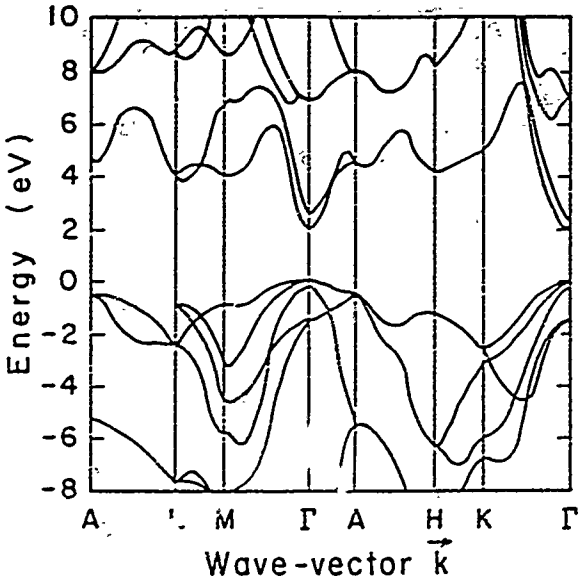


Fig. 3. Band structure of wurtzite InN, after Ref. [1], for comparison with Fig. 2.

The general structure of the valence bands is similar to that obtained both experimentally and theoretically for other wurtzite semiconductors [13], lending further credence to the results.

We are aware of only one calculation of the band structure of InN that predates this work, a calculation based on the empirical pseudopotential method [1,2]. There are significant qualitative and quantitative differences between our bands and the pseudopotential bands (Fig. 3), and we have reason to believe that the tight-binding band structure is more realistic [14]. Based on experience with zincblende [8] and the other wurtzite [6] semiconductors, we expect our valence bands and the lowest conduction bands to mimic the real band structure rather well, up to 3 eV above the valence band maximum.

We propose that experimental studies of wurtzite InN, including angle-dependent photoemission and electron energy-loss spectroscopy, be initiated to resolve the discrepancies between the two theories and to determine the correct band structure of this potential orange light emitter.

3. Zincblende InN

It is interesting to speculate about the properties of zincblende-structure InN, on the grounds that it might be possible to grow metastable zincblende InN in a layered structure, much as zincblende $Zn_{1-x}Mn_xSe$ was grown between ZnSe layers for $0.3 < x < 0.7$ by Kolodziejski et al. [15], despite its preference for the wurtzite phase for $0.3 < x < 0.6$. Because of the small covalent radius of N, few of the common semiconductor materials lattice-match well to InN; for example, SiC, AlN, BP, GaN, and Si have bond-length mismatches with InN of 14%, 13%, 9%, 8%, and -8%, respectively. Clearly the stabilization of the zincblende form of InN in a suitable artificial layered structure represents quite a challenge.

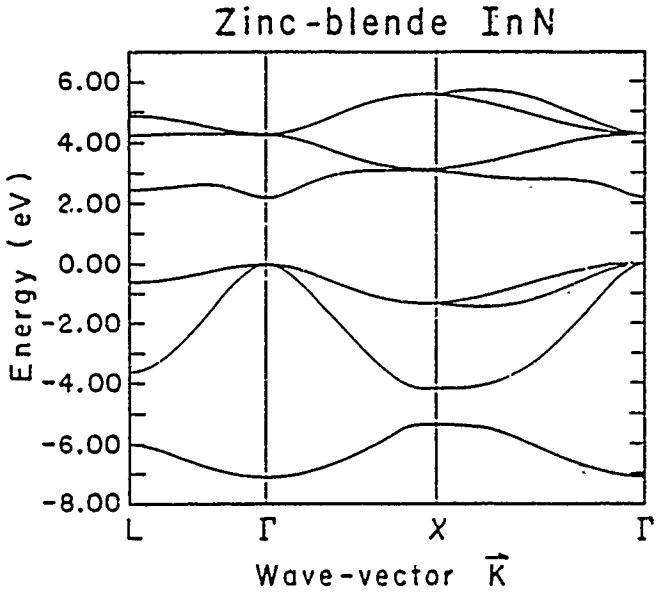


Fig. 4. Predicted band structure of zincblende InN.

In the hope of stimulating efforts to grow zincblende InN we have predicted in Fig. 4 its band structure, using the model of Vogl et al. [3], with the s and p Hamiltonian matrix elements of Table I and s* matrix elements [7] obtained by the same method of extrapolation used for the wurtzite matrix elements. (The zincblende Brillouin zone is displayed in Fig. 1b.) The fundamental band gap is the same in the zincblende structure as in the wurtzite structure for the sp^3 wurtzite nearest-neighbor tight-binding theory and the sp^3s^* Vogl model of the zincblende structure, namely 2.2 eV, or

$$2\,E_{gap} = E(s,a) + E(s,c) - E(p,a) - E(p,c) + [\,(E(s,a)-E(s,c))^2 + 4\,V_{s,s}^{\,2}\,]^{1/2} + [\,(E(p,a)-E(p,c))^2 + 4\,V_{x,x}^{\,2}\,]^{1/2}.$$

Here we have used the notation of Ref. [8].

One unique feature of the predicted band structures is the width of the valence bands, which for both wurtzite and zincblende InN is about half what one finds for other similar semiconductors. These narrow widths may be an artifact of the model, however: for example, a similar tight-binding theory of AlN produced a valence band width larger than subsequently predicted by the local-density pseudo-function method [16]. In this regard, we note that Tsai et al. [17] and Christensen [18] have very recently independently communicated to us local-density band structures which feature broader valence bands. Local-density theory, however, generally produces a fundamental band gap that is much too small, and so the issue of which theory is superior should be resolved experimentally. We do note that we can adjust the parameters of the model to obtain broader valence bands while retaining the 2 eV gap, but such adjusted parameters would not have the *a priori* status of the present parameters. Such adjustments would not greatly affect predictions for either the band gaps of InN-based alloys (such as $In_{1-x}Ga_xN$) or the electronic states near the fundamental band gap associated with localized perturbations, such as deep levels [19]. Therefore, the model should prove useful, even if it is shown to predict valence bands that are somewhat narrow.

4. Summary

In summary, InN in either the wurzite or the zincblende form is expected to be an interesting optoelectronic material with a band gap in or near the orange part of the spectrum. It would be interesting if either of these materials could be grown with electronic-grade quality, and the predictions of this paper tested.

Acknowledgment -- We gratefully acknowledge the generous support of the Office of Naval Research (Contract No. N00014-84-K-0352). We are especially grateful to Dr. K. B. Hathaway, who made several helpful suggestions, including the suggestion that we predict the band structure of zincblende InN. We thank M.-H. Tsai and N. E. Christensen for their comments on the manuscript and for communications concerning their local-density calculations of InN band structures. We also want to thank M. A. Bowen for handling the editorial chores concerning this manuscript.

REFERENCES

[1] T. L. Tansley and C. P Foley, in Proc. 3rd. Internatl. Conf. on Semi-Insulating III-V Materials, Warm Springs, Oregon, 1984, Ed J. Blakemore (Shiva, London) pp. 497-500.

[2] C. P. Foley and T. L. Tansley, Phys. Rev. B 33, 1430 (1986). The room temperature band gap is about 1.89 eV, but appears larger due to a Burstein shift. Here we compute the low-temperature gap, which should be slightly larger than the room-temperature gap.

[3] See also C. P Foley and T. L. Tansley, Appl. Surf. Sci. 22, 663 (1985); K. Osamura, S. Naka, and Y. Murakami, J. Appl. Phys. 46, 3432 (1975); V. V. Sobolev, S. G. Kroitoru, A. F. Andreeva, V. Ya. Malakhov, Fiz. Tekh. Poluprovodn. 13, 833 (1979) [Engl. transl.: Sov. Phys. Semicond. 13, 485 (1979)]; V. A. Tyagai, A. M. Evstigneev, A. N. Krasiko, A. F. Andreeva, and V. Ya. Malakhov, Fiz. Tekh. Poluprovodn. 11, 2142 (1977) [Engl. transl.: Sov. Phys. Semicond. 11, 1257 (1977)]; K. Osamura, K. Nakajima, Y. Murakami, T. H. Shingu, and A. Ohtsuki, Solid State Commun. 11, 617 (1972); and B. R. Natarajan, A. H. Eltoukhy, and J. E. Greene, Thin Solid Films 69, 201 (1980).

[4] R. W. Wyckoff, Crystal Structures, Vol. I supplement V. (Interscience Publications, Inc. New York, 1960). InN has a wurtzite structure with lattice constants a = 3.533 Å and c = 5.693 Å.

[5] J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).

[6] A. Kobayashi, O. F. Sankey, S. M. Volz, and J. D. Dow, Phys. Rev. B 28, 935 (1983).

[7] For the zincblende structure, there is an extra s* basis orbital not included in the wurtzite basis. This orbital has the function of coupling to the lower-lying p-states and serves to depress the indirect conduction band minima, a feature that is important for small band-gap indirect semiconductors, such as Si, but not for large-gap InN. Here we have used the following parameters associated with the s* state: E(s*,a)=9.0138 eV, E(s*,c)=6.8154 eV, V(s*a,pc)=2.4807eV, and V(s*c,pa)=3.2109eV, in the notation of Ref. [8].

[8] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids, 44, 353 (1983).

[9] In a wurtzite structure all nearest-neighbor distances are equal if (c/a)^2 = 8/3. This ratio in InN differs from the equal bond-length limit of 8/3 by only 1.2%, and

so we assume that the bond lengths are equal.

[10] Indeed, after using this method for extrapolating the tight-binding parameters first, we tried other extrapolation schemes based on atomic energies and nuclear charges, and found band structures that either were comparable with the ones presented here or featured fundamental band gaps in disagreement with the data.

[11] W. A. Harrison, Electronic Structure and the Properties of Solids, (W. H. Freeman and Co., San Francisco, 1980), p. 47.

[12] See, for example, C. Kittel, Introduction to Solid State Physics, fifth edition (J. Wiley and Sons, New York, 1976).

[13] See, for example, Ref. [6] and references therein.

[14] The number of bands predicted in Ref. [1], the general structure of the valence bands, the absence of a relative minimum of the conduction band within 2 eV of the conduction band edge, and the lack of agreement of the displayed bands of Fig. 3 at the two Γ points ($\vec{k}=\vec{0}$) and the two A points ($\vec{k} = (2\pi/c)(0,0,1/2)$), lead us to prefer the empirical tight-binding results. We attempted to fit the band structure of Refs. [1] and [2] with an empirical tight-binding theory, and were unable to obtain a satisfactory fit, as judged by our experience fitting calculated band structures for other wurtzite semiconductors. The band structure of Ref. [2] appears to have agreement at the two Γ and the two A points, but seems to be missing a band from A to L.

[15] L. A. Kolodziejski, R. L. Gunshor, N. Otsuka, B. P. Gu, Y. Hefetz, and A. V. Nurmikko, Appl. Phys. Lett. 47, 1482 (1986).

[16] R. V. Kasowski and F. S. Ohuchi, "Electronic structure of Cu overlayers on AlN," Phys. Rev. B; to be published.

[17] M.-H. Tsai, D. W. Jenkins, J. D. Dow, and R. V. Kasowski, "Pseudo-function theory of the electronic structure of InN," to be published.

[18] N. E. Christensen, unpublished.

[19] D. W. Jenkins and J. D. Dow, unpublished.

International Journal of Modern Physics B Vol. 2, No. 6 (1988) 1483–1493
© World Scientific Publishing Company

X-RAY SPECTRA OF ORDERED AND DISORDERED BINARY COMPOUNDS

MARSHALL A. BOWEN
Department of Computer Science, Western Illinois University, Macomb, IL 61455, USA

and

JOHN D. DOW
Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA

Received 24 June 1988
PACS Nos.: 78.70, Dm; 71.50+t

The X-ray photoemission, absorption, and emission spectra of a one-dimensional, one-orbital-per-site model of an ordered binary compound AB and its disordered counterparts $A_{1-x}B_{1-x}B_xA_x$ are calculated. With increasing x, the band gap of this model changes from being large (insulating) to very small (semiconductive). Final-state interaction effects that produce electronic excitation of valence electrons across the gap are incorporated in a change-of-mean-field approximation. The spectra exhibit sidebands associated with disorder and distinctive features associated with the character of the one-electron states and excitonic states.

1. Introduction

In this paper we report calculations of the X-ray photoemission, absorption, and emission spectra of a one-dimensional model of an ordered diatomic compound AB, and of its disordered counterparts $A_{1-x}B_{1-x}B_xA_x$ with various concentrations x of antisite defects. With increasing x this model changes from having a large band gap (an insulator) to having a sufficiently small gap and density of states in the vicinity of the Fermi surface that it is semiconductive. The calculations are based on a one-electron and one-orbital per site tight-binding model, and incorporate many-electron final-state interaction effects in a change-of-mean-field approximation.[1] To our knowledge, this is the first treatment of a model for such multi-electron processes in diatomic compounds, either ordered or disordered.

2. Model

In the change of mean-field model, we treat only the valence and conduction electrons, which in the initial state for X-ray photoemission are described by the N-electron Hamiltonian

$$H_I = \sum_{i=1}^{N} h_i , \tag{1}$$

where we have the tight-binding one-electron Hamiltonian

$$h_i = \sum_{n=1}^{M} \varepsilon(n) |n)(n| + \beta |n)(n+1| + \beta |n+1)(n| . \tag{2}$$

Here $|n)$ refers to the orbital at the n-th site of M sites, $\varepsilon(n)$ is either ε_A or ε_B, and β is the transfer matrix element. For the disordered solid, $\varepsilon(n)$ is ε_A with probability $1-x$ and ε_B with probability x, if n is even; $\varepsilon(n)$ is ε_B with probability $1-x$ and ε_A with probability x, if n is odd. For the ordered solid, we have $x = 0$.

The final-state Hamiltonian is

$$H_F = \sum_{i=1}^{N} h_i' , \tag{3}$$

where h_i' is identical to h_i, except at the site R of the core hole -- which has an additional electron-hole interaction term $V = V_0 |R)(R|$, with $V_0 < 0$. The initial many-electron state $|I\rangle$ is a Slater determinant of the N lowest-energy single-particle orbital eigenfunctions $|\phi)$ of h. The various final-states $|Fv\rangle$ are determinants of N eigenfunctions $|\psi)$ of h'.

Here we consider only a single-spin channel (up), because the effects of the spin-down channel can be included *a posteriori* by convolution[1] and merely serve to asymmetrically broaden the single-channel spectra slightly, without altering the essential physics. Hence, in this paper, the number of electrons N is half the number of sites M; the Fermi energy is halfway between ε_A and ε_B, which we define to be zero: $\varepsilon_A = -\varepsilon_B$.

In the change-of-mean-field model the X-ray photoemission spectrum (XPS) for photoelectrons of energy E is

$$I(E) = \sum_{v} |\langle I|Fv\rangle|^2 \, \delta(E + E_{Fv} - E_I - \hbar\omega - \varepsilon_{core}) . \tag{4}$$

Here E_{Fv} and E_I are the final- and initial-state energies of the conduction electron gas and ε_{core} is the core hole energy relative to the center of the conduction band. The XPS recoil energy is

$$E_{Fv} - E_I = \sum_{i}^{N} \varepsilon_i' - \sum_{i}^{N} \varepsilon_i . \tag{5}$$

where the sums are over occupied final- and initial-state single-particle energies.

Likewise the X-ray emission spectrum for photons of energy E is

$$\kappa(E) = \Sigma_v |\langle Fv|M|I\rangle|^2 \, \delta(E - E_I + E_{Fv}) . \tag{6}$$

where the initial state $|I\rangle$ is a determinant of $N+1$ $|\psi)$'s, and the various final states $|Fv\rangle$ are configurations described by determinants of $N|\phi)$'s and one core orbital:

$$\langle F|vM|I\rangle = M_0 \begin{vmatrix} \psi_1(\mathbf{R}) & \psi_2(\mathbf{R}) & \dots & \psi_{N+1}(\mathbf{R}) \\ (\phi_{v,1}|\psi_1) & & & \\ (\phi_{v,2}|\psi_1) & & & \\ \dots & & & \\ (\phi_{v,N}|\psi_1) & \dots & & |(\phi_{v,N}|\psi_{N+1}) \end{vmatrix} . \tag{7}$$

Here $(\phi|\psi)$ is a scalar product, $\mathbf{R}$ is the core-hole site, and we have assumed that the core radius is negligible[2] and M_0 is a constant. A similar expression holds for X-ray aborption.[2,3]

We have the sum rules for XPS[1]

$$\int_{-\infty}^{\infty} I(E)\,dE = 1\,, \tag{8}$$

and for emission[2]

$$\int_{-\infty}^{\infty} \kappa(E)\,dE = M_0^2 \Sigma_i |\psi_i(\mathbf{R})|^2\,, \tag{9}$$

where the sum is over occupied orbitals of the initial state.

The lineshapes we display are ensemble averages (denoted by $<<\dots>>$) over all core hole sites (typically 1000 such sites); e.g.,

$$\chi(E) \equiv << \Sigma_v |\langle F|M|I\rangle|^2 M_0^{-2}\, \delta(E - E_I + E_{Fv}) >>\,. \tag{10}$$

The procedure for executing the calculations has been described elsewhere[3,4]. Briefly, we have directly diagonalized the Hamiltonian for $N = 20$ electrons and $M = 40$ sites, with the core-hole site being one of the ten innermost sites.

3. Results

The results of our calculations (for $\beta = -1/2$, $V_0 = -2|\beta|$, $\varepsilon_B = 2|\beta|$, and $\varepsilon_A = -\varepsilon_B$) are given in Figs. 1 and 2, 3 and 4, and 5 and 6, for X-ray photoemission, X-ray absorption, and X-ray emission, respectively.

These spectra can be understood by referring to the density of states[5,6] (Fig. 7), which feature a filled A-like valence band and an empty B-like conduction band (for $x = 0$) which both develop sidebands associated with antisite defects (e.g., ABAAAB) and clusters of defects as a result of disorder (for $x \neq 0$). For sufficient disorder ($x > 0.1$) the gap between the valence and the conduction bands becomes

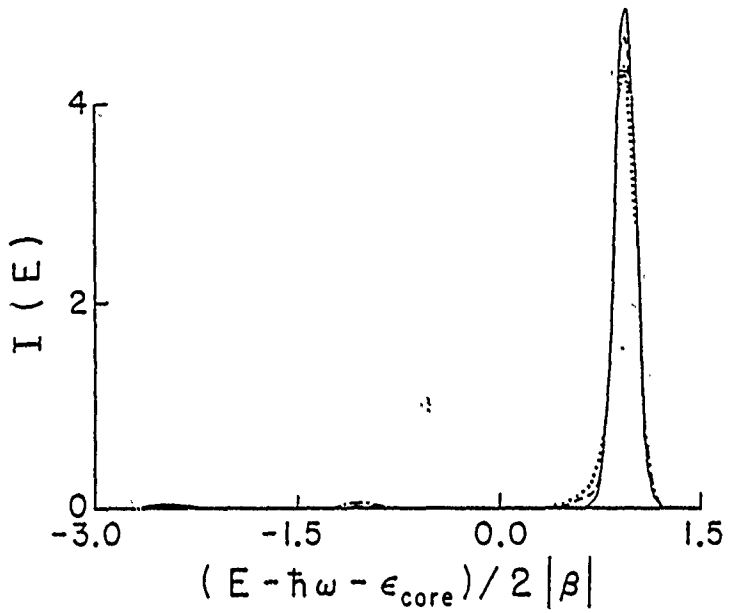


Fig. 1. Predicted X-ray photoemission spectra for core excitation of the A-site as a function of the emitted electron's energy E for $A_{1-x}B_{1-x}B_xA_x$, for $x = 0$ (solid line), for $x = 0.1$ (dashed line), $x = 0.2$ (dotted), $x = 0.3$ (dashed-dotted), $x = 0.4$ (dash-double-dotted), and $x = 0.5$ (double-dash dotted).

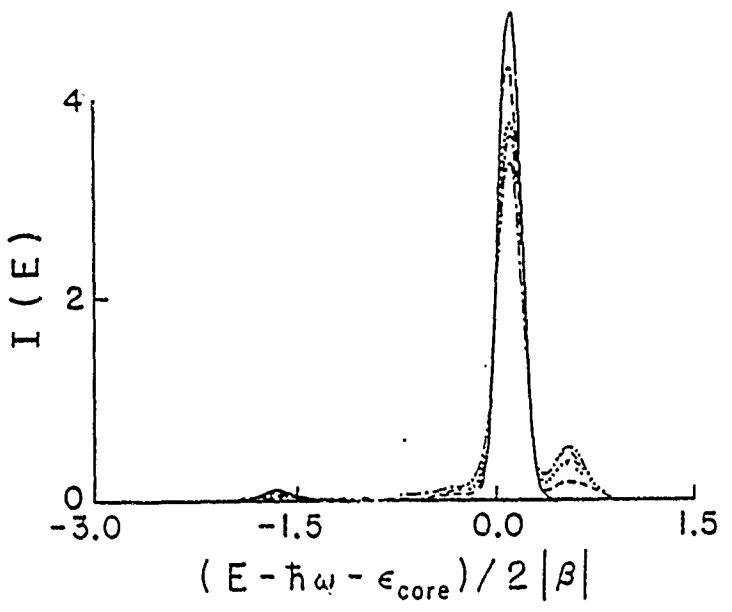


Fig. 2. Predicted X-ray photoemission spectra for excitation of the B-site, as in Fig. 1.

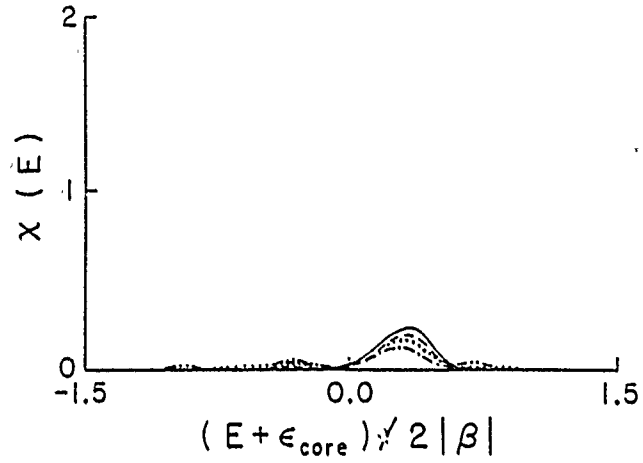


Fig. 3. Predicted X-ray absorption spectra $\chi(E)$ for excitation of a core level at the A-site, as in Fig. 1.

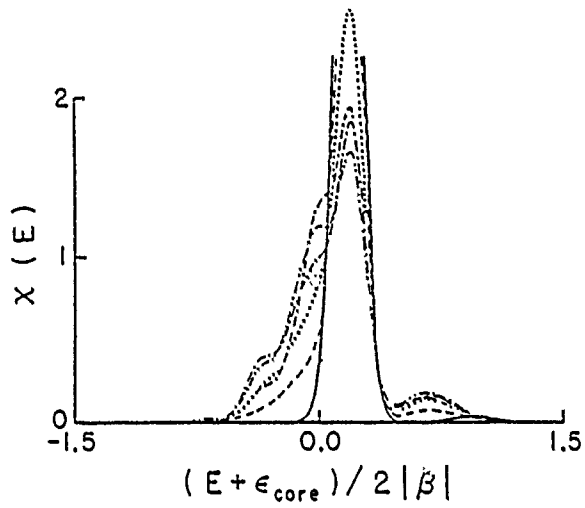


Fig. 4. Predicted X-ray absorption spectra $\chi(E)$ for excitation of a core level at the B-site, as in Fig. 1.

very small. That is, the disorder converts the model material from insulating to semiconductive.

Further aids to understanding the spectra are Figs. 8 and 9, which display for the ordered diatomic material AB the total density of states, the local densities of states at the A and B sites, and the hole-perturbed densities of states at the A and B sites. Note that the hole-perturbed local density of states at the B-site has two bound states, one associated with each band, but that at the A-site there is only an A-band bound state.

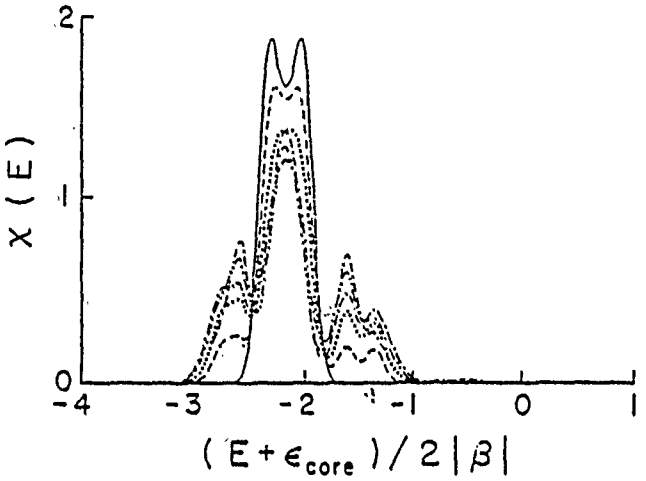


Fig. 5. Predicted A-site X-ray emission spectra $\chi(E)$, as in Fig. 1.

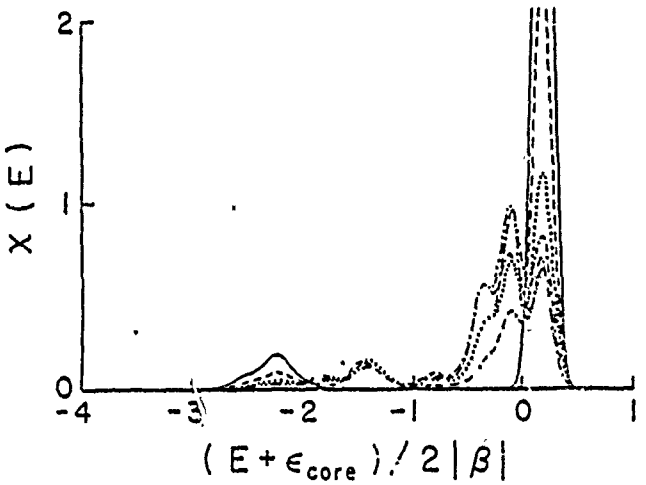


Fig. 6. Predicted B-site X-ray emission spectra $\chi(E)$, as in Fig. 1.

a) XPS spectra

The X-ray photoemission spectra (XPS) for creation of a core hole at site A (Fig. 1) each have a large peak [at $(E-\hbar\omega-\varepsilon_{core})/2|\beta| \simeq 1$] associated with photoemission from the core. For $x = 0$, this peak is symmetric. There is a low-energy sideband [at $(E-\hbar\omega-\varepsilon_{core})/2|\beta| \simeq -2.5$] corresponding to accompanying excitations of the electron gas from the bound state of the (filled) A band to the empty B band (see Fig. 9). For $x \neq 0$, the structure near $(E-\hbar\omega-\varepsilon_{core}) \simeq -2|\beta|$ corresponds to disorder-activated excitations from near the top of the

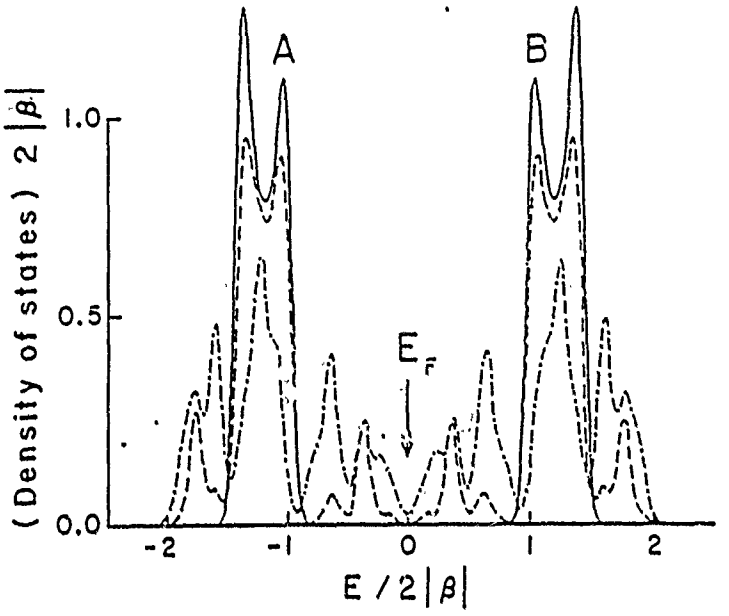


Fig. 7. Broadened densities[5] of states (times $2|\beta|$, the unit of energy) versus $E/2|\beta|$ for $A_{1-x}B_{1-x}B_xA_x$ with $x = 0.1$ (solid line), $x = 0.3$ (dashed), and $x = 0.5$ (chained). Levels with $E < 0$ lie below the Fermi energy $E_F = 0$ and are occupied at zero temperature. For $x = 0.1$, the characteristic $|E - E_0|^{-1/2}$ van Hove singularities are blurred slightly. For larger x, antisite-defect peaks are prominent, small sidebands associated with clusters of antisite defects are visible, and band-center peaks associated with disorder are present.[8]

filled A band to the empty B band (see Fig. 7). With increasing disorder the gap between the filled A band and the empty B band fills in, and the main peak becomes more asymmetric, due to low energy excitations of the electron gas across the gap from below the Fermi energy to above. (These excitations are forbidden for $x = 0$, because the model is insulating in this limit — there are no states in the band gap of $4|\beta|$.)

XPS at the B site (Fig. 2) produces a nearly recoilless peak at $E - \hbar\omega - \varepsilon_{core} \simeq 0$, and a peak at $\simeq -1.7$ corresponding to transitions from the filled A band to the empty bound state (see Fig. 8) of the B band. The sidebands of the main peak associated with disorder are due to transitions from filled to empty states. In particular the small peak at $(E - \hbar\omega - \varepsilon_{core})/2|\beta| \simeq 0.6$ is associated with shape-resonance excitations of the lowest density-of-states peak above the Fermi energy from below the Fermi energy (where it is pulled by the core hole) to above.[9]

b) Absorption spectra

At the A-site, the absorption is relatively weak and featureless because the core-hole state has A character and the A character of the conduction band is weak (Fig. 3).

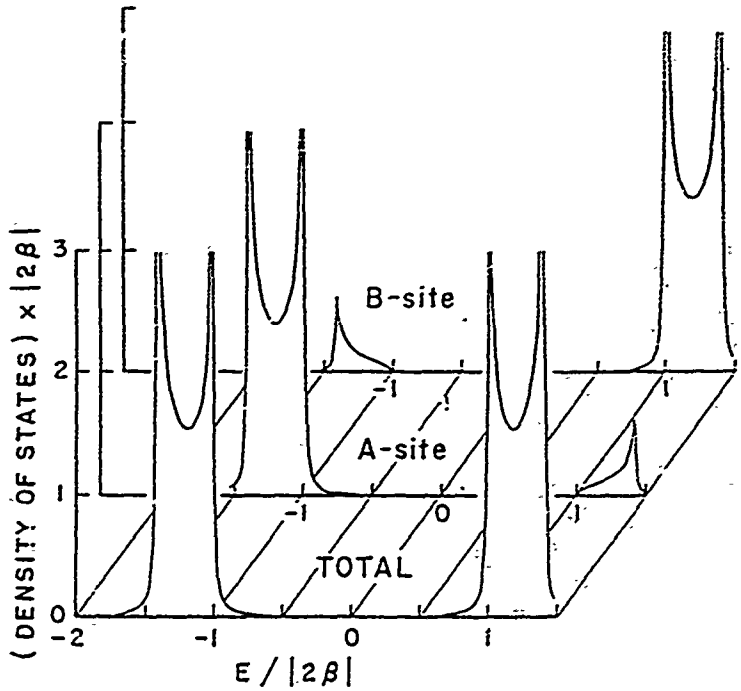


Fig. 8. Various calculated densities of states (DOS) (times $2|\beta|$) of a perfect AB crystal: total DOS; local DOS at an A-site; and local DOS at a B-site. Note that the upper part of the valence band is totally A-like and the lower part of the conduction band is B-like.

At the B-site (Fig. 4), there is a strong absorption peak for $x = 0$ corresponding to electronic transitions from the B-like core-hole level to the bound exciton state of the empty B band. This excitonic level has a strong B-like character. In addition, a higher energy peak [near $(E+\varepsilon_{core})/2|\beta| \simeq 1.0$] corresponds to transitions into the B continuum. Disorder ($x \neq 0$) merely serves to add sidebands and shape resonances[9] to the spectrum.

c) Emission spectra

For a core-hole at the A-site, the emission spectra reflect the A-like density of states of the valence band, and exhibit disorder-activated sidebands. The initial states each has one electron at the B-band minimum which prefers to project onto a final-state B-like one-electron orbital, causing the spectra to exhibit the A-like character of the remaining electrons.

However, for a core-hole at the B-site, the spectra for $x = 0$ reveal two interesting features: (i) a main peak near $E+\varepsilon_{core} \simeq 0$ corresponding to the electron in the excitonic bound state of the B band falling into the core hole and (ii) a weaker peak at $(E+\varepsilon_{core})/2|\beta| \simeq -2$ associated with transitions in which the core hole captures an electron from the bound state below the A-band; its

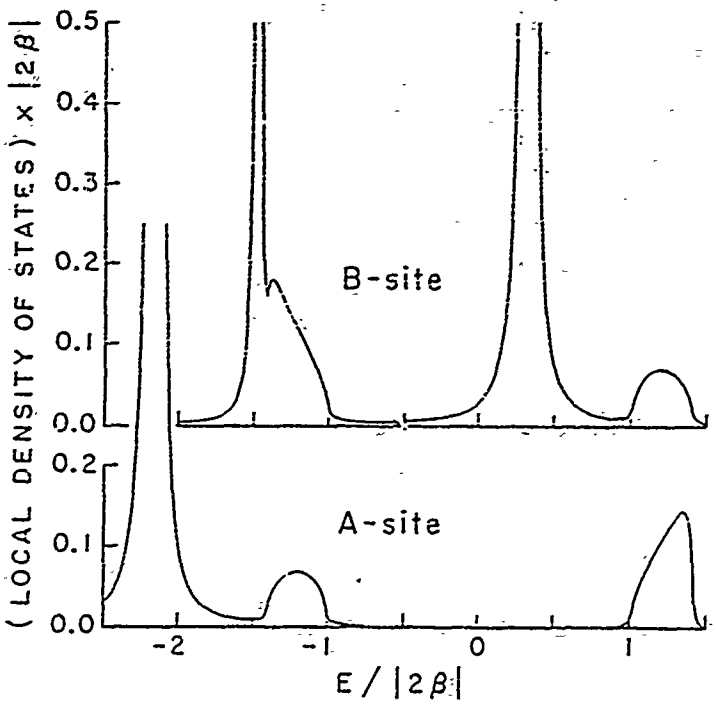


Fig. 9. The core-hole perturbed densities of states (times $2|\beta|$) at the core-hole site in a perfect AB lattice, for a core-hole at an A-site and a B-site. Note that the core-hole produces two bound states when it occupies a B-site, but only one when on an A-site.

small size reflects the reduced B-like character in that state. This peak disappears with disorder, which weakens the bound state below the A-band. Disorder activates additional sideband features.

d) Discussion

The X-ray spectra discussed here exhibit the effects of final-state electron-hole interactions as well as the effects of disorder. For an insulating model, the spectra have excitonic features. However, no prominent "X-ray edge anomaly"[10] manifests itself even in the semiconducting regime.

One interesting feature of the XPS spectra is that the major lines are nearly symmetric — a feature once observed[11] for sodium-tungsten bronzes with low local state densities at the core hole site. At the time of that observation, the only available theory was for free-electron metals,[12] which were predicted to have considerably asymmetric XPS lines for the case of strong electron-hole interactions — contrary to the observations.[13] The theory presented here suggests that nearly symmetric XPS lines can be expected in systems with low densities of states at the Fermi energy.

In summary, the X-ray spectra of $A_{1-x}B_{1-x}B_xA_x$ are rich in structure associated with disorder, but primarily determined by the densities of states.

Appendix A

Matrix elements of the retarded Green's function for the perfect AB lattice between two A states in cells i and $i + n$ are given by

$$G_{AA}(n,E) = \langle A, i|G(E)|A, i+n\rangle$$

$$= \pm (E-\varepsilon_b)(\alpha^2-4\beta^4)^{-1/2}\{[\alpha/2\beta^2] \mp \{[\alpha/2\beta^2]^2-1\}^{1/2}\}^{|n|},$$

where we have $\alpha = (E-\varepsilon_a)(E-\varepsilon_b) - 2\beta^2$ and the lower set of signs is applicable from the bottom of the upper band (ε_b) to the bottom of the lower band $[\varepsilon_a + \varepsilon_b - \{(\varepsilon_b-\varepsilon_a)^2 + 16\beta^2\}^{1/2}]/2$.

Other matrix elements are given by

$$\langle B, i|G(E)|B, i+n\rangle = [(E-\varepsilon_a)/(E-\varepsilon_b)]\, G_{AA}(n,E),$$

$$\langle A, i|G(E)|B, i+n\rangle = [\beta/(E-\varepsilon_b)]\,[G_{AA}(n,E) + G_{AA}(n+1,E)],$$

and

$$\langle B, i|G(E)|A, i+n\rangle = [\beta/(E-\varepsilon_b)]\,[G_{AA}(n,E) + G_{AA}(n-1,E)].$$

Using these matrix elements one may examine the local density of states of various defects by using Dyson's equation $G = G_0(1 - VG_0)^{-1}$. In particular new states outside the bands may be found by looking for the zeros of the determinant of $[1 - VG_0]$. We have carried out this search for the parameters given and find that filling two A sites with B-like states gives:

Cell spacing	*Energies*
1	1.88, 1.56, 0.55, 0.15
2	1.78, 1.74, 0.39, 0.32
3	1.76, 1.76, 0.36, 0.36

The results of filling two B sites with A-like states is just the negative of the above energies. Interchanging an A and a B state gives:

Cell spacing	*Energies*
1	±1.61, ±0.64
2	±1.74, ±0.29
3	±1.76, ±0.36

These results are well correlated with the dominant features and concentration dependence of the numerically calculated densities of states.

Acknowledgment

We are grateful to the U. S. Office of Naval Research for their generous support (Contract No. N00014-84-K-0352).

References

1. J. D. Dow and C. P. Flynn, *J. Phys.* C13 (1980) 1341.
2. C. A. Swarts, J. D. Dow, and C. P. Flynn, *Phys. Rev. Lett.* 43 (1979) 158; C. A. Swarts and J. D. Dow, to be published.
3. M. A. Bowen and J. D. Dow, *Phys. Rev.* B32 (1985) 5119.
4. S. Satpathy, J. D. Dow, and M. A. Bowen, *Phys. Rev.* B28 (1983) 4255.
5. The densities of states (per site) are calculated using the negative eigenvalue method for 10000 sites,[6] and Gaussian broadened[7] with $\Gamma = 0.03\,(2|\beta|)$.
6. P. Dean, *Rev. Mod. Phys.* 44 (1972) 127; *Proc. Roy. Soc. London*, Ser. A 260 (1961) 263; *Proc. Phys. Soc., London* 84 (1964) 727; D. N. Payton and W. M. Visscher, *Phys. Rev.* 154 (1967) 802; 156 (1967) 1032; 175 (1968) 1201.
7. The broadening function for the spectra is $(2\pi\Gamma^2)^{-1/2}\exp(-x^2/2\Gamma^2)$, with $\Gamma = 0.08|2\beta|$. For the densities of states in Figs. 8 and 9, we use $\Gamma = 0.03|2\beta|$.
8. Any slight asymmetries of these spectra about $E=0$ are due to the Monte Carlo method employed in calculating them.
9. E. Mehreteab and J. D. Dow, *Solid State Commun.* 43 (1982) 837; *Phys. Rev.* B26 (1982) 2261.
10. J. Friedel, *Comments on Solid State Physics* 2 (1969) 21; P. Nozieres and C. T. de Dominicis, *Phys. Rev.* 178 (1969) 1097.
11. M. Campagna, G. K. Wertheim, H. R. Shanks, F. Zumsteg, and E. Banks, *Phys. Rev. Lett.* 34 (1975) 738.
12. S. Doniach and M. Sunjic, *J. Phys.* C3 (1970) 285.
13. J. D. Dow, *Phil. Mag.* 35 (1977) 837.

DEFECTS IN SUPERLATTICES UNDER PRESSURE

RUN-DI HONG, DAVID W. JENKINS, S. Y. REN, AND JOHN D. DOW
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 U.S.A.

ABSTRACT

We report theoretical calculations of deep levels in GaAs/$Al_xGa_{1-x}As$ superlattices under hydrostatic pressure. We predict phase diagrams for DX centers: for a given composition x there is a function p(a), which relates pressure p and GaAs quantum-well width a, and defines a phase boundary between two regions: one in which DX is a deep trap in the fundamental band gap and another in which the DX deep level lies in the conduction band.

I. INTRODUCTION

For a superlattice in which the crystal structure is continuous across the interfaces between alternate layers of different semiconductors, there is no difficulty, in principle, in calculating the energy bands and the deep impurity levels [1]. In this paper, we present the results of such calculations for superlattices under hydrostatic pressure, based on an empirical tight-binding model. Deep impurity levels are treated by a Green's function method [2], using special points to perform k-space sums [3]. To account for the effects of pressure, we fit the pressure dependences of empirical matrix elements of the Hamiltonian to the observed pressure variations of the bulk band structures [4]. The valence band offset at the interface is treated by using the experimental results of Wolford et al. [5].

In the present study we consider a DX center in an $Al_yGa_{1-y}As$ layer of an $Al_yGa_{1-y}As/Al_xGa_{1-x}As$ superlattice (for $y < x$), and we show that (i) reducing the thickness of the $Al_yGa_{1-y}As$ layer, (ii) increasing the alloy composition y, and (iii) the application of hydrostatic pressure all have similar effects on the "DX" center [6]: the deep level is driven from the conduction band of $Al_yGa_{1-y}As$ into the fundamental band gap. Thus this work adopts the pioneering viewpoint of Hjalmarson et al. [7], that the DX center is associated with a substitutional donor (such as Si_{Ga}) and produces a deep level resonant with the conduction band in GaAs, but the level can be driven into the fundamental band gap by either (i) increasing the alloy composition [6], increasing pressure [8], or reducing the thickness of the GaAs layer containing the donor [1]. The persistent photoconductivity of the DX center is presumably explained by the weak electron-phonon coupling model of Hjalmarson and Drummond [9,10]. We shall refer to this substitutional donor as a DX center hereafter, although many workers believe the DX center is a donor-vacancy pair [6].

This paper is organized as follows: Section II discusses the theory, while Section III describes our results for the DX center when subjected to a change of alloy composition x in $Al_xGa_{1-x}As$, an increase of pressure, or a decrease of the quantum-well thickness. The phenomenon of a shallow-deep

Mat. Res. Soc. Symp. Proc. Vol. 77. 1987 Materials Research Society

transition is discussed, and the notion of a critical combination of pressure and quantum-well thickness is introduced: for pressures larger than this critical pressure or for quantum-wells thinner than the critical width, the DX center will be a deep trap in the fundamental band gap.

II. THEORY

We begin with a minimal basis-set LCAO-type model (with one s-, one excited s*-, and three p-orbitals centered on each atom) and include only nearest-neighbor interaction matrix elements [11]. This is sufficient to give a reasonable description of the lowest conduction band as well as the valence bands for both homogeneous semiconductors GaAs and AlAs and alloys $Al_xGa_{1-x}As$. Hence the same type of Hamiltonian should describe the $GaAs/Al_xGa_{1-x}As$ superlattice [1]. (Alloys are treated in a virtual crystal approximation [12].) A universal set of parameters for the Hamiltonian is available [11]; these have accurately reproduced the band structures of GaAs and AlAs. We perform our calculations for $GaAs/Al_xGa_{1-x}As$ superlattices whose layers are perpendicular to the (001) direction. We assume GaAs and $Al_xGa_{1-x}As$ are perfectly lattice-matched and we consider N_1 layers of GaAs and N_2 layers of $Al_xGa_{1-x}As$ repeated periodically. The dimension of the superlattice Hamiltonian matrix (for a given surface wave-vector K in the plane of the layers) equals the number of orbitals per atom times $2(N_1+N_2)$. The valence band edge discontinuity, which has some effects on the positions of deep levels in the superlattice, is chosen to reproduce the recent experimental results by Wolford et al. (1/3 of the band offset is in the valence band) [5]. This offset is incorporated in the model by adding a constant to all of the diagonal matrix elements of the Hamiltonian of GaAs, because these tight-binding parameters [11] are defined with respect to the top of the valence band. The full Hamiltonian, which produces the superlattice band structure as well as all s- and p-bonded deep impurity levels, is given in detail in the paper by Ren et al. [1]. We follow the techniques of that work and the theory of deep levels in pressurized bulk semiconductors [4] in performing our own calculations for the presssure dependences of the deep levels in superlattices.

The impurity levels are evaluated following the Green's function theory of Hjalmarson et al. [2], which solves the secular equation for the deep level energy E:

$$\det(1-GV) = 0.$$

Here V is the defect potential matrix, which is diagonal in the Vogl sp^3s^* basis (in the case of zero lattice relaxation) and has matrix elements related to the defect and host atomic energies [2]. The Green's function operator is $G = (E-H)^{-1}$, where H is the host tight-binding Hamiltonian operator and E has a small positive imaginary part when E lies outside the band gap. We consider substitutional donor impurities in this paper, and they generally have C_{2v} point group symmetry in the $GaAs/Al_xGa_{1-x}As$ superlattice. Details of the calculational procedure may be found in Refs. [1] and [4].

The pressure dependences dE/dp of substitutional deep point defect levels in $GaAs/Al_xGa_{1-x}As$ are deduced using the fact that the hydrostatic pressure preserves crystal symmetries while altering bond lengths. In the

model of Vogl et al., the diagonal matrix elements of the host Hamiltonian H and the on-site defect potential V are independent of changes in the bond length, while the off-diagonal matrix elements of H depend on bond length d according to Harrison's rule:

$$H_{\alpha,\beta} = H_{\alpha,\beta}^{0} \, (d_0/d)^{n(\alpha,\beta)} \quad ,$$

where d_0 is the zero-pressure bond length. The exponents $n(\alpha,\beta)$ (for α,β = s, p, and s*) are obtained by fitting the observed pressure dependences of the direct band gaps at Γ, X, and L, and the indirect gaps from the valence band maximum to L and X, using least-square methods. A discussion of the effects of hydrostatic pressure on bulk electronic structures of III-V semiconductors is given in Ref. [4]. We use the exponents $n(\alpha,\beta)$ obtained for the bulk semiconductors as input to calculate the pressure dependences of the electronic structures of superlattice GaAs/$Al_xGa_{1-x}As$ and its deep levels.

III. RESULTS AND DISCUSSION

In this section we assess (i) the effect of increased alloy composition x on a deep donor level in bulk $Al_xGa_{1-x}As$, (ii) the effect of hydrostatic pressure on the DX center, and (iii) the effect of decreased quantum-well thickness on a DX center or substitutional point defect in a GaAs well of a GaAs/$Al_xGa_{1-x}As$ superlattice. We take as our model of a DX center, a single Si impurity on a cation site (while acknowledging that DX behavior can be associated with a wider class of donors, including donor defect complexes).

We assume that the defect potential of our DX center produces a deep level with the property that the level descends into the fundamental band gap at an alloy composition near x=0.3 [6] in $Al_xGa_{1-x}As$. The defect potential V that produces such a level within the context of the deep level theory of Hjalmarson et al. is diagonal with V_s = -1.047 eV and V_p = -0.844 eV, essentially equal to the Si defect potential (within the theoretical uncertainty). The resulting alloy dependence of that deep level in $Al_xGa_{1-x}As$ is displayed in Fig. 1, and was first predicted by Hjalmarson [2,7], who also first provided the picture of Si as a potential DX center -- an idea that has since been developed by Yamaguchi [13], who has also compared with data the predictions of the Hjalmarson theory for the dependence of the DX center on alloy composition x and pressure p. Note that this center undergoes a shallow-deep transition. The s-like A_1-symmetric deep level in GaAs lies in the conduction band. Therefore the ground state has the extra Si electron trapped in a shallow donor level. But for x > 0.3 in $Al_xGa_{1-x}As$, the deep level lies in the gap, below the shallow donor levels, and this deep level is occupied by one electron in the ground state. Thus a second electron of opposite spin can be trapped by the deep level: when the deep level lies in the gap, the Si no longer is a donor but instead removes electrons from the conduction band. Thus Si or the DX center undergoes a shallow-deep transition from an impurity that produces n-type shallow-donor semiconductive behavior to one that is a deep trap and leads to semi-insulating properties. This same type of behavior occurs as a function of increasing pressure for the DX center in GaAs [14], and was predicted theoretically [4]. (See also Ref. [13].) Finally, Ren et al. [1] have developed a theory of deep levels in superlattices which shows that such a DX center in a GaAs quantum well of a GaAs/$Al_xGa_{1-x}As$ superlattice can descend into the gap as the GaAs well-thickness a decreases because the deep level stays relatively constant in energy while the conduction band edge of the superlattice increases (due to quantum confinement) as the layer thickness decreases -- until it passes above the deep level, causing the DX level to lie in the gap as an electron trap. See Fig. 2.

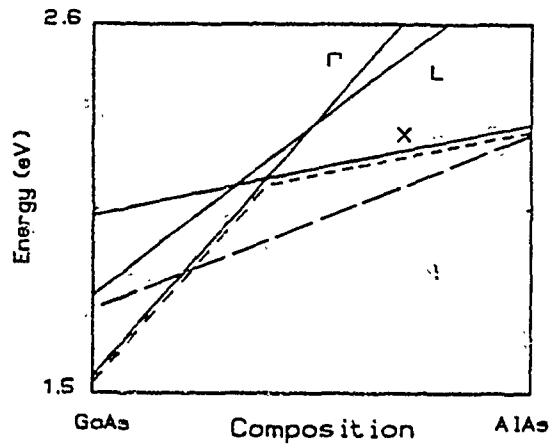


Fig. 1. Energies versus alloy composition x in $Al_xGa_{1-x}As$ alloys of (i) the conduction band-edge minima at Γ, X, and L (solid), (ii) shallow donor states (short dashed), and (iii) DX center (long dashed), after Refs. [2] and [7].

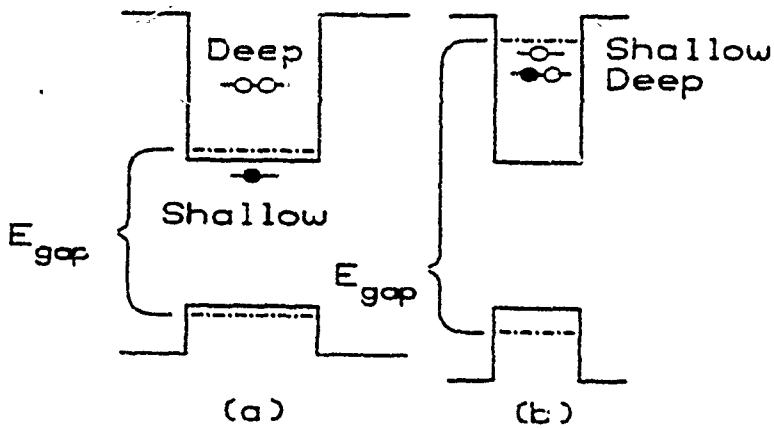


Fig. 2. Illustration of a shallow-deep transition as the width of a GaAs quantum well decreases. In a thick well (a) the deep donor level lies above the conduction-band edge, so that the electron occupies the lower shallow level; in a thin well (b) the shallow level lies just below the superlattice conduction band edge and above the deep level, which is occupied by the electron.

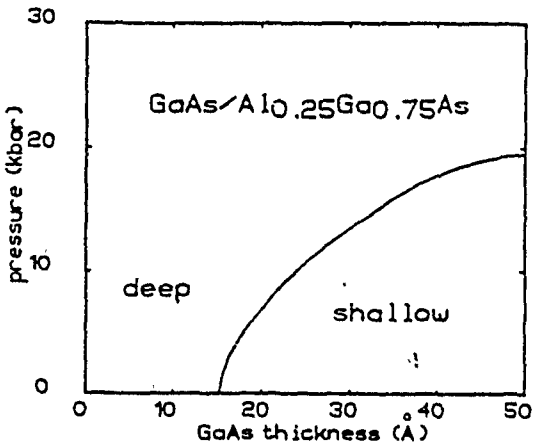


Fig. 3. Predicted critical pressure versus GaAs quantum-well width for a GaAs/$Al_{0.25}Ga_{0.75}As$ well. Here we have N_2=10.

In Fig. 3 we show our predictions for the "phase diagram" of a DX center in the center of a GaAs well of GaAs/$Al_{0.25}Ga_{0.75}As$, namely the function of pressure and well-thickness that causes the A_1 deep level of the DX center to coincide with the conduction band edge. Thus, in the region marked "shallow", the DX impurity produces a deep level in the conduction band of the superlattice, and is a shallow donor; in the "deep" region the DX center is a deep trap. A word of caution about the theory: the general shape of the curve $p(\underline{a})$, where p is the pressure and $\underline{a}$ is the well thickness, is reliable, whereas the precise values may not be. Indeed, we suspect that at zero pressure the critical well-thickness for a DX center to produce a deep level in the gap is somewhat smaller than predicted.

ACKNOWLEDGMENTS

We are grateful to M. F. Li for a discussion of the DX center data, and to the Office of Naval Research for their generous support (Contract No. N00014-84-K-0352).

REFERENCES

[1] S. Y. Ren, J. D. Dow, and J. Shen, "Deep impurity levels in semiconductor superlattices," to be published.
[2] H. P. Hjalmarson, P. Vogl, J. D. Dow, and D. J. Wolford, Phys. Rev. Lett. 44, 810 (1980).
[3] D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).
[4] S. Y. Ren, J. D. Dow, and D. J. Wolford, Phys. Rev. B 25, 7661 (1982).
[5] D. J. Wolford, T. F. Kuech, J. A. Bradley, M. A. Gell, D. Ninno, and M. Jaros, J. Vac. Sci. Technol. B4, 1043 (1986).
[6] D. V. Lang and R. A. Logan, Phys. Rev. Lett. 39, 635 (1977).
[7] H. P. Hjalmarson, private communication. See also Ref. [2].
[8] M. Mizuta, M. Tachikawa, H. Kukimoto, and S. Minomura, Jpn. J. Appl. Phys. 24, L1431 (1985).
[9] H. P. Hjalmarson and T. J. Drummond, Appl. Phys. Lett. 48, 656 (1986).
[10] See also M. F. Li and P. Yu, Solid State Commun. 61, 13 (1987) for an opposing viewpoint.
[11] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[12] In the virtual crystal approximation, the diagonal matrix elements of AlAs and GaAs are averaged, with weights x and 1-x, and the off-diagonal matrix elements times the square of the bond length are also averaged. See, for example, Ref. [2].
[13] E. Yamaguchi, Jpn. J. Appl. Phys. 25, L643 (1986).
[14] The critical pressure p_c for GaAs is approximately 25 kbar.

Solid State Communications, Vol.64,No.4, pp.431-43-, 1987.
Printed in Great Britain.

0038-1098/87 $3.00 + .00
Pergamon Journals Ltd.

THEORY OF RAMAN SPECTRA OF CORRELATED SUBSTITUTIONAL ALLOYS

Andrew C. Redfield* and John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556, U.S.A.

(Received 3 June 1987 by A. A. Maradudin)

A method for computing the Raman spectra of substitutional crystalline III-V alloys, combining Monte Carlo and Recursion methods, is presented and applied to $In_{0.5}Ga_{0.5}As_{0.5}Sb_{0.5}$.

1. Introduction

This paper describes a simple method for computing Raman scattering spectra for substitutional crystalline semiconductor alloys with correlations between atoms occupying neighboring sites. Most theories assume that such alloys are random, and omit correlations between the atoms on neighboring sites: The probabilities of having different types of atoms on a given site depend only on the concentrations of the constituents and not on the types of atoms on neighboring sites. In reality, however, alloys often have a tendency to cluster and form regions with preferential bonding. For example, Islam and Bunker [1] have recently observed preferential Ga-As and In-Sb bonding in $In_{1-x}Ga_xAs_ySb_{1-y}$ alloys; as a result, the fraction of Ga-As bonds differs significantly from the random value xy. With the increasing technological importance of III-V ternary and quaternary alloys -- alloys that exhibit different amounts of clustering and correlations for different growth conditions -- it has become important to develop a theory capable of predicting the electronic and vibrational properties of correlated alloys. Such a theory is needed especially for phonon spectra, which unlike electronic spectra are normally persistent [2] or two-mode [3] in character, and most often are not well approximated by a virtual crystal approximation (which, with few exceptions [4-6], provides an adequate description of the electronic states near the fundamental band gaps of common semiconductors).

In this paper, we present such a theory of phonons in $In_{1-x}Ga_xAs_ySb_{1-y}$; the theory has three distinct elements: (i) the determination of an 8,000-atom cluster of atoms that has the desired nearest-neighbor correlations, (ii) the approximation of the Raman spectrum by an appropriate projected density of states of the alloy [7], and (iii) evaluation of this density of states for the alloy using the recursion method [8] and a Born-von Karman model of the lattice vibrations [7,9-12].

2. Correlations

The first task in developing our theory of correlations is to determine the site occupations of a large 8,000-atom cluster such that the cluster has the desired nearest-neighbor correlations. We place atoms Ga or In on cation sites (with probabilities x and 1-x, respectively) and As or Sb on anion sites. However, constructing a cluster with the desired nearest-neighbor correlations is somewhat more difficult than creating an uncorrelated cluster. In the case of an uncorrelated cluster, one merely deposits atoms on sites with the prescribed probabilities, using a random number generator to determine which atom occupies a given site. However, if this approach were applied to create a correlated cluster, the correlations in different parts of the cluster would be different: when putting down the first atom, none of its neighbors is known; when depositing the second atom, at most one of its neighbors is known; and so on, until all of the neighbors of the last atom are known. As a result, higher-order correlations are present which depend on the sequence in which the sites are occupied.

To circumvent this problem, we employ a four-component Ising-like model of the cluster, and apply the standard Monte Carlo procedure [13,14] to solve the model for an alloy configuration with the desired average alloy composition and nearest-neighbor correlations. The energy E of the configuration is

$$E/(k_BT) = \Sigma_{[\vec{R},\vec{R}']} J(\vec{R},\vec{R}') + \Sigma_{\vec{R}} h(\vec{R}),$$

where we have $h(\vec{R}) = H_\nu$ if atom-type ν is at site $\vec{R}$, and $J(\vec{R},\vec{R}') = J_{\mu,\nu}$ if atom-type μ is at $\vec{R}$ and ν is at $\vec{R}'$. (Here $J_{\mu,\nu}$ is zero unless $\vec{R}$ and $\vec{R}'$ are nearest-neighbors.) The numbers H_ν and $J_{\mu,\nu}$ are independent parameters which can be adjusted to provide the desired nearest-neighbor correlations. The probability of a given 8,000-atom alloy configuration is proportional to $\exp(-E/k_BT)$. Examining the above expression for E, we see that, if we change a single atom in the configuration, the probability that the new atom is of type ν depends only on that atom and its four nearest-neighbors, through H_ν and $J_{\mu,\nu}$.

Our procedure is to first solve this Ising-like model (assuming values of $J_{\mu,\nu}$ and H_ν) for an equilibrium alloy configuration, using ordinary Monte Carlo techniques [14], and then to adjust the parameters $J_{\mu,\nu}$ and H_ν by trial and error [15] until we find an equilibrium configuration with the desired

* Present address: School of Physics, Georgia Institute of Technology, Atlanta, GA 30332.

nearest-neighbor correlations. As usual, enough iterations of the Monte Carlo scheme are performed to achieve convergent values of the average alloy compositions x and y and the nearest-neighbor correlations $N_{\mu,\nu}$; and the results are independent of the initial configuration and detailed method of relaxation. Here, for example, $N_{Ga,As}$ is the fraction of GaAs bonds [16].

For the quaternary alloys considered here, we assume that only Ga and In atoms occupy cation sites, and only As and Sb occupy anion sites of a zincblende lattice. This condition is achieved in all configurations by choosing the initial configuration to satisfy it, and by taking $J_{cation,cation}$ and $J_{anion,anion}$ to be infinite.

The net result of the Monte Carlo procedure is a single final cluster configuration that has the desired average composition and nearest-neighbor correlations. If the correlations are so strong that there is considerable clustering, a single 8,000-atom cluster may not mimic the real alloy sufficiently well, and it may be necessary to generate an ensemble of such clusters. In such a case, the calculations described below would be repeated for each cluster and then ensemble-averaged.

3 Raman Spectra

An explicit expression for the intensity $I(\omega')$ due to Raman scattering of light into a solid angle $d\Omega'$ is derived in detail in the textbook of Born and Huang [7]. In particular, they show that only the electronic portion of the polarizability P_{jk} is relevant to Raman processes [17]. This contribution should be very similar on all cation (anion) sites, since different III (V) elements are electronically similar. Therefore, to a good approximation P_{jk} in a zincblende crystalline alloy will be the same as and have the same symmetries as, P_{jk} for a pure zincblende crystal. Moreover, only the terms P_{jk} of first order in the displacements u_ℓ are needed for lowest-order Raman scattering. These terms can be written simply in terms of their decomposition into displacements $u_\ell(n,\sigma)$ at individual sites in cell n, and atom σ [18] as

$$\delta P_{jk} = \Sigma_{n,\sigma,\ell}\, P_{jk,\ell}(0,\sigma) u_\ell(n,\sigma)$$

In the zincblende structure, there are two atoms per cell (σ=1,2) and, from translational symmetry [19], they must have $P_{jk,\ell}$ that are equal in magnitude and opposite in sign. The form of $P_{jk,\ell}$ is determined largely by the symmetry of the crystal. The Raman-allowed modes and corresponding forms of $P_{jk,\ell}$ for various crystal structures have been tabulated by Hayes and Loudon [20]. For the zincblende structure (symmetry $\bar{4}3m$), one finds that the only Raman-active optical mode is Γ_5 and the corresponding tensor $P_{jk,\ell}$ has components P_0 (an unknown constant) whenever j, k, and ℓ are all different, and zero otherwise.

Born and Huang [21] show that the Raman scattering intensity at temperature T for one-phonon absorption (i.e., the lowest-order anti-Stokes process), in terms of the creation and annihilation operators E^+ and E^- for the incoming photon of energy $\hbar\omega$ and the polarization vector $\bar{e}$ of the outgoing photon of energy $\hbar\omega'$, is

$$I(\omega') = (\omega^4/2\pi c^2)\, \Sigma_{j,k,\ell,m}\, e'_j e'_\ell\, i_{jk,\ell m}\, E^-_k E^+_m,$$

where we have

$$i_{jk,\ell m} = \frac{P_{jk}(\lambda) P_{\ell m}(\lambda)\, \hbar}{(2\omega_\lambda)(\exp(\hbar\omega_\lambda/k_B T) - 1)}.$$

Here $P_{jk}(\lambda)$ is the projection of the polarization tensor discussed above onto the λ-th eigenstate, of frequency ω_λ [22]. Using the form of $P_{jk,\lambda}$ discussed above, the scattering is independent of direction and polarization, except for an overall scale factor, so the results can be discussed without reference to the crystal axes.

4. Recursion Method

The preceeding discussion demonstrates that the Raman spectrum is simply related to the projection of the vibrational density of states onto the polarizability tensor [22]. For a disordered alloy, there is no way to determine this density of states exactly. However, the recursion method [8] provides a simple, fast method to obtain approximate solutions numerically for a large cluster. In fact, what it really provides is not a total density of states, but the projected density of states onto a given state. Thus the method is ideal for calculations of the Raman scattering, with the projected state the polarizability tensor found above [22]. The resulting calculation is really just a spectral density of states computation, and the detailed method of application of the recursion technique to such problems has been discussed by Davis [6].

The phonon states are modeled by a simple Born-von Karman model with first- and second-nearest-neighbor force constants; i.e., by the eigenvalue equation for the eigenfrequencies ω_λ:

$$M(n,\sigma)\, \omega_\lambda^2\, \bar{u}(n,\sigma) =$$

$$\Sigma_{n',\sigma'}\, \underline{F}(n,\sigma;n',\sigma').[\bar{u}(n',\sigma') - \bar{u}(n,\sigma)].$$

Here the F's are taken to be the most general force constant matrices consistent with the point group symmetry of atom n,σ, and $M(n,\sigma)$ is its mass. The force constant matrix contains two independent parameters for nearest-neighbors, and three for second-nearest-neighbors [9-12]. There are no long-ranged forces in this model, and so the usual Lyddane-Sachs-Teller splitting of the longitudinal and transverse optical modes at zero crystal momentum vanishes. Following earlier work [11,12], we use the force constants of GaAs and treat the alloy disorder as occurring exclusively in the mass matrix.

5. Illustrative Results and Conclusions

The calculated one-phonon absorption Raman spectrum for $In_{0.5}Ga_{0.5}As_{0.5}Sb_{0.5}$ at room temperature is plotted in Fig. . Three curves are displayed, corresponding to three possible correlations. The three peaks in the curves are due to In-Sb (heavy-heavy), In-As and Ga-Sb (heavy-light), and Ga-As (light-light) bonds,

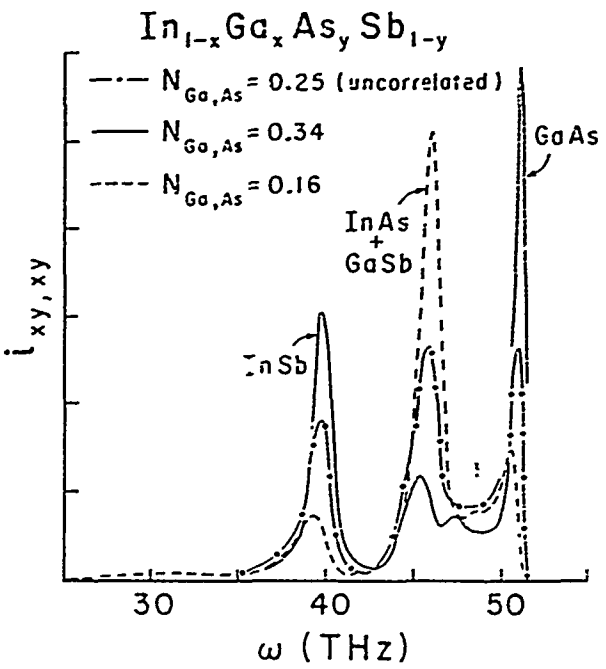


Fig. 1. Theoretical Raman spectrum $i_{xy,xy}$ for anti-Stokes scattering at room temperature in $In_{0.5}Ga_{0.5}As_{0.5}Sb_{0.5}$. The chained curve is for Ga and its nearest-neighbors 25% of the time (i.e., no correlations), the solid curve 34%, and the dashed curve 16%.

respectively. The dotted curve corresponds to no correlations; i.e., the allowed types of atoms are placed randomly on each site. The solid curve corresponds to more Ga-As and In-Sb bonds, and the corresponding peaks are higher while the central one is lower. (This is the actual situation according to the extended x-ray absorption fine structure experiments [1].) The dashed curve corresponds to more In-As and Ga-Sb bonds, and shows a higher central peak and reduced side ones. This is qualitatively as expected: the area under any given peak is roughly proportional to the number of bonds of the corresponding type(s), as is true for the corresponding peaks in the density of states.

However, the relative areas under different peaks depend quantitatively on a number of factors, and differ markedly from the corresponding areas for the density of states curves. First, there are various peaks in the density of states (at lower energies, in particular) that are Raman forbidden and do not show up at all in the Raman spectra. Second, the relative areas under different peaks are affected by the ratio of the phonon energy to the temperature. Third, the relative areas depend on the projection of the states onto the optical state of zero crystal momentum. This depends primarily on the ratio of the peak separation to the optical band-width (of the pure material), and explains the reduction in size of the lower-frequency peaks [2].

These considerations are important in attempting to determine the near-neighbor correlations from Raman data. Until now, experimentalists have, in the absence of any theoretical calculations for comparison, tried to estimate the number of nearest-neighbors of different atomic types by comparing the areas under the different peaks. While this provides a correct order-of-magnitude estimate of numbers of neighbor pairs, it ignores the factors discussed above and therefore does not provide accurate estimates of these correlations. The techniques described in this paper can be used to determine quantitatively the Raman spectra for various correlations. These can then be compared with experimental results to allow an accurate determination of the actual correlations [23].

The present approach allows one to determine the effects of clustering and correlations on the Raman spectra of III-V alloys, and can easily be generalized to predict densities of states and spectra of a variety of other alloys. The method is not restricted solely to the treatment of vibrational properties or to III-V alloys, but is generally applicable to a wide range of problems in many different substitutional crystalline alloys.

Acknowledgment -- We are grateful to the U.S. Office of Naval Research (Contract No. N00014-84-K-0352) and to the U.S. Air Force Office of Scientific Research (Contract No. AFOSR-85-0331) for their support. We also wish to thank Craig S. Lent, Kathie E. Newman, and Bruce A. Bunker for helpful discussions about correlated clusters, and Andrew Zangwill for making useful suggestions concerning the manuscript.

REFERENCES

[1] S. M. Islam and S. A. Bunker, Bull. Amer. Phys. Soc. 30, 603 (1985); 31, 667 (1986); and to be published.

[2] Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968).

[3] G. Lucovsky, K. Y. Cheng, and G. L. Pearson, Phys. Rev. B 12, 4135 (1975).

[4] A.-B. Chen and A. Sher, Phys. Rev. Lett. 49, 948 (1982); W. E. Spicer, J. A. Silberman, J. Morgan, I. Lindau, J. A. Wilson, A.-B. Chen, and A. Sher, Phys. Rev. Lett. 49, 948 (1982).

[5] K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B 27, 1088 (1983).

[6] L. C. Davis, Phys. Rev. B 28, 6961 (1983).

[7] M. Born and K. Huang, Dynamical Theory of Crystal Lattices, (Oxford University Press, London, England, 1954).

[8] See R. Haydock in Solid State Physics, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, New York, 1980), vol. 35, p. 215, for a discussion of the recursion method and the Cambridge Recursion Library.

[9] R. Banerjee and Y. P. Varshni, Can. J. Phys. 47, 451 (1969).
[10] A sixth force constant, δ in the notation of Ref. [9], is assumed to be zero.
[11] A. Kobayashi, J. D. Dow, & E. P. O'Reilly, Superlattices & Microstructures, 1, 471 (1985).
[12] A. Kobayashi, K. E. Newman, and J. D. Dow, Phys. Rev. B 32, 5312 (1985).
[13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. 12, 1087 (1953).
[14] For a review of Monte Carlo methods and application to Ising-type models, see K. Binder, ed., Monte Carlo Methods in Statistical Physics (Springer-Verlag, New York, 1979).
[15] Typical values of $J_{\mu,\nu}$ and H_ν range from -0.6 to +0.6.
[16] $N_{Ga,As} = (8N)^{-1} \Sigma_{\bar{R},\bar{R}'} n_{Ga}(\bar{R}) n_{As}(\bar{R}') \Delta(\bar{R},\bar{R}')$, where $\Delta(\bar{R},\bar{R}')$ vanishes unless $\bar{R}$ and $\bar{R}'$ are nearest-neighbors, $n_\mu(\bar{R})$ is the occupation number for atom of type μ at site $\bar{R}$ (and is either zero or unity), and the sums run over both anion and cation sites $\bar{R}$.
[17] Ref. [7], p. 207.
[18] Ref. [7], Eqs. 23.3 and 39.12.
[19] Ref. [7], Eq. 23.27.
[20] W. Hayes and R. Loudon, Scattering of Light by Crystals (John Wiley & Sons, New York, 1978), Tables 3.1 and 1.2.
[21] Ref. [7], Eqs. 49.1 and 21.8.
[22] Ref. [7], Eqs. 39.12 and 39.14.
[23] In comparison with experiment, the computed widths of the peaks here (especially the narrow peaks) cannot be trusted, since effects such as the finite lifetime of the quasi-particle states have been ignored. For instance, for a pure material, the calculations here give a δ-function at the top of the optical band, while experimental curves usually have a rather sizeable width to the peak.

Scanning tunneling microscope tip structures

Ruth Nicolaides
U.S. Army ARDEC, Dover, New Jersey 07801

Yong Liang,[a] William E. Packard,[a] Zhou-Wu Fu,[a] Howard A. Blackstead,[a] K. K. Chin,[a] John D. Dow,[a] Jacek K. Furdyna,[a] Wei Min Hu,[a] Robert C. Jaklevic,[a,b] William J. Kaiser,[a,c] Alan R. Pelton,[d] and Mary V. Zeller[e]
University of Notre Dame, Notre Dame, Indiana 46556

Joseph Bellina, Jr.
Department of Physics and Chemistry, St. Mary's College, Notre Dame, Indiana 46556

(Received 24 September 1987; accepted 29 October 1987)

Studies of electrochemically etched tungsten scanning tunneling microscope tips, using scanning electron microscopy, show that (i) the tips are often convoived or bent if the mass of the tungsten wire submerged in the etchant is large (an effect ascribed to surface plastic flow), (ii) bent tips nevertheless often produce good quality scanning tunneling microscopy images of Au films in air, but (iii) tips, once crashed clumsily into the Au films, no longer produce images.

I. INTRODUCTION

In an ideal scanning tunneling microscope, electron tunneling occurs between the surface being studied and a single atom at the end of a sharp tunneling tip. In practice, it is rare to prepare tips even resembling the sharp, single-atom ideal. Often, in order to obtain scanning tunneling microscope images from layered compounds, one must first crash the tungsten tip into the surface. This initial crash very likely "spears" a layer of the material being studied, which then can act as a tunneling tip. For example, scanning tunneling microscope studies of layered compounds, such as graphite, using tungsten tips suggest that the "tip" may in reality be a layer of graphite stuck on the tungsten.[1,2] Colton *et al.*[1] and Mizes and Harrison[3] have shown rather dramatically that many of the different images reported for graphite surfaces can be obtained by having more than one atom acting as a tunneling site. In III–V semiconductors, Feenstra and Fein[4] have shown that images of defects on the GaAs (110) surface depend on the character of the tip as much as on the defect. Biegelsen *et al.*[5] have published studies of tip structures and have found that ion milling improves the sharpness of a tip, removes oxide, and enhances the tip's reliability. Clearly the role of the tip and its geometry in forming scanning tunneling microscope images is incompletely understood.

In this paper, we report some elementary studies of tungsten scanning tunneling microscope tips. These include studies of scanning electron microscope images of tips, the dependence of tip geometry on tip etching and growth conditions, and the quality of scanning tunneling microscope images obtained from each tip. As our touchstone of comparison, we use images of Au films in air. Surface Au atoms have a high mobility, forming nearly planar surfaces, and the steps on these surfaces are easily visible with our microscope. We use Au rather than graphite as our standard because graphite layers are too easily peeled from the surface. We find, not surprisingly, that once our tips crash into the surface of Au, unless the crash is rather gentle,[6] the tips no longer produce good images; however, we also find that tip geometry, as observed with a scanning electron microscope, can be a deceptive predictor of scanning tunneling microscope image quality. In particular, some tips can be terribly "bent" or convolved geometrically and yet produce rather good images.

II. TIP PREPARATION

Each tip was prepared by placing several millimeters of the lower end of a tungsten wire (0.025 in. diameter) into an aqueous 1M NaOH etching solution and applying a 12-V potential to the tungsten wire (with respect to a stainless-steel electrode inserted into the solution). The etch was continued until the submerged portion of the wire dropped off into the bath, leaving the usable tip suspended near the liquid/air interface. By electronically monitoring the etching current (typically 10 mA) with a comparator circuit, the 12-V potential was shut off when the wire separated. This prevented further etching of the tip. After the separation, the etch voltage was pulsed "on" for 1 s, to remove any irregularities at the end of the tip.

In order to prevent unnecessary etching, we covered a large portion of the wire submerged in the solution with Teflon insulation. This kept the current density in the etching region approximately constant and permitted better determination of the mass of the submerged portion of the wire (for correlation of tip shape with the mass of the submerged portion, see below).

III. SCANNING ELECTRON MICROSCOPE IMAGES

The tips that we etched generally exhibited nearly exponential shapes (Fig. 1), rather than the nearly parabolic shapes reported by some authors.[7] We find that this exponential shape results when the current density (and hence the reaction rate) is high. We have observed that, with a longer length of wire (~1 cm or more) exposed to the etchant, the profile of the tip tended to become more parabolic. We have also found that more parabolic shapes result from electrochemical etching with alternating rather than direct current.

A number of our tips had, in addition to the nearly expo-

 © 1988 American Vacuum Society

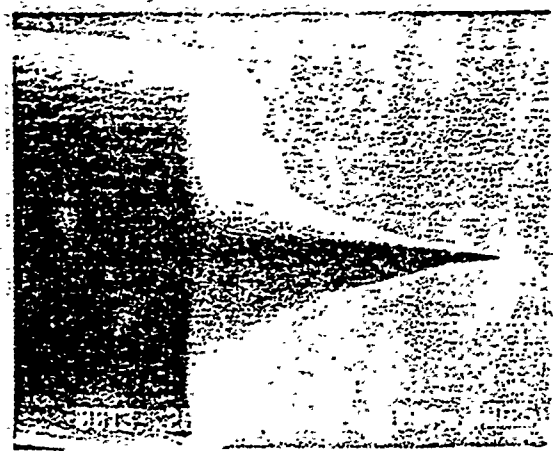


FIG. 1. Scanning electron microscope image of a pointed tungsten tip. The tip was etched in an aqueous NaOH solution. Note the 100-μm scale.

nential overall shape, highly convolved or bent points (Fig. 2), although such tips had never been allowed contact with any surface. By carefully controlling the etching conditions, we learned that such bent points tend to occur when the mass (length) of the tungsten wire in the etchant is large, a condition indicative of (i) plastic flow of the tungsten wire as the tip is formed and (ii) some recoil of the tip at the instant of tip formation, when the wire in the etchant drops off.

Just before the bottom portion of the wire separates, plastic flow occurs at the narrowest region of the wire when the stress induced by the wire's weight is greater than the yield stress. Rough estimates indicate that the weight of several millimeters of tungsten wire in the etchant bath is sufficient to allow plastic deformation at a necking diameter of about 1 μm. Furthermore, the mechanical energy stored in the neck region of the stretched tungsten wire is released when the wire separates. This energy, although perhaps an order of magnitude too small to plastically deform the entire volume of the thicker portion of the wire, is nevertheless sufficient to deform small surface regions, leading to tip recoil and bending, with the yield stress apparently being exceeded locally at certain surface regions.

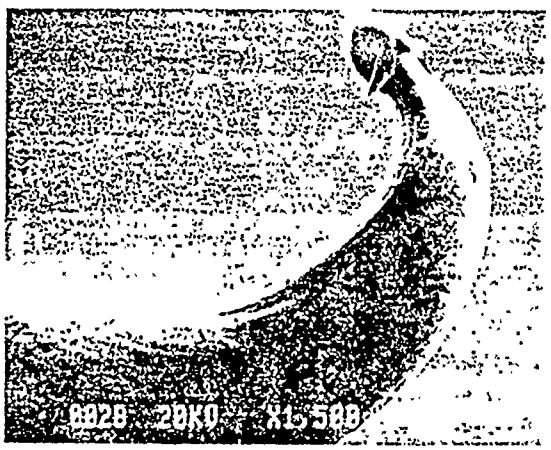


FIG. 2. Scanning electron microscope image of a contorted or bent tungsten tip. The contortion is not due to the tip's having been crashed, but rather is due to tip deformation during etching. Such bent tips occur when a large mass of wire is submerged in the etchant and are ascribed to recoil after fracture resulting from plastic deformation. Note the 10-μm scale. Note also the "dirt" on the tip, residual NaOH.

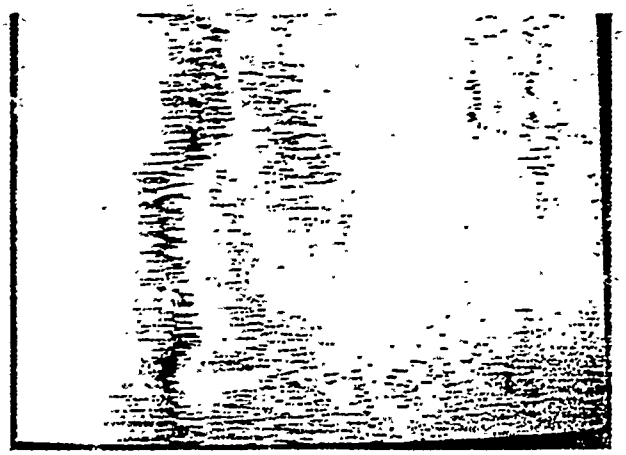


FIG. 3. Scanning tunneling microscope image of a Au film, taken using the tip of Fig. 1. This is a tunneling current image over a 150 Å × 150 Å area of the film. Comparable quality images are obtained with both pointed and bent tips, provided the tips have not been crashed.

The possibility of plastic flow playing a role in the formation of tips has been raised previously by Müller and Tsong,[8] but those authors ascribed the tip bending to the action of gas bubbles. We virtually eliminate such bubbles by using a direct-current etch (alternating current produces many bubbles), but still obtain bent tips when the conditions of significant plastic flow are met.

IV. SCANNING TUNNELING MICROSCOPE IMAGES

Surprisingly, the bent or convolved tips often produced decent scanning tunneling microscope images—of comparable quality with images produced by "pointed" tips, such as the one in Fig. 3. Subjectively, the pointed tips may have produced slightly sharper scanning tunneling microscope images, but the variation of image quality for various pointed tips was comparable with the differences between images for pointed and bent tips.

In contrast, tips that were crashed clumsily into the surface no longer produced images. ("Controlled and gentle crashes, however, can leave the tips capable of forming subsequent images.")

V. CONCLUSIONS

Thus we conclude that the best tungsten tips are formed when only a small portion of the wire is suspended in the etchant, and that the sharpness of a tip on the ~ 10-μm scale of a scanning electron microscope image may not be a good indicator of tip imaging quality. Nevertheless, as a matter of good experimental practice, bent tips should be avoided, and so only a small portion of the tungsten wire should be submerged during the etching process.

ACKNOWLEDGMENTS

We are grateful for the generous support of the U.S. Army Research Office, the Defense Advanced Research Projects Agency, and the Office of Naval Research (Contracts No.

DAALO3-86G-0179. No: N0530-0716-05. and No. N00014-84-K-0352). One of us (W.J.K.) acknowledges the support of SDIO/IST.

[a] Department of Physics.
[b] Also Department of Physics; permanent address. Ford Motor Company, Dearborn, MI 48121.
[c] Present address: Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.
[d] Department of Materials Science.
[e] College of Engineering.

[1] R. J. Colton, S. M. Baker, R. Driscoll, J. D. Baldeschwieler, and W. J. Kaiser, J. Vac. Sci. Technol., A 6, 349 (1988) (these proceedings).
[2] R. J. Colton, S. M. Baker, J. D. Baldeschwieler, and W. J. Kaiser, Appl. Phys. Lett. 51, 305 (1987).
[3] H. A. Mizes and W. A. Harrison, J. Vac. Sci. Technol. A 6, 300 (1988) (these proceedings).
[4] R. M. Feenstra and A. P. Fein, Phys. Rev. B 32, 1394. See also, R. M. Feenstra and J. A. Stroscio, J. Vac. Sci. Technol. B 5, 923 (1987).
[5] D. U. Biegelsen, F. A. Ponce, J. C. Tramontana, and S. M. Koch, Appl. Phys. Lett. 50, 696 (1987).
[6] R. C. Jaklevic and L. Elie, Phys. Rev. Lett. (to be published).
[7] C. J. Chen, J. Vac. Sci. Technol. A 6, 319 (1988) (these proceedings).
[8] E. W. Müller and T. T. Tsong, *Field Ion Microscopy* (Elsevier, New York, 1969), p. 122; see also Y. Yashino, Oyo Butsuri (Appl. Phys. Jpn.) 33, 912 (1964).

Electronic structures and doping of InN, $In_xGa_{1-x}N$, and $In_xAl_{1-x}N$

David W. Jenkins* and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 23 December 1987)

The electronic structures of InN, $In_xGa_{1-x}N$, and $In_xAl_{1-x}N$ are predicted and these materials are found to be direct-band-gap semiconductors with fundamental band gaps ranging from orange through the blue-green to the ultraviolet. The deep levels associated with substitutional *s*- and *p*-bonded impurities are predicted, and, for InN we find (i) that the native defect responsible for naturally occurring *n*-type InN is a nitrogen vacancy (not N_{In}); (ii) that the nitrogen vacancy also produces a deep level just below the conduction-band edge, which is responsible for an observed 0.2-eV optical-absorption feature; (iii) that *p*-type doping should be achievable by inserting column-II impurities on In sites; (iv) that *n*-type conductivity should result from oxygen atoms on N sites; (v) that In_N produces *s*- and *p*-like deep levels near midgap that are responsible for an optical-absorption feature near 1 eV; (vi) that column IV impurities on either anion or cation sites will tend to make the material semi-insulating, and (vii) that an isoelectronic electron trap should be produced by B_{In}, whereas column-V impurities on the N site should produce deep isoelectronic hole traps. Similar results hold for the alloys $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$. Some impurities undergo shallow-deep transitions in the alloys, as functions of alloy composition.

I. INTRODUCTION

High-mobility InN has recently been grown in a polycrystalline hexagonal structure.[1,2] The band gap of this material is direct and the optical absorption threshold lies in the orange portion of the visible spectrum (the band gap plus the Burstein shift is 2.05 eV). This is an exciting experimental result for at least two reasons. (i) the high mobility suggests that electronic-grade material may eventually be fabricated, and (ii) the orange color indicates that InN and alloys based on InN could be candidates for efficient semiconducting large-band-gap visible-light emitters and lasers. Until now much of the emphasis on developing large-band-gap visible solid-state light emitters has focused on II-VI compound semiconductors, materials that in many cases have proven difficult to dope both *n* type and (especially) *p* type[3] and hence do not form good diodes, much less light-emitting diodes. The origin of the II-VI compound doping problem is often ascribed to "self-compensation"—common dopants purportedly distort off site and produce accompanying vacancies which compensate them.[4] It is widely believed that such self-compensation problems do not plague III-V compound semiconductors, and so the existence of the isoanionic semiconductors InN, GaN, and AlN, all with large band gaps [2 eV (orange), 3.5 eV (ultraviolet), and 6 eV (ultraviolet), respectively], raises the possibility of fabricating alloys whose band gaps range from the orange to the ultraviolet.

Assuming that the problem of growing electronic-grade material can be solved, there will remain five major criteria that the material must meet. (i) the band gap must be the desired color (orange, blue, etc.); (ii) the band gap must be direct so that the crystal-momentum selection rule governing light emission[5] will be satisfied; (iii) the material must be crystalline and, if it is an alloy, must be relatively strain free—because large strains produce dislocations, and dislocations tend to quench luminescence (as well as trap and scatter carriers and degrade mobility);[6,7] (iv) the materials must be relatively free of deep levels in the band gap that might trap electrons or holes, leading to enhanced nonradiative transition rates and luminescence degradation, and (v) schemes for doping the material both *n* and *p* type must be found. The purpose of this paper is to provide theoretical guidance concerning these five issues, in the hope of stimulating efforts to grow electronic-grade InN and InN-based alloys.

In Sec. II we discuss the band structures of (wurtzite) InN and alloys of InN and GaN and AlN. We show that these alloys can be described by the virtual-crystal approximation, have direct band gaps that range from the orange to the ultraviolet, and should be relatively strain free because they are moderately well lattice matched. In Sec. III we discuss the deep levels associated with *s*- and *p*-bonded substitutional impurities in InN, with particular emphasis on the native defects (i.e., vacancies and antisite defects) and the dopants from columns II, IV, and VI of the Periodic Table. Section IV is devoted to a comparable discussion for the alloys. Our conclusions are summarized in Sec. V.

II. BAND STRUCTURES

The band structures are obtained using a nearest-neighbor tight-binding model of the electronic structures, based on the Slater-Koster[8] theory. The resulting Hamil-

 ©1989 The American Physical Society

tonian is

$$H=\begin{bmatrix} H_a & 0 & H_{1,3} & H_{1,4} \\ 0 & H_a & H_{1,4} & H_{2,4} \\ H_{1,3}^{\dagger} & H_{1,4}^{\dagger} & H_c & 0 \\ H_{1,4}^{\dagger} & H_{2,4}^{\dagger} & 0 & H_c \end{bmatrix},$$

where H_b is

	$\lvert s,4\rangle$	$\lvert p_z,4\rangle$	$\lvert p_x,4\rangle$	$\lvert p_y,4\rangle$
$\lvert s,1\rangle$	$U(s,s)$	$U(s,z)$	0	0
$\lvert p_z,1\rangle$	$U(z,s)$	$U(z,z)$	0	0
$\lvert p_x,1\rangle$	0	0	$U(x,x)$	0
$\lvert p_y,1\rangle$	0	0	0	$U(y,y)$

$M_{2,4}$ is

	$\lvert s,4\rangle$	$\lvert p_z,4\rangle$	$\lvert p_x,4\rangle$	$\lvert p_y,4\rangle$
$\lvert s,2\rangle$	$f_0^* U'(s,s)$	$f_0^* U'(s,z)$	$-f_1^* U'(s,x)$	$-f_-^* U'(s,x)$
$\lvert p_z,2\rangle$	$f_0^* U'(z,s)$	$f_0^* U'(z,z)$	$-f_1^* U'(z,x)$	$-f_-^* U'(z,x)$
$\lvert p_x,2\rangle$	$-f_1^* U'(x,s)$	$-f_1^* U'(x,z)$	$f_0^* U'(x,x)+f_+^*[U'(x,x)+U'(y,y)]$	$f_-^*[U'(y,y)-U'(x,x)]/2$
$\lvert p_y,2\rangle$	$-f_-^* U'(x,s)$	$-f_-^* U'(x,z)$	$f_-^*[U'(y,y)-U'(x,x)]/2$	$U'(y,y)f_1^*+f_+^*[U'(x,x)+U'(y,y)]/2$

$M_{1,3}$ is

	$\lvert s,4\rangle$	$\lvert p_z,4\rangle$	$\lvert p_x,4\rangle$	$\lvert p_y,4\rangle$
$\lvert s,2\rangle$	$f_0 U'(s,s)$	$f_0 U'(s,z)$	$f_1 U'(s,x)$	$f_- U'(s,x)$
$\lvert p_z,2\rangle$	$f_0 U'(z,s)$	$f_0 U'(z,z)$	$f_1 U'(z,x)$	$f_- U'(z,x)$
$\lvert p_x,2\rangle$	$f_1 U'(x,s)$	$f_1 U'(x,z)$	$f_0 U'(x,x)+f_+[U'(x,x)+U'(y,y)]$	$f_-[U'(y,y)-U'(x,x)]/2$
$\lvert p_y,2\rangle$	$f_- U'(x,s)$	$f_- U'(x,z)$	$f_-[U'(y,y)-U'(x,x)]/2$	$U'(y,y)f_1+f_+[U'(x,x)+U'(y,y)]$

and

$$g_1(\mathbf{k})=\exp[i(-k_1a/3+k_2a/3+k_3c/8)],$$

$$g_2(\mathbf{k})=\exp[i(k_1a/3-k_2a/3+k_3c/8)],$$

$$g_3(\mathbf{k})=\exp(-i3k_3c/8),$$

$$f_0(\mathbf{k})=\exp(+ik_1a)+1+\exp(-ik_2a),$$

$$f_1(\mathbf{k})=\exp(+ik_1a)-[1+\exp(-ik_2a)]/2,$$

$$f_+(\mathbf{k})=\tfrac{3}{4}[1+\exp(-ik_2a)],$$

and

$$f_-(\mathbf{k})=(\tfrac{3}{4})^{1/2}[1-\exp(-ik_2a)].$$

Here we have $\mathbf{k}=k_1\mathbf{b}_1+k_2\mathbf{b}_2+k_3\mathbf{b}_3$; $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$ are the reciprocal-lattice vectors divided by 2π, namely $((2/\sqrt{3})/a,0,0)$, $((1/\sqrt{3})/a,1/a,0)$, and $(0,0,1/c)$, respectively. The parameters of the Hamiltonian for AlN, GaN, and InN have been published,[9-11] and are reproduced in Table I along with the wurtzite lattice constants a and c, and the c/a ratio. Since the c/a ratio, to within 2.1%, is the ideal value of $(\frac{8}{3})^{1/2}$, we simplify the model by assuming the ideal value.

	$\lvert s,1\rangle$	$\lvert p_z,1\rangle$	$\lvert p_x,1\rangle$	$\lvert p_y,1\rangle$
$\lvert s,1\rangle$	$E(s,b)$	0	0	0
$\lvert p_z,1\rangle$	0	$E(p,b)$	0	0
$\lvert p_x,1\rangle$	0	0	$E(p,b)$	0
$\lvert p_y,1\rangle$	0	0	0	$E(p,b)$

for $b=a$ or c (anion site or cation site). The off-diagonal matrices are $H_{1,3}=g_1(\mathbf{k})M_{1,3}$, $H_{2,4}=g_2(\mathbf{k})M_{2,4}$, and $H_{1,4}=g_3(\mathbf{k})M_{1,4}$, where $M_{1,4}$ is

The band structures of the alloys are obtained in the virtual-crystal approximation.[12] We implement the

TABLE I. Tight-binding parameters in eV for AlN, GaN, and InN in the notation of Ref. 12. Parameters are taken from Refs. 9, 10, and 11, respectively.

	AlN	GaN	InN
$E(s,a)$	−12.104	−13.114	−4.984
$E(p,a)$	3.581	1.269	0.565
$E(s,c)$	−0.096	−1.786	0.254
$E(p,c)$	9.419	7.131	3.895
$V(s,s)$	−10.735	−9.371	−3.841
$V(x,x)$	5.808	3.008	1.347
$V(x,y)$	8.486	6.535	3.033
$V(sa,pc)$	8.092	4.889	1.595
$V(pa,sc)$	9.755	10.867	4.000
a	3.104	3.180	3.563
c	4.963	5.166	5.693
c/a	1.599	1.635	1.611

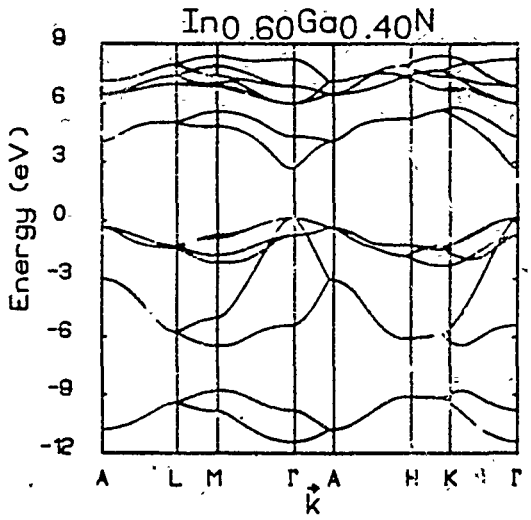


FIG. 1 Predicted band structure of $In_{0.60}Ga_{0.40}N$. The direct energy-band gap lies in the blue-green part of the spectrum.

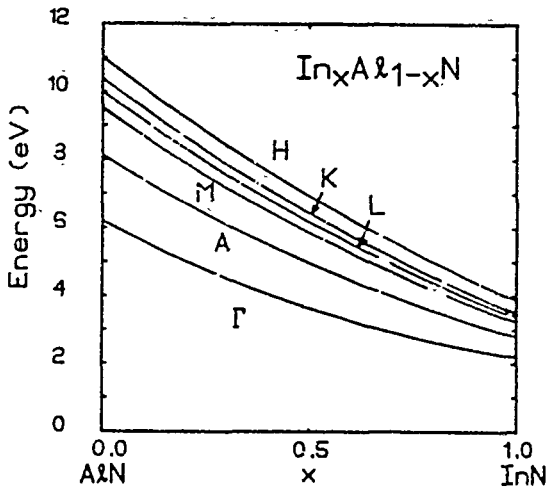


FIG. 3. Energies of principle conduction-band minima vs alloy composition for $In_{1-x}Al_xN$.

virtual-crystal approximation by taking weighted averages of the matrix elements. For example, in the case of $In_xGa_{1-x}N$, we average the various diagonal matrix elements E,

$$E(In_xGa_{1-x}N)=(1-x)E(GaN)+xE(InN)\ .$$

The off-diagonal matrix elements V, multiplied by the square of the bond length d, are also averaged this way (according to Harrison's rule[13]), with the bond length obtained from Vegard's law,[14]

$$d(In_xGa_{1-x}N)=(1-x)d(GaN)+xd(InN)\ .$$

Thus we assume that virtually all of the In and Ga atoms occupy cation sites, while anion sites are overwhelmingly occupied by N. The lattice mismatch, $1-d(GaN)/d(InN)$, is 9.3%; the corresponding mismatch for AlN and InN is 12.8%.[12]

The resulting predicted band structure for $In_{0.60}Ga_{0.40}N$ is given in Fig. 1. $In_xGa_{1-x}N$ is a direct-band-gap semiconductor for all compositions x, and has a band gap ranging from 2 eV for InN to 3.5 eV for GaN. The principal features of the predicted band structures of $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$, namely the energies of the Γ, A, M, L, K, and H conduction-band minima,[15] are plotted as functions of alloy composition x in Figs. 2 and 3.

The principal conclusion to be drawn from these calculations is that $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$ should be direct-band-gap semiconductors and hence potential light emitters for all compositions x.

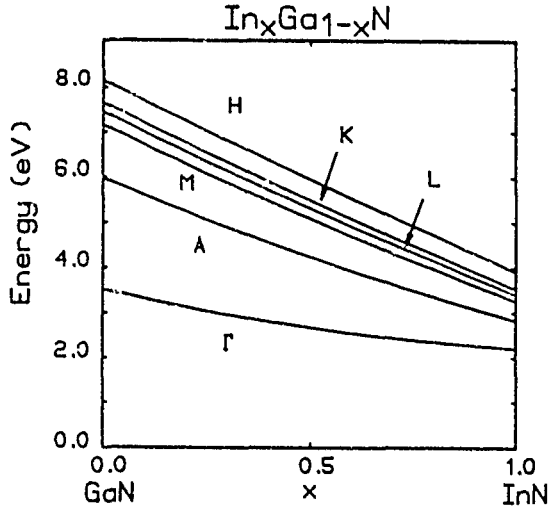


FIG. 2. Energies of principle conduction-band minima vs alloy composition for $In_{1-x}Ga_xN$. The symmetry points H, K, L, M, A, and Γ have the usual Brouckaert-Smoluchowski-Wigner definitions (Ref. 15).

III. DEEP LEVELS IN InN

A. General considerations

Every s- and p-bonded impurity produces both deep levels associated with its central-cell potential and shallow levels caused by any nonzero valence difference between the impurity and the host atom it replaces (although the "deep" levels often do not lie in the fundamental band gap as once believed, but often can be resonant with the host bands). The deep levels can be computed using the theory of Hjalmarson *et al.*[16] The one-electron Schrödinger equation for the deep levels E can be rewritten,

$$\det[1-G_0(E)V]=0\ ,$$

where $G_0(E)$ is the host Green's-function operator,

$$G_0=(E-H_0)^{-1}\ ,$$

and H_0 is the Hamiltonian operator that generates the band structure of the host. E is assumed to have a posi-

tive imaginary part whenever E is not in a band gap. The defect potential is $V \equiv H - H_0$, where H is the Hamiltonian of the crystal with an impurity.

If we follow established custom[16] and (i) neglect the weak long-ranged Coulomb potential responsible for shallow levels, and (ii) ignore lattice relaxation around the impurity (effects of order 0.1 eV), the defect potential V is diagonal in a basis of localized Löwdin orbitals centered at the impurity site, and the secular equation for the energy levels reduces to two[17,18] scalar equations,

$$G_{A_1}(E) = 1/V_s, \quad G_{E_2}(E) = 1/V_p .$$

Here, $G_{A_1}(E)$ and $G_{E_2}(E)$ are the host Green's functions for the s like A_1 and p-like E_2 states,[17] respectively, and V_s and V_p are the defect potentials for the s-like and p-like states, respectively.

Using the scaling rules for the matrix elements of H_0, namely that diagonal matrix elements depend on atomic energies, whereas off-diagonal matrix elements vary inversely as the square of the bond length,[11,13,19] we find that

$$V_s = \beta_s (w_{s,\text{imp}} - w_{s,\text{host}})$$

and

$$V_p = \beta_p (w_{p,\text{imp}} - w_{p,\text{host}}) ,$$

where $w_{l,\text{imp}}$ and $w_{l,\text{host}}$ are atomic-orbital energies in the solid[19] for impurity and host, respectively. We have $\beta_s = 0.8$ and $\beta_p = 0.6$. These equations can be solved for deep levels of energy E in the fundamental band gap by computing the Green's function $G_l(E)$ and plotting E versus $[G_l(E)]^{-1} = V_l$.

B. Native defects

The first question the theory should answer is "Why is InN n type?" Tansley and Foley[1] had speculated some years ago that the n-type behavior is caused by an antisite defect: N on an In site (N_{In}), which they had suggested might be a double donor. However we find that this defect produces both s-like and p-like[17] deep levels deep in the gap (see Fig. 4)—closer to the valence-band edge than to the conduction-band edge.

The s-like state is occupied by the two extra N electrons and is too far from the conduction-band edge to be thermally ionized—even if one makes allowances for a few-tenths-of-an-eV theoretical uncertainty in the predicted deep levels. The p-like states are far from the band edge, empty, and together can trap six electrons. (Of course, Coulombic charge-state splitting, omitted from the model, will raise these neutral-impurity levels as each additional electron is added.[20]) Thus N_{In} is a deep trap for both electrons and holes. Its natural occurrence in InN cannot explain the material's n-type character.

The In_N antisite defect produces deep s-like and p-like levels near the center of the gap (Fig. 4). Six electrons occupy the lowest of the eight spin orbitals associated with this defect, making the neutral defect unquestionably a deep trap for both (two) electrons and (six) holes. Thus In_N can compensate N_{In}, but does not dope InN either n type or p type.

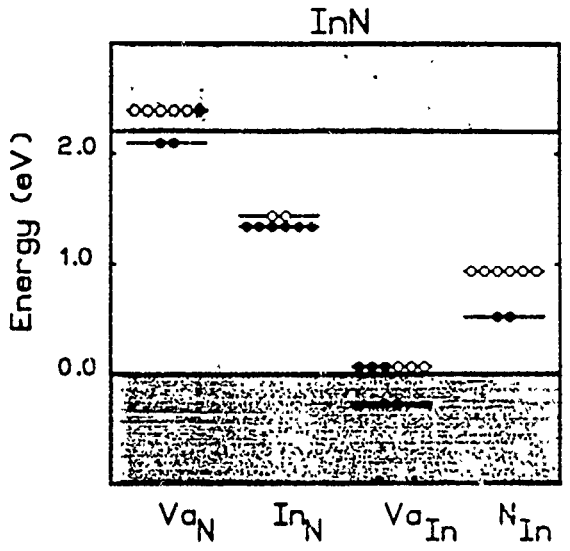


FIG. 4. Energy levels and electron occupancies of neutral native defects in InN. Holes are denoted by open circles and electrons are denoted by solid circles. For the nitrogen vacancy, V_N, the electron in the level resonant with the conduction band decays to the band edge, where it becomes a shallow-donor electron. The energies of levels resonant with the host bands are merely schematic, are not to be taken as quantitative, and are merely to illustrate that there are resonances in the bands.

Neither antisite defect can explain the observed n-type character of InN.

The theory for In_N does provide a simple and natural explanation for the optical absorption data of Tansley and Foley:[2] They find a deep level in n-type InN, with p-like character lying ~ 1 eV below the conduction-band edge, which they attribute to an In_N antisite defect. Our theory (Fig. 4) does seem to be in excellent agreement with their data.

The In-site vacancy Va_{In} (Fig. 4) produces s-like and p-like levels near the valence-band maximum, with the s-like level doubly occupied and the p-like neutral vacancy level containing three electrons and three holes. The theory, taken literally, places the p-like level in the gap, where it can trap both electrons and holes, and the s-like level in the valence band. (It is conceivable that the p-like level actually lies below the valence-band maximum, in which case the In vacancy would be a triple shallow acceptor, because the holes would bubble up to the valence-band edge.) Clearly, the In vacancy cannot account for the observed n-type character of InN either.

The N vacancy can (Fig. 4). (Tansley and Foley have also speculated that the N vacancy might be the defect responsible for the natural n-type character of InN.[21]) N produces an s-like level (containing two electrons) near the conduction-band edge and a p-like level (containing one electron) above the conduction-band edge. Since the p-like level is resonant, its electron is autoionized, decays to the conduction-band edge, and dopes InN n type (one electron per vacancy). It is also possible that the s-like deep states lie a bit higher than predicted (not in the gap) and are resonant with the conduction band, donating

their electrons to the conduction band. In this second case the N vacancy would be a triple donor. Thus we conclude that the N vacancy is most likely responsible for the *n*-type behavior of InN, and is most likely a single donor (but possibly donates three electrons to the conduction band). Moreover, the nitrogen vacancy, if it is a simple donor with its *s*-like deep level in the gap just below the conduction-band edge, provides a natural explanation of the 0.2-eV feature in the Tansley-Foley optical absorption data.[22] Thus we propose that the N vacancy both dopes InN *n* type and produces a deep level 0.2 eV below the conduction-band minimum which has been detected in the optical absorption.

Another defect possibly responsible for the *n*-type character of InN is oxygen on a N site, which is not a native defect, but is nevertheless likely to be present in significant concentrations (Fig. 5).

C. Dopants

1. p type

Since InN occurs *n* type naturally, the central question concerning doping is whether it can be doped *p* type. If, as we predict, InN is naturally *n* type because of N vacancies, then *p*-type material must be relatively free of these vacancies or contain a sufficiently large number of acceptors to compensate them.

The best candidate for an acceptor is a column-II impurity on an In site (Fig. 6). Such an impurity will be a shallow acceptor in the classic sense. There is a problem with ordinary acceptors in InN, however, because this large-band-gap semiconductor should have a moderately small dielectric constant, estimated to be $\epsilon = 8.3$,[22] and a rather large (calculated) valence-band effective mass, $m^* = 1.6$,[22] causing the acceptor's effective-mass-theory binding energy to be rather large, ~ 0.3 eV. Thus unless these crude estimates of m^*/ϵ^2 are too large by a factor of 3 or more, the shallow-impurity binding energy is large enough to inhibit thermal ionization of holes at room temperature. Ultimately, the fact that the shallow-acceptor binding energy is so large, not the purported difficulty of incorporating shallow acceptors, may be the reason InN cannot be fabricated sufficiently *p* type.

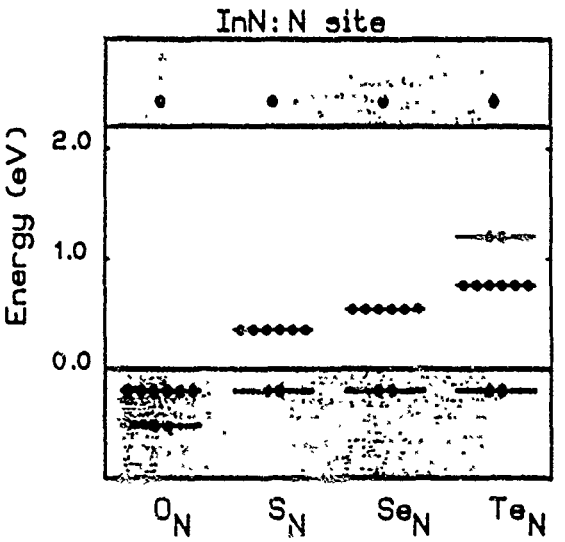


FIG. 5. Energy levels and occupancies of neutral column-VI impurities on the N site in InN. Electrons are denoted by solid circles. All column-VI impurities on the N site are predicted to be donors. The extra donor electron is denoted by a solid circle in the conduction band.

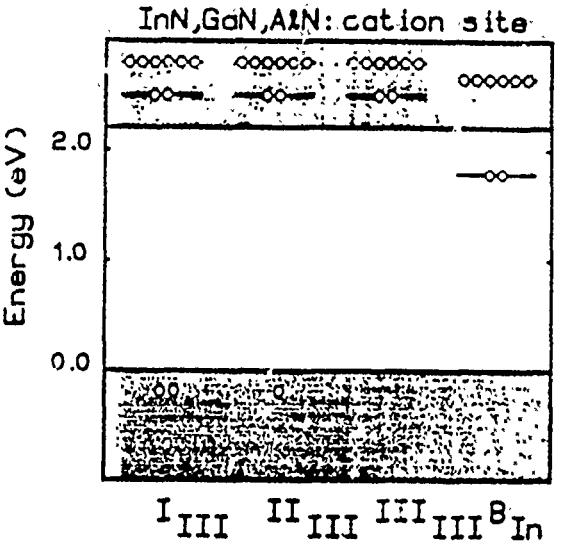


FIG. 6. Energy levels and occupancies of neutral impurities from columns I, II, and III on the cation site and B on the In site in InN, GaN, and AlN. Holes are denoted by open circles. All column-II impurities on the III site are predicted to yield shallow acceptors. Column-I impurities are predicted to yield double acceptors. Isoelectronic impurities on the III site are predicted to be inert, except for B_{In} in InN only, which is predicted to be a trap.

Column-IV impurities on the N site will very likely not produce shallow acceptors, but instead will produce both *s*-like (except perhaps for C) and *p*-like deep levels in the fundamental band gap—with one hole and five electrons in the upper (*p*-like) level and two electrons in the *s*-like level (Fig. 7)—except for Pb, which has its *s*-like and *p*-like levels reversed. Thus neutral column-IV impurities on N sites are deep traps for both electrons (one) and

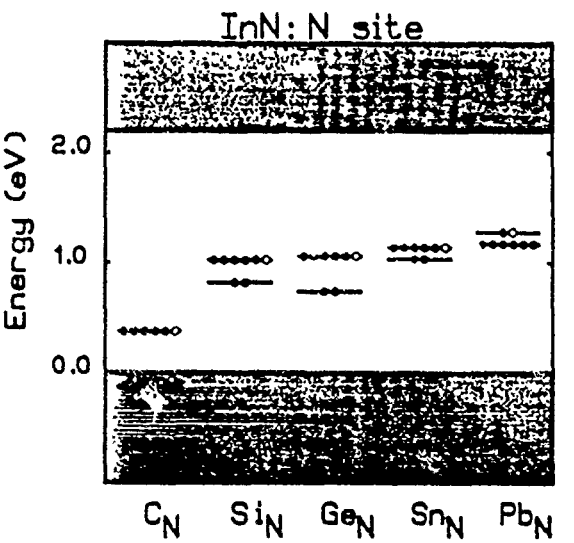


FIG. 7. Energy levels and occupancies of neutral column-IV impurities on the N site in InN.

holes (five).

We also note that the In vacancy, if its *p*-like level actually lies below the valence-band maximum, could dope InN *p* type (see Fig. 4).

2. *n*-type

For *n*-type doping, the best candidate, other than the N vacancy, is oxygen or some other chalcogen on a N site. Oxygen produces no deep levels in the fundamental band gap, and so is a classic shallow donor (Fig. 5). Neutral S, Se, and Te, in addition to producing the shallow donor, each also yield a *p*-like fully occupied deep level in the gap, which is driven up from the valence band because the *s* atomic-orbital energies of S, Se, and Te are higher than that of N.[23]

Column-IV impurities on the In site are not good candidates for donors, since they are predicted to have *s*-like deep levels in the gap, and so, when neutral, could either trap electrons or holes (Fig. 8). The predicted *s*-like levels for Pb and possibly Sn are close enough to the conduction-band edge that, allowing for a small uncertainty in the theory, these levels could lie resonant with the conduction band, and so could conceivably lead to shallow-donor behavior.

Thus we predict that the best dopants for InN are column-II impurities on the In site for *p*-type doping, and either a vacancy or oxygen on the N site for *n*-type doping. Column-IV dopants on either or both sites will tend to produce semi-insulating material.

D. Isoelectronic impurities

Isoelectronic impurities, namely impurities from the same column of the Periodic Table as the host atom they replace, are normally thought of as electronically inert. Rather spectacular counterexamples to this thinking are the N isoelectronic traps in $GaAs_xP_{1-x}$ and $Al_{1-x}Ga_xAs$ alloys, electron traps which play major roles in localizing electrons and enhancing the intensity of recombination radiation.

Isoelectronic impurities (unlike heteroelectronic donors and acceptors) can often trap one carrier without repelling a carrier of opposite sign, as in the case of N_P in GaP, which traps an electron—and the electron is subsequently able to capture a hole and to form an impurity-bound exciton.

In InN, according to the theory, neutral B on an In site produces such an isoelectronic trap, an *s*-like level slightly below the conduction-band edge (Fig. 6). (The *p*-like level of B is predicted to be resonant with the conduction band.) The remaining In-site isoelectronic traps are electronically inert, according to the theory, with their deep levels all being resonant.

On the N site, the Bi and Sb isoelectronic impurities produce both *s*-like and *p*-like deep levels in the band gap (fully occupied by electrons for the neutral defect) and hence are deep hole traps (Fig. 9). Similarly, As and P on the N site have (full) *p*-like deep levels in the gap, while their (full) *s*-like levels lie just below the valence-band maximum. They too are deep hole traps.

One of the interesting features of these isoelectronic traps is that they bind one carrier in a localized orbital, and so can bind an exciton by binding one carrier which binds the second through the electron-hole interaction. For example, B_{In} can bind an electron which, in turn, can bind a hole. Similarly, Bi_N, Sb_N, and As_N, or P_N, can bind a hole which can attract an electron. By localizing an exciton this way, an isoelectronic trap can enhance the intensity of the recombination luminescence, because the recombination rate for a localized state is generally much larger than for a delocalized state.

E. Other deep levels

Impurities two or more columns of the Periodic Table distant from the host atom they replace tend to be rather insoluble; nevertheless, their solubilities are not zero, and we include their predicted deep levels here for completeness.

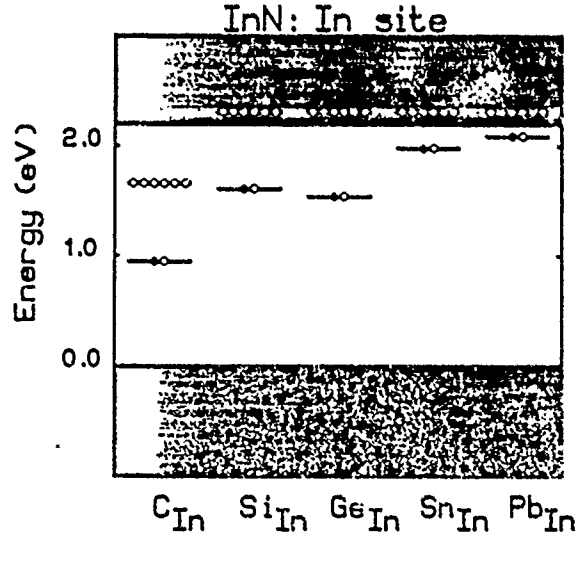


FIG. 8. Energy levels and occupancies of neutral column-IV impurities on the In site in InN.

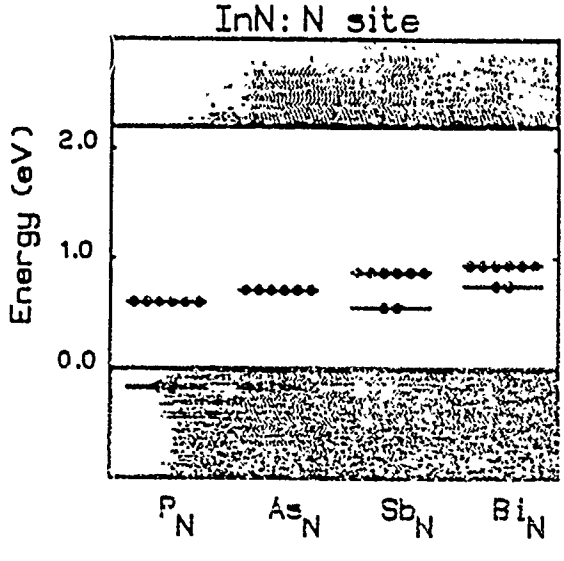


FIG. 9. Energy levels and occupancies of neutral isoelectronic impurities on the N site in InN.

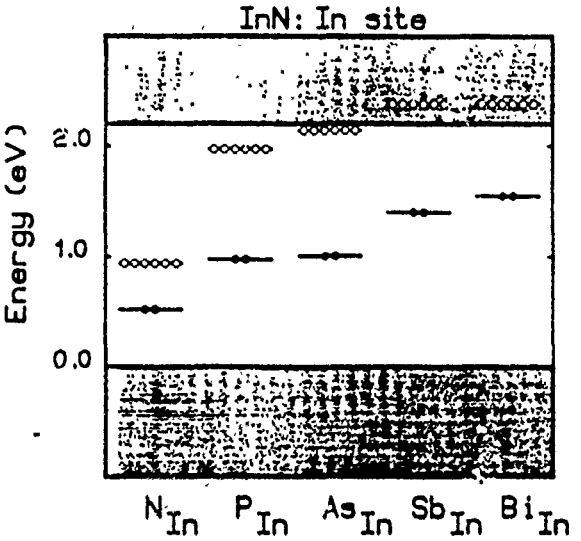


FIG. 10. Energy levels and occupancies of neutral column-V impurities on the In site in InN.

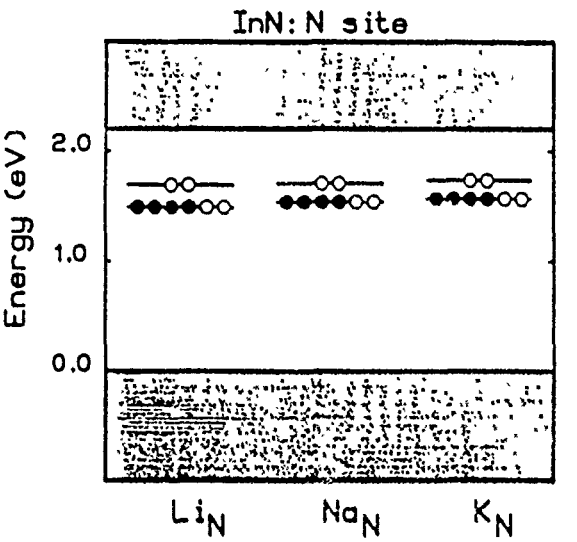


FIG. 13. Energy levels and occupancies of neutral column-I impurities on the N site in InN.

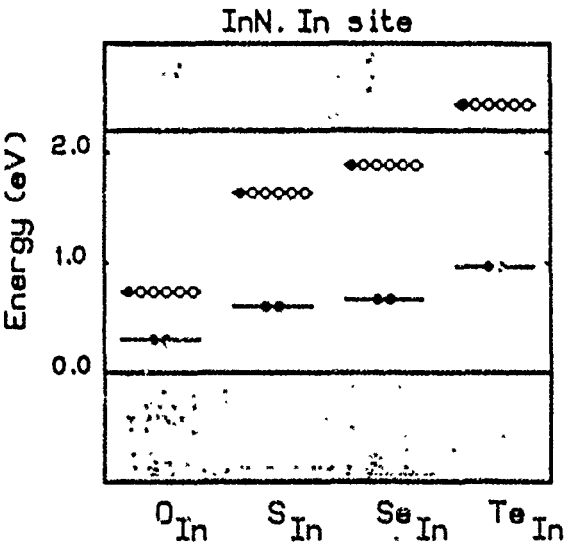


FIG. 11 Energy levels and occupancies of neutral column-VI impurities on the In site in InN.

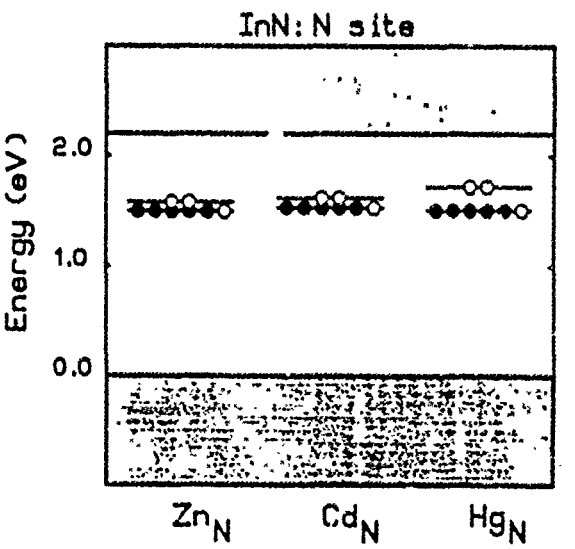


FIG. 14. Energy levels and occupancies of neutral column-II impurities on the N site in InN.

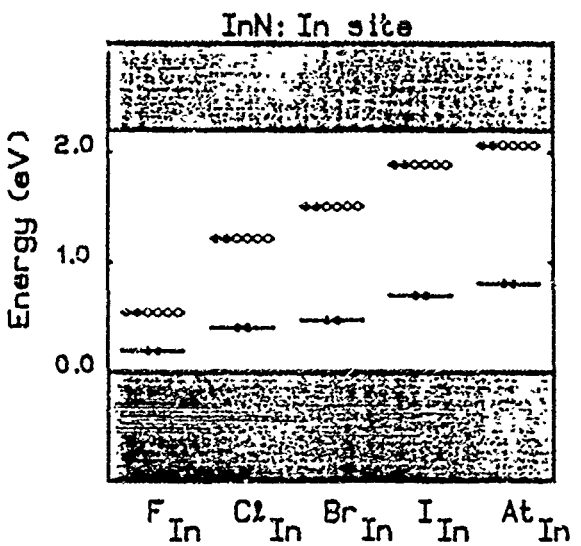


FIG 12. Energy levels and occupancies of neutral column-VII impurities on the In site in InN.

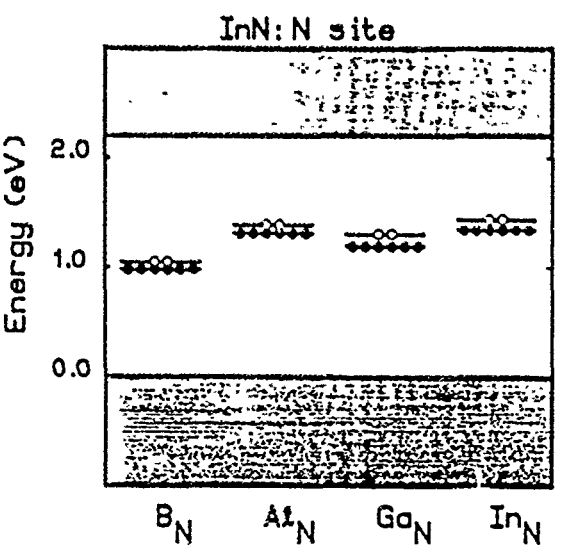


FIG. 15. Energy levels and occupancies of neutral column-III impurities on the N site in InN.

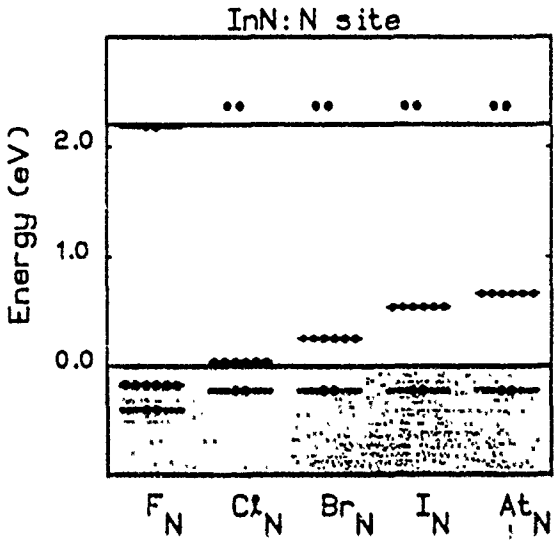


FIG. 16. Energy levels and occupancies of neutral column-VII impurities on the N site in InN.

On the In site column-I and -II impurities are double acceptors and single acceptors, respectively, with their s and p deep levels in the conduction band (see Fig. 6). The column-V impurities produce deep traps: doubly occupied deep s levels in the gap and empty p levels that are in the gap for P_{In} and As_{In} but in the conduction band for Sb_{In} and Bi_{In} (Fig. 10). The column-VI impurities (Fig. 11) produce doubly occupied s levels deep in the gap plus singly occupied p levels that are deep in the gap, except for Te (which should produce a shallow donor!). The halogens on the In site should produce deep s and p levels in the gap (Fig. 12).

On the N site, columns-I, -II, and -III impurities all produce s-like and p-like deep levels in the gap of InN, all at about the same energy (Figs. 13–15). Column-VII impurities (Fig. 16) are all double donors, except possibly for F, which the theory predicts to be inert (which, within the theoretical uncertainty, may also be a double donor).

IV. DEEP LEVELS IN $In_xGa_{1-x}N$ AND $In_xAl_{1-x}N$

A. Doping anomalies

The deep levels in the alloys $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$ are similar to those in InN. As functions of alloy composition they vary in energy rather smoothly. In many cases deep levels that lie in the fundamental band gap for InN move out of the gap as a function of alloy composition and lie resonant with the host bands of GaN or AlN. When this happens, the character of the impurity changes (e.g., from a deep trap to a shallow donor) and a "doping anomaly" occurs.

There are two common types of doping anomaly: (i) false valences, and (ii) deep-shallow transitions.

False valences occur when, as a function of alloy composition, a deep level completely crosses the fundamental band gap. False valences do not occur in $In_xGa_{1-x}N$ or $In_xAl_{1-x}N$, but, to understand the concept of a false valence, suppose that the s-like level of Si on a cation site (Fig. 8) were to descend from the conduction band of GaN through the gap to the valence band of InN (it does not). Then the hole in the deep level would bubble up to the valence-band maximum, and neutral Si would become a single acceptor rather than a deep trap for both electrons and holes, and would have a false valence of -1 instead of its normal valence $(+1)$ with respect to the column-III cation. The reason false valences do not occur in $In_xGa_{1-x}N$ or $In_xAl_{1-x}N$ is that the vacancies have both s-like and p-like levels in or very near the fundamental band gap (Fig. 4). Since the vacancies correspond to infinite defect potentials,[24] they separate the impurity levels that originate from the conduction band with finite defect potentials from those that come from the valence band—and, if the vacancy levels lie in the gap for all alloy compositions, prevent impurity levels from crossing the gap.

Deep-shallow transitions occur when a deep level in the gap moves out of the gap (as a function of x). Consider, as an example, a column-IV impurity such as Si on a cation site in $In_xGa_{1-x}N$. In InN, neutral Si on an In site produces an s-like deep level in the gap occupied by one electron and one hole (Fig. 8). Therefore neutral Si_{In} is a deep electron and hole trap in InN, but in GaN or AlN, Si on a cation site produces an s-like level degenerate with the conduction band (Figs. 17 and 18). The electron that occupied this level in InN is autoionized in GaN or AlN and falls to the conduction-band edge (where the long-ranged Coulomb potential omitted in this paper traps the electron in a shallow-donor state). As a result cation-site Si in GaN or AlN is a shallow-donor impurity: its ground state has the extra electron in a shallow level (whereas in InN this extra electron occupies a deep level). For some intermediate alloy composition between InN and GaN, the deep level of Si_{In} in InN passes through the conduction-band edge, and the Si impurity changes its character from a deep trap to a shallow donor (Fig. 19).

The predicted dependences on alloy composition x of substitutional deep levels in $In_xGa_{1-x}N$ are displayed in Figs. 19–22.

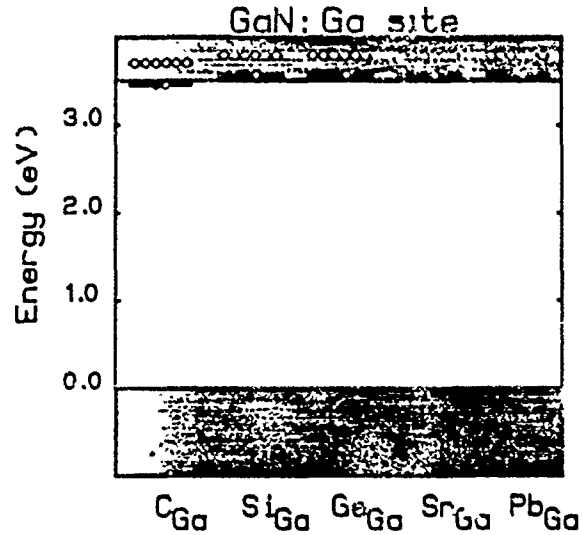


FIG. 17. Energy levels and electron occupancies of neutral column-IV impurities on the Ga site in GaN.

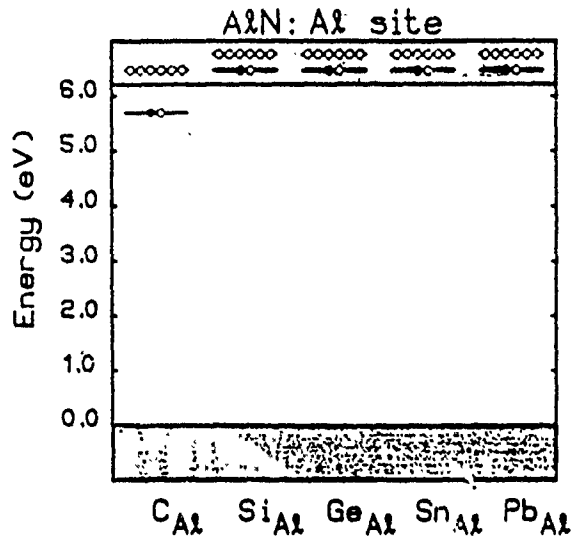


FIG. 18. Energy levels and electron occupancies of neutral column-IV impurities on the Al site in AlN.

The principal deep levels for substitutional defects in GaN and AlN are given in Figs. 17 and 18 and 23–30. The variation with alloy composition x of deep levels in $In_xAl_{1-x}N$ is given in Figs. 31–34. In many cases, the qualitative level structure for a specific impurity on a given site is the same for all alloy compositions of $In_xGa_{1-x}N$ and $In_xAl_{1-x}N$ (e.g., for oxygen on an anion site), and no deep levels cross either the valence-band or conduction-band edge as alloy composition is varied. In such cases, the qualitative doping character of the defect does not change, although its quantitative energy levels do, according to Figs. 19–22 and 31–34. Here we focus on those defects whose qualitative characters do change with alloy composition, defects that undergo deep-shallow transitions.

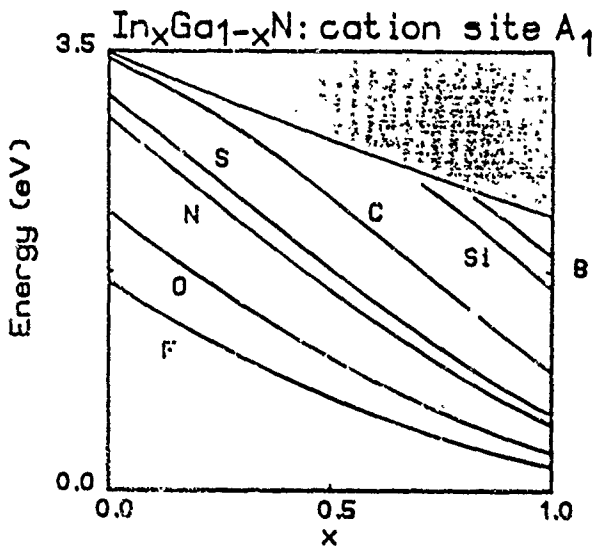


FIG. 19. A_1 (s-like) defect levels vs alloy composition x for impurities on the cation site in $In_{1-x}Ga_xN$. Levels for impurities from columns V, VI, and VII are completely filled, levels for impurities from column IV have one electron and one hole, and levels for isoelectronic impurities are unoccupied by electrons. Other impurities are acceptors.

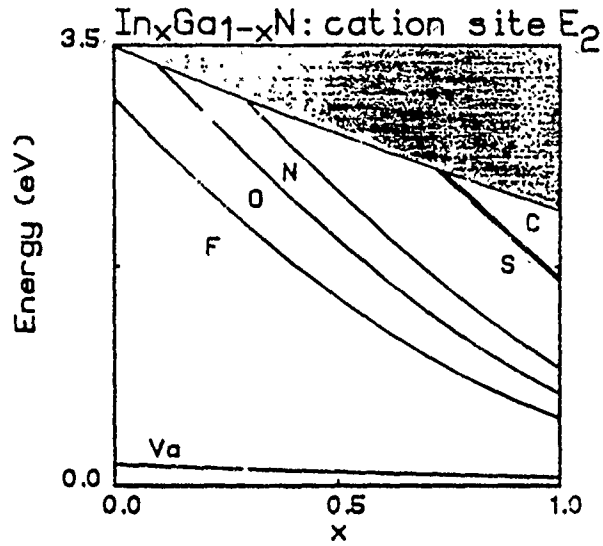


FIG. 20. E_2 (p-like) defect levels for impurities on the cation site in $In_{1-x}Ga_xN$. Only levels for impurities from columns VI and VII are partially filled by electrons; other impurity levels are unoccupied.

B. Native defects

The antisite defects, N on a cation site and a cation on a N site, have the same qualitative level structures in GaN and AlN, but a different one in InN (Figs. 4, 23, and 27). In InN the (neutral) N_{In} defect has both a filled (doubly occupied) s-like level and an empty p-like deep level in the gap. The p-like level is in the conduction band for GaN and AlN. Thus N_{In} is a deep trap for both electrons

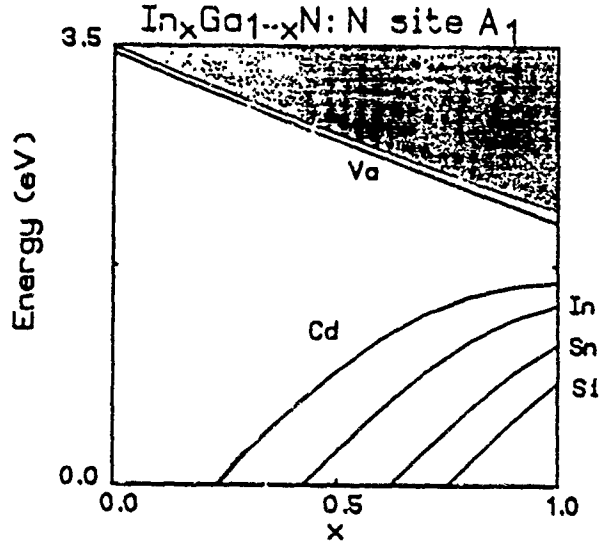


FIG. 21. A_1 (s-like) defect levels for impurities on the N site in $In_{1-x}Ga_xN$. All the levels shown are occupied by two electrons.

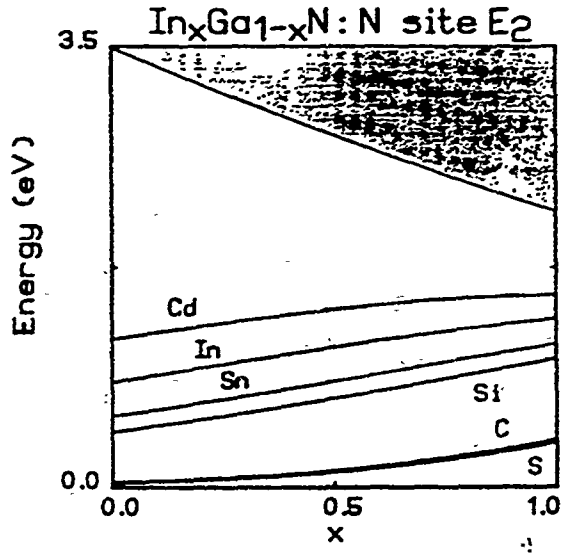


FIG. 22. E_2 (p-like) defect levels for impurities on the N site in $In_{1-x}Ga_xN$. Levels due to impurities which have fewer valence electrons than N, such as Si, are partially occupied by electrons and trap both electrons and holes. Impurities which have more valence electrons than N, such as O, are donors: their levels are completely occupied by electrons and have extra electron in the conduction band.

and holes, whereas neutral N_{Ga} in GaN and N_{Al} are deep-hole traps.

The cation-vacancy p-like level is barely in the gap for InN and GaN, and deep in the gap for AlN (Figs. 4, 23, and 27), whereas the s-like level is resonant with the valence band for InN and GaN, but in the gap for AlN. This vacancy, when neutral, can trap either electrons or holes.

The N vacancy is a shallow donor in InN and GaN (Figs. 4 and 23). with its p-like level in the conduction band and its s-like deep level doubly occupied in the gap. In AlN the p-like level lies in the gap (Fig. 27), making the neutral N vacancy a deep electron trap.

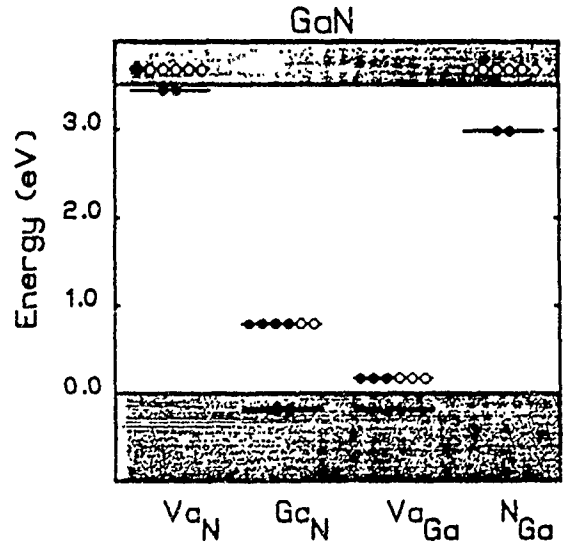


FIG. 23. Energy levels and electron occupancies of neutral native defects in GaN.

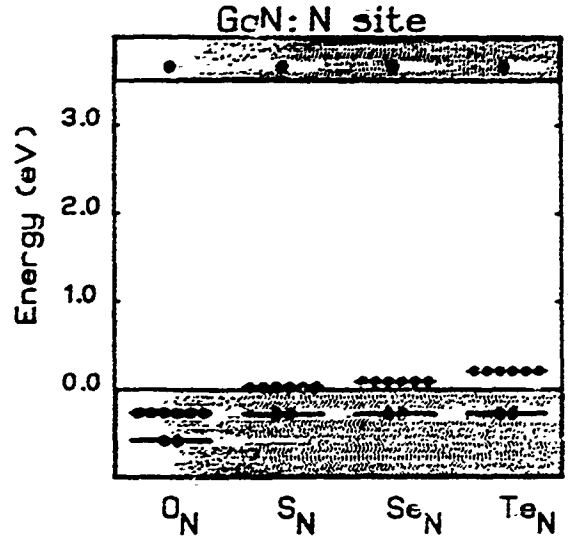


FIG. 24. Energy levels and electron occupancies of neutral column-VI impurities on the N site in GaN.

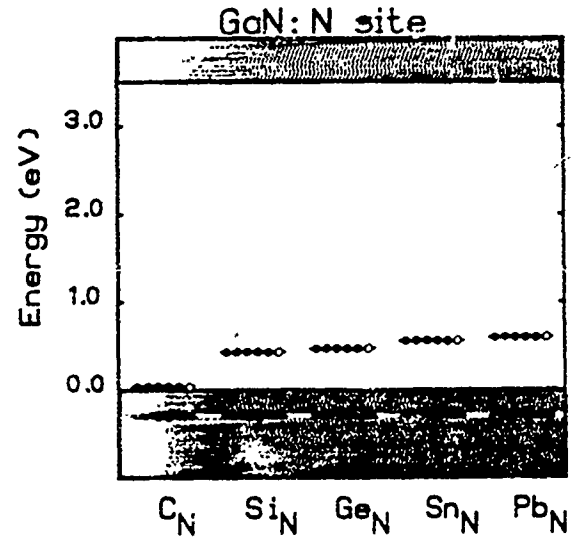


FIG. 25. Energy levels and electron occupancies of neutral column-IV impurities on the N site in GaN.

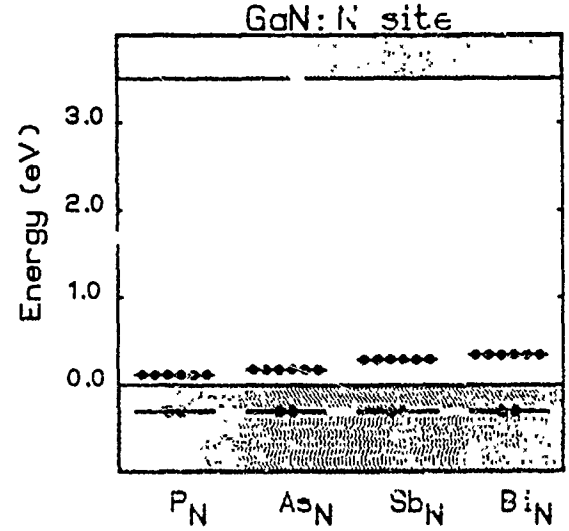


FIG. 26. Energy levels and electron occupancies of neutral isoelectronic impurities on the N site in GaN.

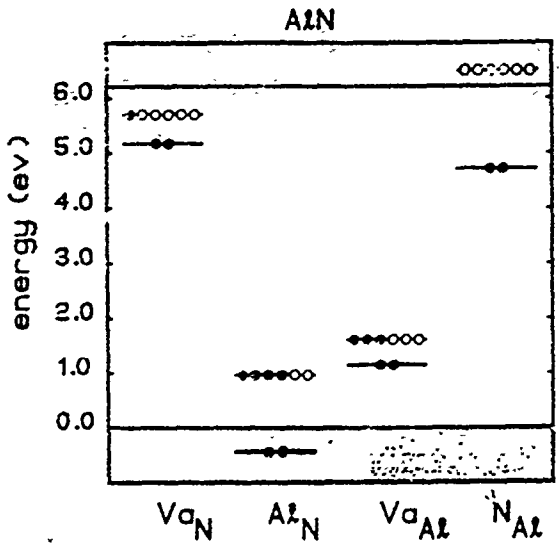


FIG. 27. Energy levels and electron occupancies of neutral native defects in AlN.

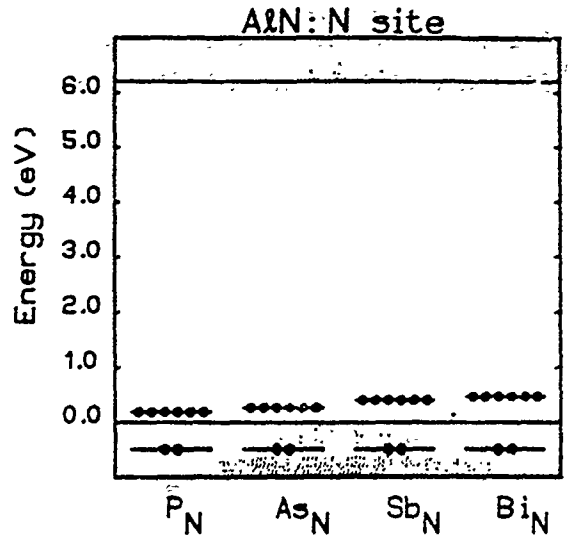


FIG. 30. Energy levels and electron occupancies of neutral isoelectronic impurities on the N site in AlN.

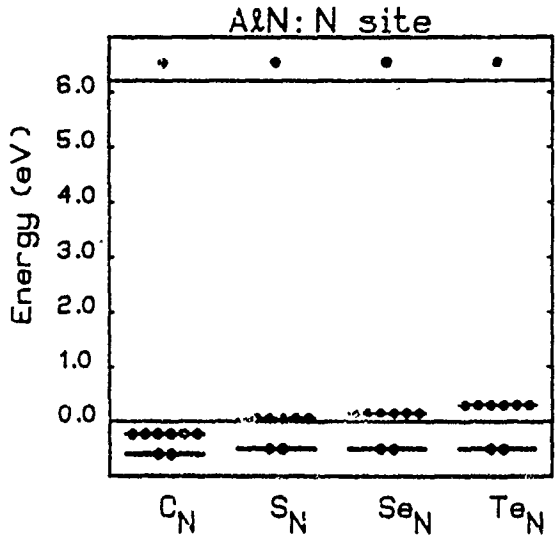


FIG. 28. Energy levels and electron occupancies of neutral column-VI impurities on the N site in AlN.

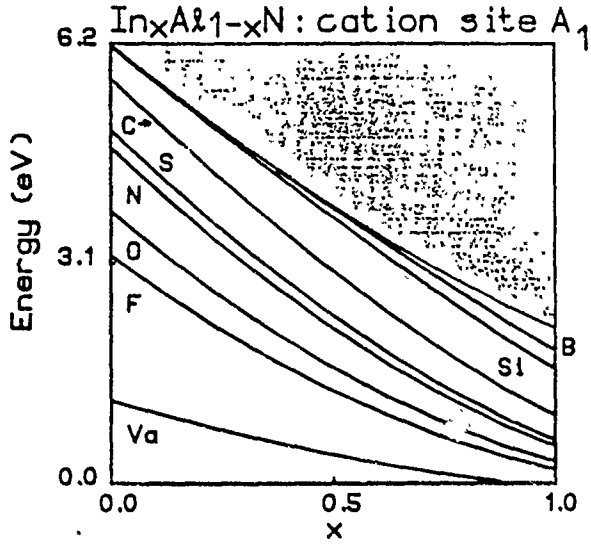


FIG. 31. A_1 (s-like) defect levels for impurities on the cation site in $In_{1-x}Al_xN$.

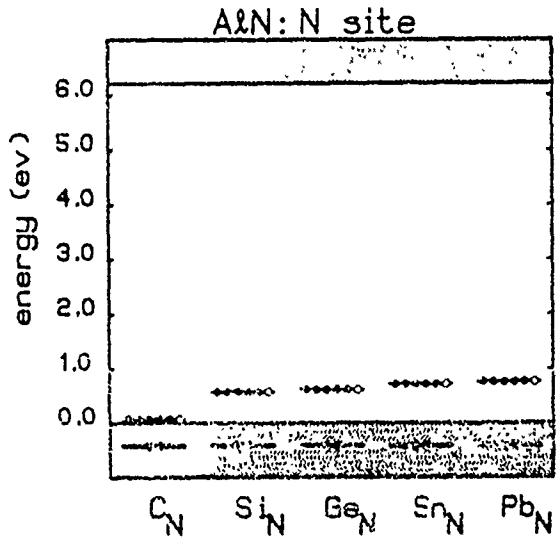


FIG. 29. Energy levels and electron occupancies of neutral column-IV impurities on the N site in AlN.

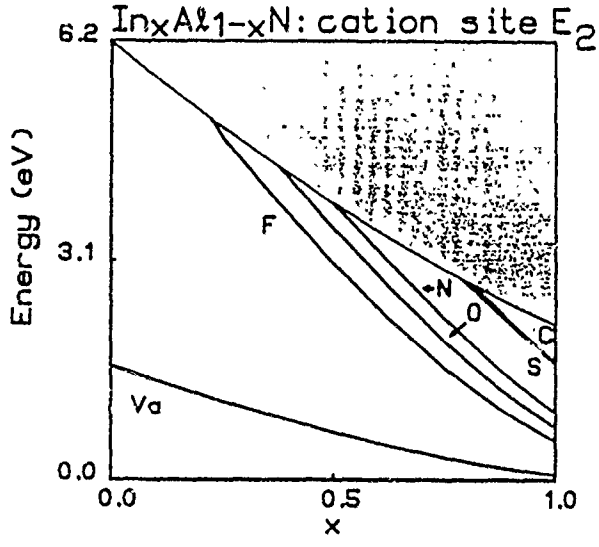


FIG. 32. E_2 (p-like) defect levels for impurities on the cation site in $In_{1-x}Al_xN$.

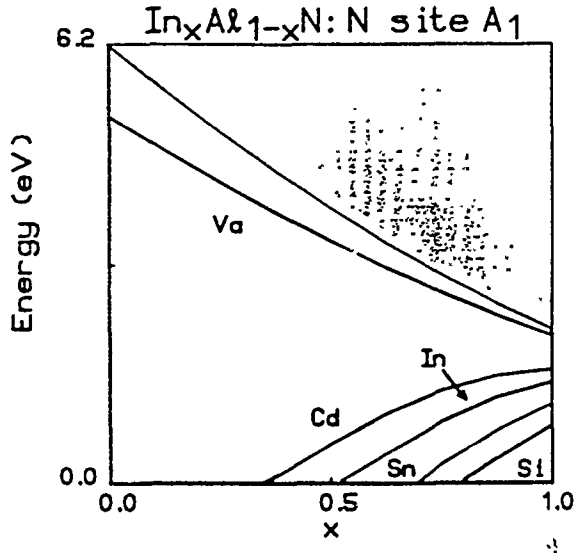


FIG. 33. A_1 (s-like) defect levels for impurities on the N site in $In_{1-x}Al_xN$.

C. Donors and acceptors

Column-II impurities on cation sites are ordinary acceptors in InN, GaN, and AlN (Fig. 6). Column-I impurities are double acceptors.

Column-VI impurities on anion sites are ordinary donors, but S, Se, and Te each have a fully occupied deep p-like level deep in the gap for InN, and barely in the gap for GaN and AlN. (This level is in the valence band for oxygen.) See Figs. 5, 24, and 28.

Column-IV impurities on a N site each produce a deep electron and hole trap due to a p-like deep level in the gap (occupied by one hole and five electrons for the neutral defect) in InN, GaN, and AlN (Figs. 7, 25, and 29). The filled s-like deep level is also in the gap for InN (except for C), but not for GaN or AlN (See Figs. 7, 25, and 29).

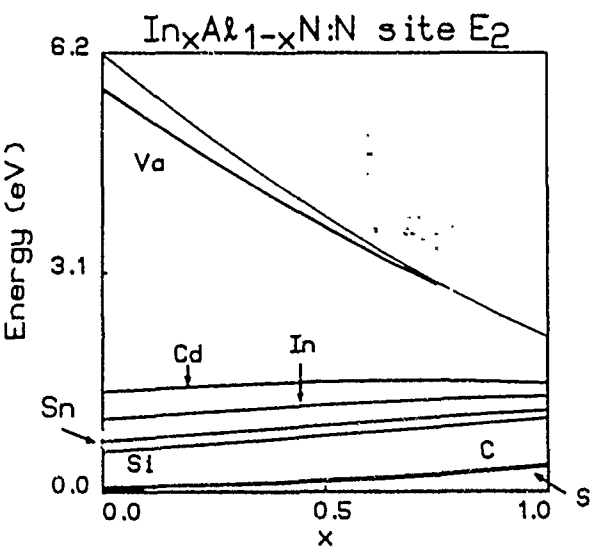


FIG. 34. E_2 (p-like) defect levels for impurities on the N site in $In_{1-x}Al_xN$.

Column-IV impurities (except C) on the Ga site in GaN and the Al site in AlN produce shallow donors (Figs. 17 and 18). In InN (Fig. 8) they produce s-like deep levels in the gap. Indeed, the p-like deep level of C_{III} even lies well in the gap of InN, while the s-like levels of neutral carbon are predicted to be deep electron and hole traps for alloys of InN, GaN, and AlN.

D. Isoelectronic defects

The isoelectronic defect B on a cation site produces an s-like deep level in the gap of InN, but this level is in the conduction band of GaN and AlN (Fig. 6). Thus with decreasing x in $In_xGa_{1-x}N$ or $Al_xGa_{1-x}N$, B undergoes a deep-inert transition (Figs. 19 and 31). (The isoelectronic impurity has no long-ranged Coulomb potential and hence no shallow levels; thus it becomes inert rather than shallow when its deep levels are all resonant with host bands.) The other column-III isoelectronic defects are inert in InN, GaN, and AlN.

Column-V impurities on the N site (except N) all produce occupied p-like levels in the gap and are hole traps in InN, GaN, and AlN (Figs. 9, 26, and 30). In addition, Bi and Sb in InN have occupied s-like levels in the gap for InN. (These levels descend into the valence bands of GaN and AlN.)

V. SUMMARY

We have predicted the electronic structures of InN, $In_xGa_{1-x}N$, and $In_xAl_{1-x}N$, and find that these materials exhibit direct band gaps ranging from orange to ultraviolet. We find that the N vacancy, not the antisite defect N_{In}, is primarily responsible for InN's n-type character as grown. We propose that the nitrogen vacancy is also responsible for the 0.2-eV absorption feature and we confirm the Tansley-Foley suggestion that In_N is responsible for the absorption attributed to the midgap defect level. We predict that column-II impurities on cation sites should produce p-type behavior, while column-IV impurities should yield semi-insulating properties. B on an In site in InN should produce an isoelectronic electron trap, while column-V impurities on the N site should yield hole traps. However, the shallow-acceptor binding energy may be too large to permit thermal ionization of large numbers of holes, and it may be difficult to prepare these materials with high concentrations of positive carriers. Various deep-shallow and deep-inert transitions occur in the $In_xGa_{1-x}N$ alloy system. We conclude that $In_{0.35}Ga_{0.65}N$ and $In_{0.85}Al_{0.15}N$, if these materials can be successfully grown, should produce blue-green luminescence, and should be dopable, both n type and p type—although the shallow-acceptor binding energy may be so large as to limit the number of holes in the valence band.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research (Contract No. N00014-84-K-0352) and the U.S. Air Force Office of Scientific Research (Contract No. AFOSR-85-0331) for their generous support, and we thank K. Hathaway, M. Yoder, B. Goldenberg, C. S. Lent, and S.-Y. Ren for stimulating conversations.

*Present address: Amoco Research Center, P.O. Box 400 F-4, Naperville, IL 60540.

[1]T. L. Tansley and C. P. Foley, in *Proceedings of the 3rd International Conference on Semi-Insulating III-V Materials*, Warm Springs, Oregon, 1984, edited by J. Blakemore (Shiva, London, 1985), pp. 497–500.

[2]See also C. P. Foley and T. L. Tansley, Appl. Surf. Lett. 22, 663 (1985).

[3]A. Kobayashi, O. F. Sankey, S. M. Volz, and J. D. Dow, Phys. Rev B 28, 935 (1983), and references therein.

[4]G. Mendel, Phys. Rev. 134, A1073 (1964).

[5]See, for example, W. A. Harrison, *Solid State Physics* (Dover, New York, 1971), p. 324.

[6]See, for example, J. I. Pankove, *Optical Properties of Semiconductors* (Dover, New York, 1979), Chaps. 1 and 3.

[7]J. D. Dow and R. E. Allen, Appl. Phys. Lett. 41, 672 (1982).

[8]J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).

[9]A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 28, 946 (1983).

[10]C. S. Lent, D. W. Jenkins, and J. D. Dow (unpublished).

[11]D. W. Jenkins, R.-D. Hong, and J. D. Dow, Superlatt. Microstruct. 3, 365 (1987).

[12]The virtual-crystal approximation can be given a modicum of justification because the Onodera-Toyozawa ratios for $In_xGa_{1-x}N$ [Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1968)], namely the ratios of In-Ga diagonal on-site matrix-element differences to valence- (conduction-) band width are 0.22 (0.98) and 0.04 (0.08) for s and p orbitals, respectively. These values place $In_xGa_{1-x}As$ in the "amalgamated" limit for the valence band and marginally for the conduction band. The values are somewhat higher than the corresponding values for $Al_{1-x}Ga_xAs$, 0.07 (0.16) and 0.005 (0.008), a material for which the virtual-crystal approximation is known to be a good model for the electronic structure; however, the theory agrees with the band-gap data of Osamura *et al.* [K. Osamura, S. Naka, and Y. Murakami, J. Appl. Phys. 46, 3432 (1975)] to within 0.3 eV. (These data were taken at 77 K, while our theory describes the material at 4 K.) Thus we conclude that this approximation will provide an adequate "zeroth-order" description of all the InN alloys. Similar values for $In_xAl_{1-x}N$ are 0.94 (0.10) and) 0.40 (1.66) for the valence (conduction) band for s and p states, respectively. These parameters vary with composition and are in the "amalgamated" regime for Al-rich alloys.

[13]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980), p. 47.

[14]L. Vegard, Z. Phys. 5, 17 (1921).

[15]The conduction-band minima occur at the symmetry points in the Brillouin zone: $\Gamma=(0,0,0)$, $H=(2\pi/a)(1/\sqrt{3},\frac{1}{3},c/2a)$, $L=(2\pi/a)(1/\sqrt{3},0,c/2a)$, $K=(2\pi/a)(1/\sqrt{3},\frac{1}{3},0)$, $M=(2\pi/a)(1/\sqrt{3},0,0)$, and $A=(2\pi/a)(0,0,c/2a)$.

[16]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980). See also W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[17]A substitutional defect in a wurtzite material has two spin-degenerate s-like (A_1) states, two spin-degenerate p_z-like states (A_1), and four spin-degenerate p_x-like and p_y-like states (E_2). We approximate the ratio between the repeat distance in the z axis c and the lattice constant in the x-y plane a to be $c/a=(\frac{8}{3})^{1/2}$. In this approximation the small splitting between the p-like A_1 and E_2 states vanishes.

[18]The Green's functions are defined by

$$G_l(E)=\sum_{n,\mathbf{k}}\int d^3k\frac{|\langle l',\mathbf{k},n|l',\mathbf{k},n\rangle|^2}{E-E_n(\mathbf{k})},$$

where $E_n(\mathbf{k})$ is the nth eigenvalue at wavevector $\mathbf{k}$ and $|l,\mathbf{k},n\rangle$ is the nth eigenvector at $\mathbf{k}$ with $l=s$ or p for the A_1 or E_2 symmetry, respectively.

[19]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[20]S. Lee, J. D. Dow, and O. F. Sankey, Phys. Rev. B 31, 3910 (1985).

[21]T. L. Tansley and C. P. Foley, J. Appl. Phys. 60, 2092 (1986).

[22]M.-H. Tsai, D. W. Jenkins, J. D. Dow, and R. V. Kasowski, Phys. Rev. B 38, 1541 (1988).

[23]K. E. Newman and J. D. Dow, Phys. Rev. B 30, 1929 (1984).

[24]M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969).

Hydrostatic-pressure dependencies of deep impurity levels in zinc-blende semiconductors

Run-Di Hong, David W. Jenkins,* Shang Yuan Ren,† and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 23 December 1987)

The derivatives with respect to hydrostatic pressure are predicted for deep levels associated with *s*- and *p*-bonded impurities in Si, Ge, AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb, and ZnSe. It is shown that by combining data, both for deep levels and for their hydrostatic-pressure derivatives, with theory it is often possible to determine (i) the site of the impurity, (ii) the symmetry of the deep level, and (iii) a quite small number of substitutional *s*- and *p*-bonded impurities that could be responsible for the data. We use this method to argue that the deep levels observed by Aladashvili *et al.* in InSb to lie in the interval between 0.1 and 0.15 eV above the valence-band maximum are probably A_1-symmetric levels associated with C_{In} and/or antisite Sb_{In} (or levels associated with defect complexes involving these defects).

I. INTRODUCTION

Several years ago, deep levels were defined as impurity states in semiconductors whose energies were more than 0.1 eV from a nearby band edge—namely levels that were not thermally ionized at room temperature. More recently this definition has been revised as a result of the recognition that deep levels can, when perturbed, cease being energetically deep in the gap and can actually pass into a band where they become resonances. The current definition of a deep level is one that is caused by the central-cell potential of the defect.[1] In fact, all *s*- and *p*-bonded substitutional impurities in zinc-blende semiconductors produce typically four such deep levels in the vicinity of the fundamental band gap: one *s*-like (A_1 symmetric) and one triply degenerate *p*-like (T_2) level. These deep levels, more often than not, are resonances that lie outside the band gap, and hence are not "deep" by the old definition.

Despite the fact that various theories of deep levels have been developed,[2-4] beginning with the classical paper by Lannoo and Lenglart on the levels associated with the Si vacancy,[5] the theories generally have not been capable of identifying a particular impurity from the energies of its observed deep levels in the fundamental band gap. This is due only in part to the fact that the best theories of deep levels have theoretical uncertainties of a few tenths of an eV for their level predictions.

Ren *et al.*[6] following the "deep-level pinning" ideas of Hjalmarson *et al.*,[2] showed for substitutional *s*- and *p*-bonded point defects on a site that all deep levels in the band gap with a particular symmetry have almost the same wave function—independent of the defect (see Fig. 1). This notion was confirmed experimentally by various electron-nucleon double resonance (ENDOR) measurements of deep-level wave functions in semiconductors.[7,8] Thus, no experiment that probes only the valence electronic properties of a deep level is capable of easily identifying the impurity responsible for the level: it is necessary to probe the nucleus [e.g., with ENDOR (Ref. 7)] or the core [e.g., with extended x-ray-absorption fine structure (EXAFS) (Ref. 9)] to achieve such a unique identification.

The situation has been further complicated due to the demonstration by Sankey *et al.*[10] that extended substitutional defects often have almost the same deep-level energies as their constituent isolated impurities. Thus analyses of observed energy levels are unlikely to reveal even if the defect producing the level is a point defect. (This complication is also a simplification, because it means that theoretically one need consider only isolated defects, since defect complexes have, to a good approximation, spectra which are the sums of their constituents' spectra.)

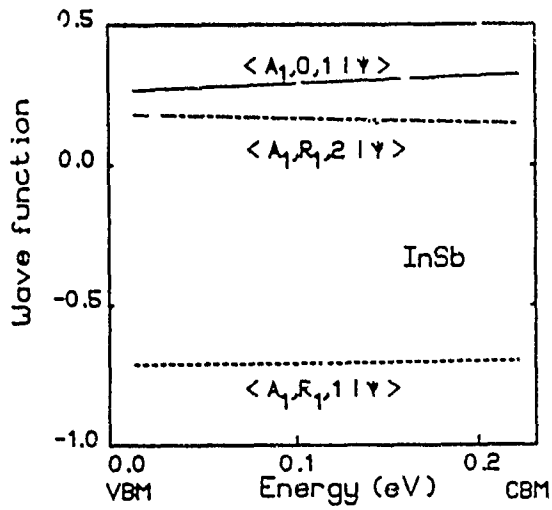


FIG. 1. Indium-site substitutional defect wave functions in InSb are shown as functions of deep energy level E. The on-site wave function $\langle A_1,0,1 \mid \Psi \rangle$ in the notation of Ref. 6 and first-shell wave functions $\langle A_1,R_{1,1} \mid \Psi \rangle$ [inward-directed hybrids (Ref. 6)] and $\langle A_1,R_{1,2} \mid \Psi \rangle$ [outward-directed hybrids (Ref. 6)] are shown by solid, dashed, and dashed-dotted lines, respectively. All these curves are nearly flat and show that the deep defect wave functions depend very little on their energy levels.

 ©1988 The American Physical Society

This insensitivity of deep impurity levels to the impurity raises the question of how much information can be gleaned from simple electronic measurements of deep levels, such as measurements of their energies and pressure derivatives.[11-13] In this paper, we show that (if we limit ourselves to s- and p-bonded substitutional impurities in zinc-blende semiconductors) combined measurements of a deep level and its change with hydrostatic pressure can usually determine both the symmetry of the level and the site (anion or cation site) of its parent impurity. Furthermore, we also show that the number of candidates for producing a particular deep level in the band gap can be reduced in number to only a few—and that this can be done even for impurities in a small-band-gap semiconductor such as InSb, whose (low-temperature) band gap of 0.23 eV is smaller than the uncertainty in most theories.

II. THEORY

Our theoretical approach is based on the model of electronic structure of Vogl *et al.*,[14] the theory of deep levels of Hjalmarson *et al.*,[2] and the work of Ren *et al.*[11] on pressure effects. We note that Ren's basic approach to pressure effects on deep levels in GaAs led to the targeting several years ago of oxygen and the antisite defect as possible constituents of the defect $EL2$— and that the role of the antisite defect is now generally acknowledged, while some (but not all) authors continue to believe that oxygen is also a constituent of $EL2$. Thus the basic theoretical approach of Ren *et al.* has a history of success, and we use that approach here for deep levels in other zinc-blende semiconductors.

The deep levels E are obtained by solving the secular determinant

$$\det[1-G_0(E)V]=0 .$$

Here, $G_0(E)$ is the Green's function, which is real in the fundamental energy band gap

$$G_0=(E-H_0)^{-1} .$$

The host Hamiltonian H_0 is taken to be the nearest-neighbor empirical tight-binding model of Vogl *et al.*[14] which is a ten-band model capable of describing both the chemistry of sp^3 bonding and the indirect-gap energy band structure of semiconductors such as Si and GaP, by virtue of its five-orbital sp^3s^* basis centered on each site. Expressed formally in terms of Löwdin orbitals[15] $|n,b,\mathbf{R}\rangle$ centered at the atom in unit cell $\mathbf{R}$ at site b (b denotes anion or cation) the Bloch-like tight-binding basis states are

$$|n,b,\mathbf{k}\rangle=N^{-1/2}\sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}}|n,b,\mathbf{R}\rangle .$$

Here, n runs over s, s^*, and the three p states. In this basis, the host Hamiltonian is a 10×10 matrix for each wave vector $\mathbf{k}$ (see Refs. 14 and 16). By diagonalizing this matrix and obtaining its eigenvalues $E(\mathbf{k},\lambda)$, namely the band structure, and its eigenvectors $|\mathbf{k},\lambda\rangle$, namely the Bloch states, one can construct the Green's-function operator

$$G_0(E)=\sum_{\mathbf{k},\lambda}(|\mathbf{k},\lambda\rangle\langle\mathbf{k},\lambda|)/[E-E(\mathbf{k},\lambda)] .$$

The defect-potential matrix V is taken to be diagonal and centered solely on the impurity site in the Löwdin basis; this approximation is now well established, and corresponds to neglecting lattice relaxation around the impurity.[2] Coulombic charge-state splittings[17] are also neglected. Since any underlying theory for predicting the deep levels of a given impurity is only accurate to a few tenths of an eV, the omission of lattice relaxation and charge-state splittings does not appreciably increase the theoretical uncertainty. Following Hjalmarson *et al.*,[2] we approximate the diagonal matrix elements of V on the impurity site as

TABLE I. Exponents $\eta_{l,l'}$ ($l,l'=s,p,s^*$) for the bond-length dependencies of the nearest-neighbor matrix elements. Exponents η are obtained by fitting the observed pressure dependencies (Table II) of the direct band gaps at Γ, L, and X, and the indirect gaps from L and X, to the valence-band maximum, using the least-squares method.

	$\eta_{s,s}$	$\eta_{s,p}$	η_{s,s^*}	$\eta_{s^*,s}$	$\eta_{s^*,p}$
Si	3.000	1.600	3.825	2.600	3.327
Ge	4.400	2.400	2.300	2.500	3.982
AlP	2.386	1.637	1.521	1.247	2.486
AlAs	3.205	1.656	2.398	1.706	3.214
AlSb	2.553	4.249	1.192	3.272	4.469
GaP	3.697	2.804	1.630	2.795	2.841
GaAs	4.144	2.341	2.220	2.596	2.665
GaSb	4.044	2.013	1.634	2.281	1.245
InP	3.100	4.443	3.049	2.366	1.207
InAs	2.539	2.812	3.757	2.825	3.014
InSb	4.012	2.987	2.533	2.751	3.134
ZnSe	1.874	1.185	1.838	1.330	3.185

TABLE II. Pressure coefficients of zinc-blende semiconductors (in meV/kbar).

Semiconductor	dE_Γ/dp	dE_L/dp	dE_X/dp	$dE_{L\Gamma}/dp$	$dE_{X\Gamma}/dp$
Si	1.0	6.2	3.0	5.5[a]	−1.5
Ge	14.2	7.8	5.5	5.0	−1.5
AlP	11.8[b]	5.21[b]	−0.62[b]	4.36[b]	−2.2[b]
AlAs	12.6[b]	6.27[b]	0.58[b]	4.82[b]	−2.5[b]
AlSb	9.00	7.5[a]	4.0[a]	6.4[a]	−1.5
GaP	10.5	5.8	1.8[a]	2.1[a]	−1.1
GaAs	10.7	5.0	4.6[a]	5.5[a]	−1.5[a]
GaSb	14.7	7.5	6.0	5.0	−1.5[c]
InP	8.5	7.5[a]	4.6[a]	6.8[a]	1.8[a]
InAs	10.15	7.0	3.5[a]	4.8[a]	−0.02[a]
InSb	15.5	8.5	6.0	8.3[a]	−1.1[c]
ZnSe	7.0	2.5[a]	−0.3[a]	1.4[a]	−2.0[a]

[a]Reference 20.
[b]R.-D. Hong, S. Lee, and J. D. Dow (unpublished).
[c]Reference 21. All other experimental values are those cited in Ref. 20.

$$V_s=\beta_s(w_s(\text{impurity})-w_s(\text{host}))\ ,$$

$$V_p=\beta_p(w_p(\text{impurity})-w_p(\text{host}))\ ,$$

and

$$V_{s^*}=0\ ,$$

where the energies w_l ($l=s$ or p) are atomic-orbital energies in the solid,[14] and β_s and β_p are constants (0.8 and 0.6, respectively[2]). These approximations to the defect-potential matrix of a specific impurity, V, are needed to associate a particular deep-level energy E or pressure derivative dE/dp with the impurity. They are not necessary to obtain a relationship between dE/dp and E, however, because this relationship depends only on the existence of such a matrix V, not on our ability to accurately predict the numerical values of its matrix elements.

Hydrostatic pressure does not affect the defect potential within the context of the Hjalmarson model, because the defect potential does not depend on the bond length. It does alter the off-diagonal two-center matrix elements of the host Hamiltonian H_0, however, because these matrix elements depend on the bond length d:

$$H_{l,l'}=H^0_{l,l'}(d_0/d)^{\eta_{l,l'}}\ .$$

Here, $H^0_{l,l'}$ and $H_{l,l'}$ ($l,l'=s$, p, and s^*) are the off-diagonal matrix elements corresponding to the bond lengths d_0 and d, respectively. d_0 is the zero-pressure bond length; the finite-pressure bond length d is obtained from the hydrostatic pressure p by using Murnaghan's[18] equation of state

$$p=[B_0/(dB_0/dp)][(d_0/d)^{3\,dB_0/dp}-1]\ ;$$

and $\eta_{l,l'}$ are exponents with values near 2, according to both Harrison[19] and Vogl *et al.*[14] We have obtained, by trial and error, sets of exponents $\eta_{l,l'}$ that reproduce rather well the observed deformation potentials or hydrostatic-pressure derivatives of the band gaps at Γ, X, and L symmetry points of the Brillouin zone. The exponents $\eta_{l,l'}$ are presented in Table I, and the experimental pressure coefficients used to determine them are given in Table II.[20,21]

III. RESULTS

For the tetrahedral (T_d) symmetry of zinc-blende semiconductors, the secular determinant reduces to two scalar equations, one for s-like A_1-symmetry deep levels,

$$(V_s)^{-1}=\sum_{\mathbf{k},\lambda}|\langle s,b,0\,|\,\mathbf{k},\lambda\rangle|^2/[E-E(\mathbf{k},\lambda)]\ ,$$

and another for p-like T_2 levels,

$$(V_p)^{-1}=\sum_{\mathbf{k},\lambda}|\langle p,b,0\,|\,\mathbf{k},\lambda\rangle|^2/[E-E(\mathbf{k},\lambda)]\ .$$

We evaluate the sums using the special-points method[22] for fixed E, and then graphically determine the defect potentials V that produce a level at that energy. This process is repeated for the pressurized semiconductor to obtain dE/dp versus V also. Then the defect potential V is eliminated to yield dE/dp as a function of E.

It should be emphasized that there are two levels of approximation in the theory for the defect-potential matrix V: (i) the diagonal form (with arbitrary matrix elements V_s and V_p) which is the only approximation entering into the determination of dE/dp as a function of E, and introduces a small theoretical uncertainty of ~ 0.5 meV/kbar into dE/dp;[11] and (ii) the expressions of V_s and V_p in terms of atomic energies, which are necessary to associate a deep level E with a particular impurity (or defect potential) and introduce an uncertainty in E of about ~ 0.3 eV. Thus, the uncertainty in dE/dp as a function of V, which is the combination of these two uncertainties, is considerably larger than the uncertainty in dE/dp as a function of E (by typically a factor of 4, as can be deduced, for ex-

ample, from Fig. 2). Thus, although the uncertainty in E *for a particular defect* is comparable with the band gap of InSb and the uncertainty in dE/dp for that defect is a significant fraction of the predicted range of possible values for dE/dp, the uncertainty in dE/dp *for a particular level* E is considerably smaller, and so that relationship can be used to determine the symmetry of the deep level and the site of its parent impurity, even though unambiguous determination of the defect, namely, highly accurate determination of $E(V)$, is not possible. Here we exploit this fact and note that similar elimination of the defect potential from the theory of ENDOR and ESR spectra of deep levels produced successful and accurate theories.[6]

The theoretical uncertainty of ~ 0.5 meV/kbar in dE/dp was first arrived at[11] by varying the major elements of the theory, such as the tight-binding matrix elements and exponents η, over the range of reasonable possibilities. It has been borne out by experiments for the A_1 deep level of a N impurity on an anion site in GaP (Refs. 12 and 23) and in GaAs (Ref. 24), which have exhibited the predicted pressure coefficients dE/dp. The theory can rather accurately predict a derivative dE/dp associated with a deep level E, even though it cannot accurately assign an impurity to a given energy.

Since the relationship dE/dp versus E depends on the site of the impurity and the symmetry of its deep level, comparison of data for dE/dp and E with theory can yield the site of the impurity and the symmetry of its deep level. Once the site and the symmetry are known, the number of candidates for producing such a level E (to within a few tenths of an eV) is greatly reduced.

We illustrate this point for deep levels in InSb, a material whose band gap is smaller than the theoretical uncertainty. The predicted pressure derivatives dE/dp are given in Fig. 2 as functions of the deep-level energies E. (The estimated theoretical uncertainty in dE/dp is roughly 0.5 meV/kbar.) The impurities associated theoretically with the deep levels E are also displayed on each curve—although this association is limited by a few-tenths-of-an-eV theoretical uncertainty in the energy E of the deep level associated with a specific impurity. Therefore we have shown on the sides of the figure those impurities (including the vacancy, denoted Va) that might have deep levels in the band gap if the theory's deep-level predictions were altered by 0.3 eV. We have considered as impurities the atoms from columns IIB and III–VIII of the Periodic Table, as well as Li, Na, K, Rb, Be, and Mg.

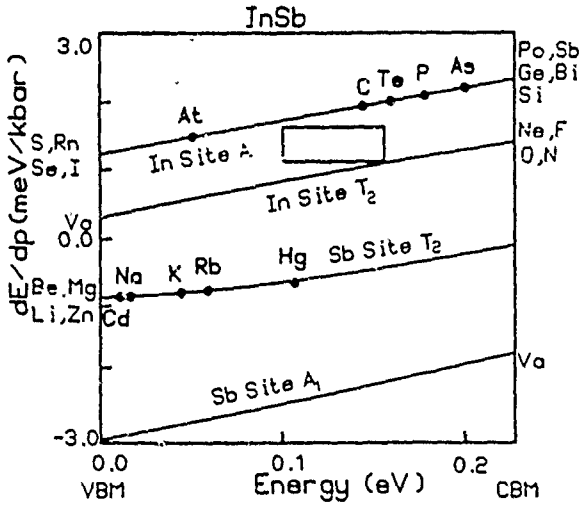


FIG. 2. Pressure coefficients dE/dp in meV/kbar for deep defect levels in InSb as functions of their energies E (in eV) in the band gap. The predicted values of dE/dp and E for specific impurities on particular sites are indicated by circles. On the sides of the figure are impurities expected to produce deep levels in the host bands, but within 0.3 eV of the gap. The boxed, shaded region corresponds to data of Aladashvili *et al.* (Ref. 25).

To illustrate how the theory can be applied to identify deep impurities, consider the deep levels observed by Aladashvili *et al.*[25] near 0.15 eV in InSb, with pressure derivatives of ~ 1 meV/kbar. These levels (see Fig. 2) correspond to either In-site T_2 levels or In-site A_1 levels within the uncertainty of 0.5 meV/kbar. However, no In-site T_2 levels lie at such energies, to within ± 0.3 eV, indicating that the only candidates from the set of s- and p-bonded substitutional impurities for producing these levels are the In-site A_1 levels: S, Rn, Se, I, At, C, Te, P, As, Po, Sb, Ge, Bi, and Si. Thus we have reduced the possibilities to some impurities from Columns IV, V, and VI of the Periodic Table on the In site, producing A_1-symmetric levels.

Two of these defects are very likely to be present in InSb: C and Sb. While C is very likely more soluble on the In site than Sb, the native antisite defect should also be easily formed. In this regard, we note that the concentration of antisite defects should be greatly increased by radiation damage, and so such studies of Aladashvili's deep levels should permit identification of the deep impurity, if it is indeed an antisite defect.

The observation of several deep levels near 0.1 eV probably is due either to the defects being clustered in complexes, or to there being different impurities, such as C and Sb with nearly equal deep-level energies.

In any case, this illustration for InSb shows how hydrostatic-pressure data, when combined with deep-level data, can determine the site of the impurity and the symmetry of the deep level, while simultaneously restricting the candidates for producing the deep level to a few impurities.

The predictions for deep impurities for other zinc-blende semiconductors are given in Figs. 3–13. For the most part we find dE/dp larger for cation-site levels than for anion-site levels, although this situation can be reversed for some (especially T_2 symmetric) levels if the levels are very near the conduction-band edge.

We can understand these calculated results qualitatively by using a defect-molecule model, as shown in Fig. 14. In Fig. 14(a) are shown the In and Sb atomic energy levels. After solid InSb is formed, the atomic energy levels become levels in the solid, as depicted in Fig. 14(b), and their ordering is different from in the free atoms, due to interaction between atoms. Under hydrostatic compres-

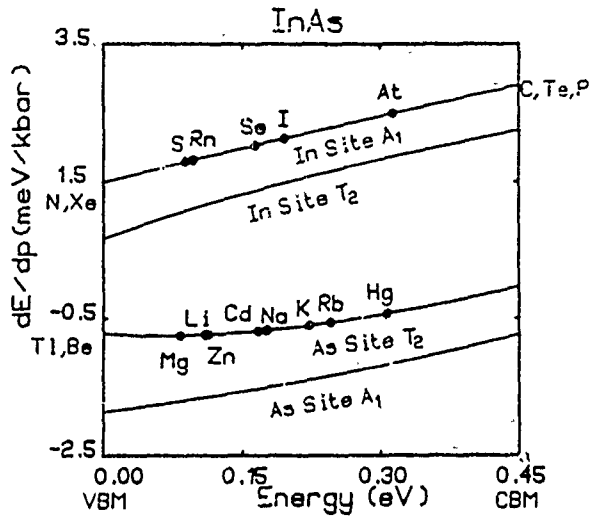


FIG. 3. Pressure coefficients dE/dp in meV/kbar for deep defect levels in InAs as functions of their energies E (in eV) in the band gap.

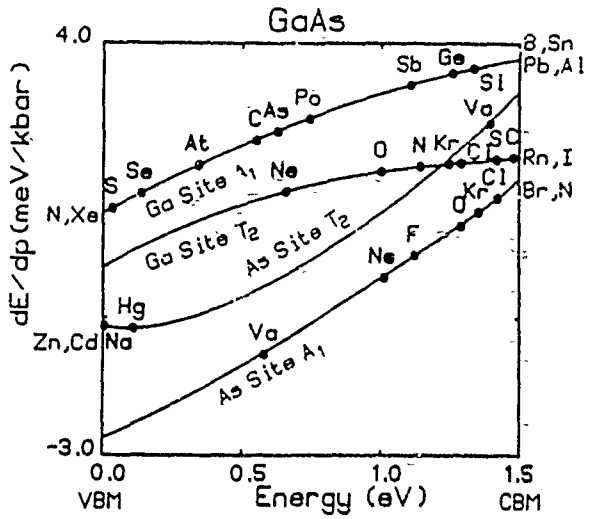


FIG. 6. Pressure coefficients dE/dp in meV/kbar for deep defect levels in GaAs as functions of their energies E (in eV) in the band gap.

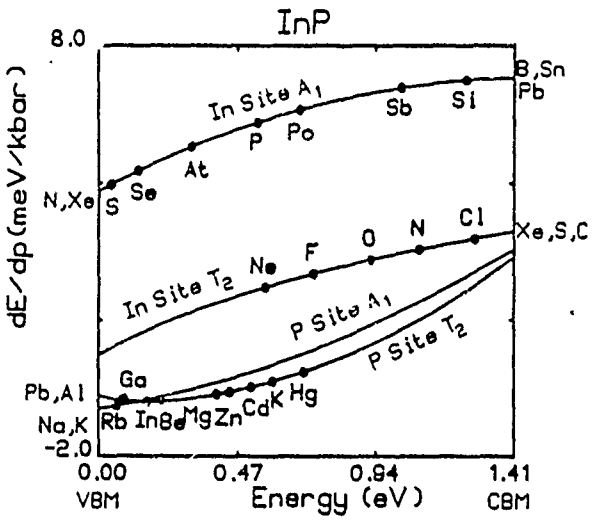


FIG. 4. Pressure coefficients dE/dp in meV/kbar for deep defect levels in InP as functions of their energies E (in eV) in the band gap.

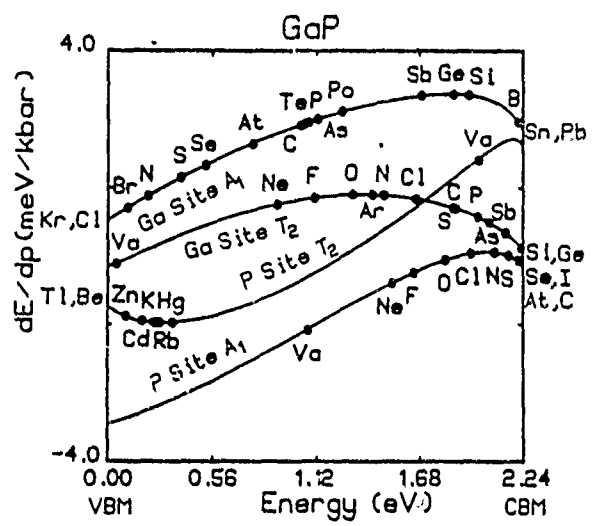


FIG. 7. Pressure coefficients dE/dp in meV/kbar for deep defect levels in GaP as functions of their energies E (in eV) in the band gap.

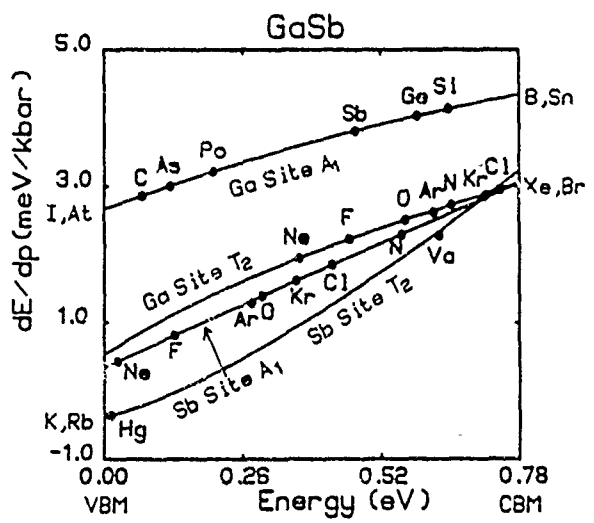


FIG. 5. Pressure coefficients dE/dp in meV/kbar for deep defect levels in GaSb as functions of their energies E (in eV) in the band gap.

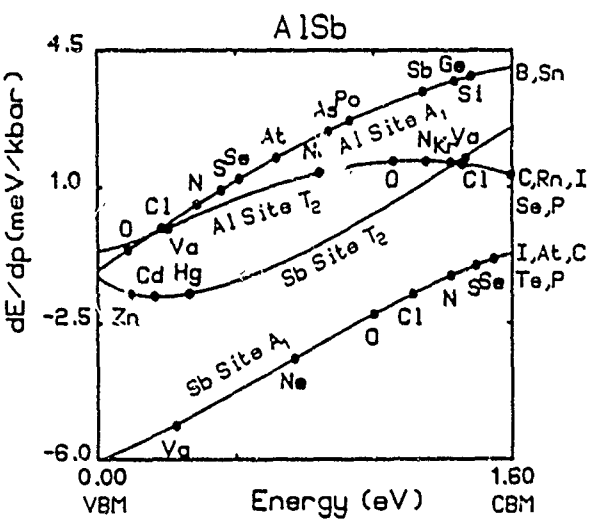


FIG. 8. Pressure coefficients dE/dp in meV/kbar for deep defect levels in AlSb as functions of their energies E (in eV) in the band gap.

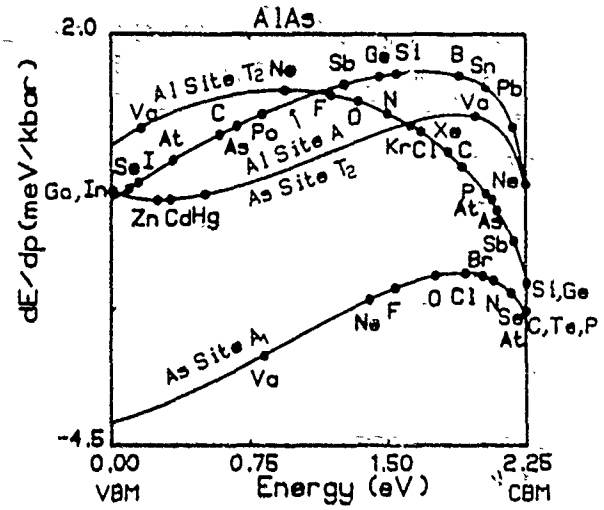


FIG. 9. Pressure coefficients dE/dp in meV/kbar for deep defect levels in AlAs as functions of their energies E (in eV) in the band gap.

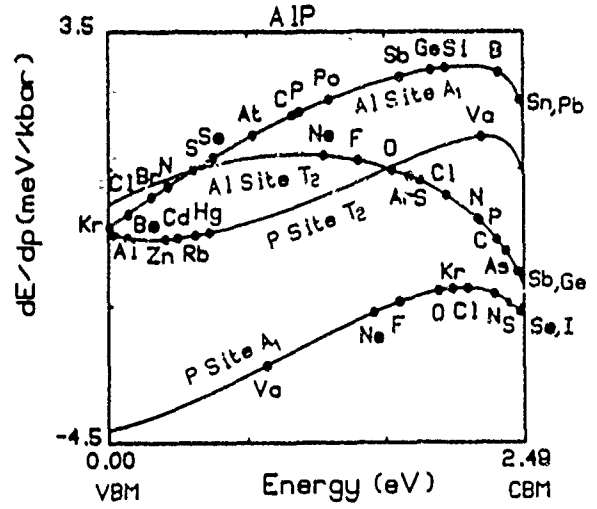


FIG. 10. Pressure coefficients dE/dp in meV/kbar for deep defect levels in AlP as functions of their energies E (in eV) in the band gap.

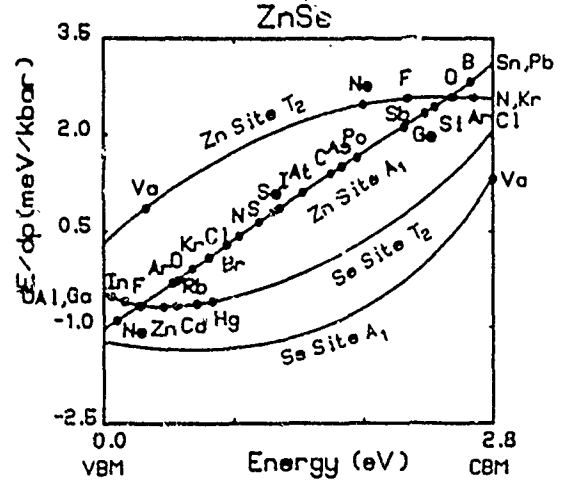


FIG. 11. Pressure coefficients dE/dp in meV/kbar for deep defect levels in ZnSe as functions of their energies E (in eV) in the band gap.

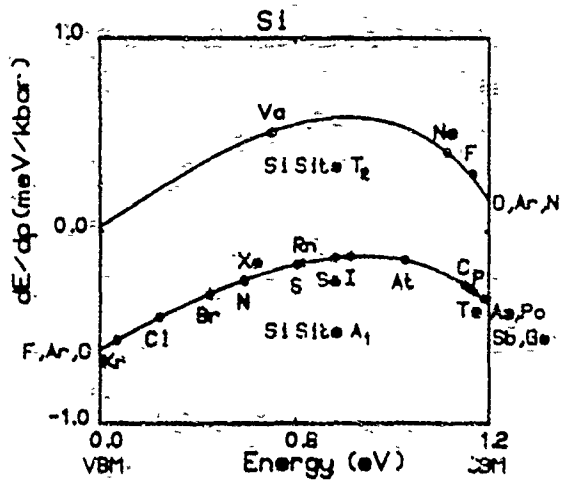


FIG. 12. Pressure coefficients dE/dp in meV/kbar for deep defect levels in Si as functions of their energies E (in eV) in the band gap.

sion, the interaction between atoms increases, and the levels shift as shown by arrows in Fig. 14(b). (In the solid, these energy levels are energy levels of a defect molecule consisting of one atom and its four nearest neighbors—and so the levels should be thought of qualitatively as representing the center of gravity of the corresponding partial densities of states.) The In T_2 level must move up in energy and the Sb A_1 level must go down, due to level repulsion. The nearest-neighbor coupling dictates that level repulsion cause the In A_1 level to move up—in the opposite direction of Sb A_1. Similarly, the Sb T_2 level moves to lower energy. Since the valence-band maximum has predominantly Sb T_2 character and is taken to be the zero of energy, the pressure

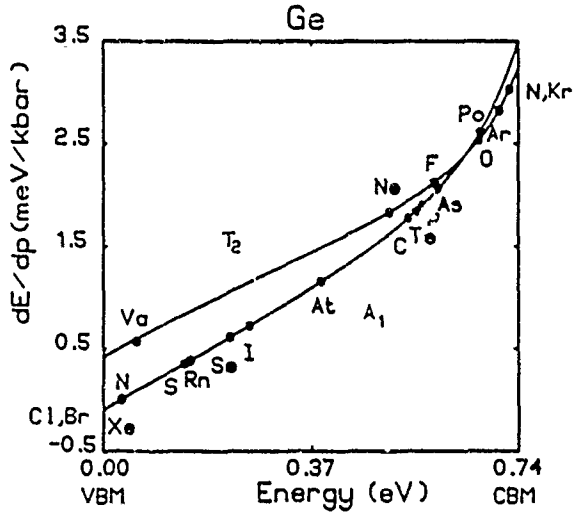


FIG. 13. Pressure coefficients dE/dp in meV/kbar for deep defect levels in Ge as functions of their energies E (in eV) in the band gap.

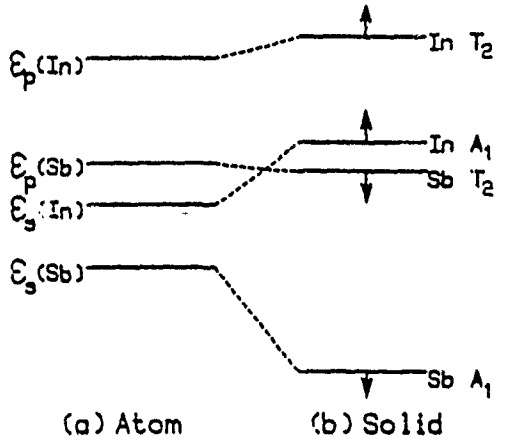


FIG. 14. Defect-molecule model for InSb. (a) Atomic energy levels (s and p orbitals) for In and Sb. (b) In the zinc-blende crystal symmetry, the atomic energy levels form singlet A_1 levels and triplet T_2 levels. The shaded region is approximately the band gap of InSb. The arrows indicate the expected directions of movement of these host levels (and the impurity levels derived from them) when hydrostatic compression is applied to the InSb crystal.

coefficient of the Sb T_2 levels are small in magnitude. The In-site A_1-symmetric impurity levels in the band gap have a larger pressure dependence than In-site T_2 levels, because the A_1 host levels are closer to the gap and hence repel impurity levels more. These general rules normally govern the pressure dependencies of the deep impurity levels—although they are sufficiently qualitative in character that exceptions to them are to be expected.

In summary, we have predicted the hydrostatic-pressure dependencies of deep levels in 12 semiconductors and have shown how pressure data can be analyzed to yield the site of the deep impurity, the symmetry of the deep level, and either the impurity itself or a handful of candidates likely to be the impurity responsible for the deep level.

ACKNOWLEDGMENTS

We thank Seongbok Lee for stimulating conversations and the U. S. Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support.

*Present address: Amoco Technology, P.O. Box 400, Mail Stop F4, Naperville, IL 60566.

Permanent address: University of Science and Technology of China, Hefei, Anhui, People's Republic of China.

[1]H. P Hjalmarson, Ph.D. thesis, University of Illinois, 1979.

[2]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980). See also W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[3]J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. Lett. 41, 895 (1978); Phys. Rev. B 21, 3545 (1980).

[4]G. A. Baraff and M. Schlüter, Phys. Rev. Lett. 41, 892 (1978); Phys. Rev. B 19, 4965 (1980).

[5]M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969).

[6]S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982); S. Y. Ren, Sci. Sin. A XXVII, 443 (1984).

[7]G. W. Ludwig, Phys. Rev. 137A, 1520 (1965); J. M. Spaeth, D. M. Hofman, B. K. Meyer, in *Microscopic Identification of Electronic Defects in Semiconductors*, Mater. Res. Soc. Symp. Proc. No. 46, edited by N. M. Johnson, S. G. Bishop, and G. D. Watkins (MRS, Pittsburgh, 1985), p. 185.

[8]H. G. Grimmeiss, E. Janzén, H. Emen, O. Schirmer, J. Schneider, R. Wörner, C. Holm, E. Sirtl, and P. Wagner, Phys. Rev. B 24, 4571 (1981).

[9]B. A. Bunker, in *Interfaces, Superlattices, and Thin Films*, Mater. Res. Soc. Symp. Proc. No. 77, edited by J. D. Dow and I. K. Schuller (MRS, Pittsburgh, 1987), p. 545.

[10]O. F. Sankey and J. D. Dow, Phys. Rev. B 27, 7641 (1983); 26, 3243 (1982).

[11]S. Y. Ren, J. D. Dow, and D. J. Wolford, Phys. Rev. B 25, 7661 (1982).

[12]S. Y. Ren, Chin. Phys. Lett. 2, 93 (1985).

[13]D. W. Jenkins, S. Y. Ren, and J. D. Dow (unpublished).

[14]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[15]P. O. Löwdin, J. Chem. Phys. 18, 365 (1950).

[16]J. D. Dow, in *Highlights of Condensed-Matter Theory*, Proceedings of the International School of Physics "Enrico Fermi," Course 89, Varenna, 1983, edited by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna/North-Holland, Amsterdam, 1985), pp. 465–494.

[17]S. Lee, J. D. Dow, and O. F. Sankey, Phys. Rev. B 31, 3910 (1985); G. Kim, J. D. Dow, and S. Lee, Phys. Rev. B (to be published).

[18]F. D. Murnaghan, Proc. Natl. Acad. Sci. U.S.A. 30, 244 (1944).

[19]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

[20]Y. F. Tsey, S. S. Mitra, and B. Bendow, Phys. Rev. B 10, 1476 (1974).

[21]J. Gorbzyca, Phys. Status Solidi B 112, 97 (1982); 125, 229 (1984).

[22]D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).

[23]J. L. Merz, A. Baldereschi, and A. M. Sergent, Phys. Rev. B 6, 3082 (1972).

[24]D. J. Wolford, J. A. Bradley, K. Fry, and J. Thompson, in *Proceedings of the 17th International Conference Physics of Semiconductors*, edited by D. J. Chadi and W. A. Harrison (Springer-Verlag, New York, 1984), p. 627. See also X. S. Zhao, G. H. Li, H. X. Han, Z. P. Wang, R. M. Tang, and R. Z. Che, Chin. Phys. 5, 337 (1985).

[25]D. I. Aladashvili, L. Konczewicz, and S. Porowski, Phys. Status Solidi A 86, 301 (1984).

Dependence on uniaxial stress of deep levels in III-V compound and group-IV elemental semiconductors

David W. Jenkins,* Shang Yuan Ren,† and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 29 June 1988)

The uniaxial-stress dependences of substitutional *s*- and *p*-bonded deep impurity levels in the semiconductors AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb, Si, and Ge are studied theoretically for stresses applied along the [100], [110], and [111] directions. We find that stress applied along the [110] direction, in particular, causes splittings and shifts of deep levels associated with a point defect that can be used (i) to determine uniquely the symmetries of the levels (*s*-like or *p*-like) and (ii) to identify the site (anion or cation) of the associated impurity.

I. INTRODUCTION

Substitutional defects in semiconductors give rise to localized states with energy levels often lying deep in the fundamental energy-band gap. Given experimental evidence of such a level, it is difficult to associate the level with a specific impurity, based on the energy of the level alone, due to the hostlike nature of substitutional impurity states. Although different impurities have defect potentials that differ by typically 1 to 10 eV,[1] their deep-level wave functions are antibonding and hostlike in character, with very little amplitude in the impurity cell. As a result their wave functions depend very little on the impurity and hence the deep energy levels of different impurities often lie within a few tenths of an eV of one another.

Since, to a good approximation, the wave functions of an *s*-like (A_1) or *p*-like (T_2) deep level associated with an impurity on a specific site (anion or cation) is *independent of the impurity*,[2] it is very difficult to perform an electronic experiment on a deep level that will identify the impurity responsible for it. Experiments that probe the core [such as extended x-ray-absorption fine structure (EXAFS) (Ref. 3)] or the nucleus [such as electron-nuclear double resonance (ENDOR) (Ref. 4)] are needed for such identification. Nevertheless, as we show here, electronic experiments that probe the deep level's energy or wave function can determine the site of the associated defect and the symmetry of its deep level, and, when information from several such electronic experiments is combined, it is often possible to eliminate all but a few impurities as candidates for producing the level.

In this paper we predict the uniaxial stress dependences of deep levels in semiconductors and show how these dependences can be used to facilitate the association of specific impurities with observed deep levels. We treat *s*- and *p*-bonded substitutional impurities in the semiconductors AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb, Si, and Ge, with the applied stress in the [100], [110], and [111] directions.

Deep impurity states are due to the central-cell potential and are fundamentally different from the well-known shallow levels described by effective-mass theory.[5] Under hydrostatic pressure, the shallow levels tend to follow the associated band edge;[6] the levels shift in energy an amount roughly equal to the shift in the associated band edge. In contrast, deep levels do not follow the band edge. For example, Wolford *et al.*[7] have shown that the conduction-band edge in GaAs rises faster in energy under hydrostatic pressure than the N deep level. At zero pressure this level is actually resonant with the conduction band and emerges into the fundamental energy-band gap at a pressure of 22 kbar. The level clearly does not follow the conduction-band edge and so is necessarily a deep level. (Here, we use the new definition of a deep level[1] as one caused by the defect's central-cell potential, and not the old definition: a level more than 0.1 eV deep in the gap. As a result, we term the *s*-like N A_1-symmetric level in GaAs deep, although it lies above the conduction-band edge.)

Hydrostatic-pressure measurements can confirm that a level is not attached to any nearby band edge, and therefore is indeed deep. Moreover, Ren *et al.*[8] and Hong *et al.*[9] have shown that quantitative analyses of such measurements can determine the site, anion or cation, and the symmetry, A_1 (*s*-like) or T_2 (*p*-like), for some, but not all, deep levels. In contrast to hydrostatic pressure, which merely shifts the energies of the deep levels, uniaxial stress causes T_2-symmetric deep levels to split, while the orbitally nondegenerate A_1 levels shift. Hence uniaxial stress can be even better than hydrostatic pressure for determining the symmetry of a deep level. Moreover, the size of the uniaxial-stress-induced splittings depends both on the site of the impurity and on the magnitude and direction (relative to the crystal axes) of the applied stress. Hence, analyses of stress-induced deep-level splittings can help determine not only the symmetry of the level, but also whether the impurity is on the anion or the cation site.

II. THEORY

Our calculations of the dependencies on uniaxial stress of deep levels are based on the theory of Hjalmarson

 ©1989 The American Physical Society

et al.[1] and extensions of those theories by Ren *et al.*[8] Our basic approach is to construct an empirical tight-binding Hamiltonian fo the stressed semiconductor, based on the ideas of Vogl *et al.*[10] In order to do this we must first determine how a particular uniaxial stress changes bond lengths in the semiconductor. We consider two types of stress-induced changes of atomic positions and bond lengths: (i) If there is not a shear strain, the changes of the relative positions of all the atoms in the crystal are given simply by first-order elasticity theory,[11] and (ii) if there is a shear strain, "internal displacements"[12] must also be considered.

In terms of first-order elasticity theory, the change in the x coordinate of an atom initially at position $\mathbf{r}=(x_0,y_0,x_0)$ relative to the origin is

$$\Delta x = P\{[c_{12}-(c_{11}+2c_{12})n_x^2]x_0/(c_{11}+2c_{12})(c_{11}-c_{12}) - n_x n_y y_0/2c_{44} - n_x n_z z_0/2c_{44}\}, \quad (1)$$

where c_{11} and c_{12} are the longitudinal and transverse elastic constants and c_{44} is the shear elastic constant. Similar expressions hold for other components of the displacement $\Delta\mathbf{r}$.[13] Here, the stress is

$$\mathbf{P}=P(n_x,n_y,n_z), \quad (2)$$

where P is the magnitude of the applied force per unit area and n_x,n_y,n_z are the cosines defining the direction of the applied stress.

Under uniaxial stress, the Brillouin zone in momentum space also changes due to changes in the atomic positions. In terms of first-order elasticity theory, the change in the x component of the wave vector $\mathbf{k}=(k_x^0,k_y^0,k_z^0)$ is

$$\Delta k_x = P\{[(c_{11}+2c_{12})n_x^2-c_{12}]k_x^0/(c_{11}+2c_{12})(c_{11}-c_{12}) + n_x n_y k_y^0/2c_{44} + n_x n_z k_z^0/2c_{44}\}. \quad (3)$$

Similar expressions hold for other components of the displacement $\Delta\mathbf{k}$.[13]

Elasticity theory adequately describes the change in positions, due to applied stress, of all the atoms in the semiconductor *except* for the atom at the origin [see Eq. (1)]. Kleinman[12] pointed out that if we choose the origin at the unstrained position of an atom, that atom moves a distance proportional to the applied stress P, when the stress involves shear, as is the case for pressures directed along the [110] or [111] axes (but not for the [100] stress). Segmüller and Neyer confirmed Kleinman's internal-displacement prediction for [111] stress in Ge and Si (see Figs. 1–3).[14] Under [111] stress, the internal displacement is $-\zeta P a_L(1,1,1)/12c_{44}$, where P is the magnitude of the stress, c_{44} is the elastic constant for shear strains, a_L is the lattice constant, and ζ is Kleinman's internal-displacement parameter[12] (Fig. 3). Under [110] stress, the internal displacement is $-\zeta P a_L(0,0,1)/8c_{44}$.

Harrison has calculated Kleinman's parameter for a number of zinc-blende III-V compound and group-IV elemental semiconductors, and has found all to be close to 0.6.[15] Thus we take $\zeta=0.6$ for all semiconductors studied here. We then compute the nearest-neighbor bond lengths d given in Table I (see Figs. 1–3), including both the effects of first-order elasticity theory and internal-displacement theory.

With the stress-perturbed bond lengths determined from first-order elasticity theory and internal-displacement theory, it is possible to construct an empirical tight-binding Hamiltonian for the stressed semicon-

TABLE I. Relative change in bond lengths, d/d_0, between atoms at the origin (zero-stress) and the nearest-neighbor atoms in a zinc-blende crystal structure under stress applied in the [100], [110], and [111] directions (see Figs. 1–3). The bulk modulus is defined in elasticity theory by $B=(c_{11}+2c_{12})/3$

Stress direction	Zero-stress position of neighbor	Relative change in bond length d/d_0
[100]	$\frac{1}{4}a_L(1,1,1)$	$(1-P/9B)$
[100]	$\frac{1}{4}a_L(1,-1,-1)$	$(1-P/9B)$
[100]	$\frac{1}{4}a_L(-1,1,-1)$	$(1-P/9B)$
[100]	$\frac{1}{4}a_L(-1,-1,1)$	$(1-P/9B)$
[110]	$\frac{1}{4}a_L(1,1,1)$	$(1-(P/9B)\{1+3B[(1-\zeta)/2c_{44}]\})$
[110]	$\frac{1}{4}a_L(1,-1,-1)$	$(1-(P/9B)\{1-3B[(1-\zeta)/2c_{44}]\})$
[110]	$\frac{1}{4}a_L(-1,1,-1)$	$(1-(P/9B)\{1-3B[(1-\zeta)/2c_{44}]\})$
[110]	$\frac{1}{4}a_L(-1,-1,1)$	$(1-(P/9B)\{1+3B[(1-\zeta)/2c_{44}]\})$
[111]	$\frac{1}{4}a_L(1,1,1)$	$(1-(P/9B)\{1+3B[(1-\zeta)/c_{44}]\})$
[111]	$\frac{1}{4}a_L(1,-1,-1)$	$(1-(P/9B)\{1-B[(1+\zeta)/c_{44}]\})$
[111]	$\frac{1}{4}a_L(-1,1,-1)$	$(1-(P/9B)\{1-B[(1+\zeta)/c_{44}]\})$
[111]	$\frac{1}{4}a_L(-1,-1,1)$	$(1-(P/9B)\{1-B[(1+\zeta)/c_{44}]\})$

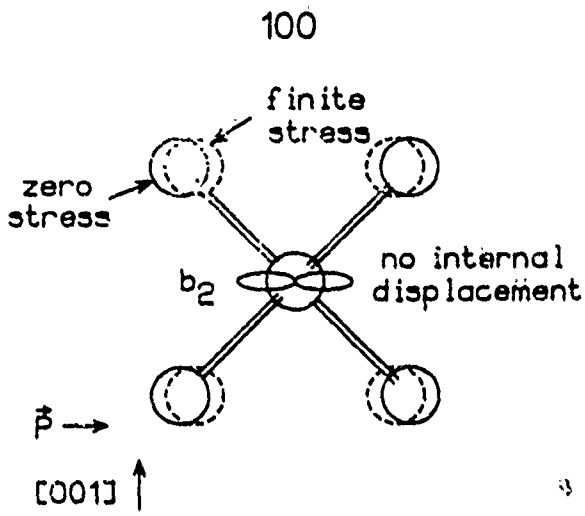


FIG. 1. Schematic illustration of change of atomic positions of an impurity and its neighbors, induced by uniaxial stress in the [100] direction. Solid (dashed) circles are the positions of atoms under zero (finite) stress. Each of the neighbors shift closer to the impurity, by the same distance. Each neighbor also shifts closer to another neighbor. The impurity does not move under [100] applied stress. The stress causes p-like T_2 impurity levels to split: one state, b_2, is directed along the stress direction and the other two states (not shown) are perpendicular to the stress direction.

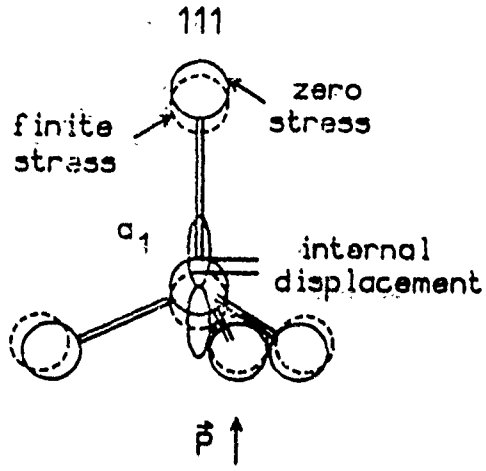


FIG. 3. Schematic illustration of change of atomic positions of an impurity and its neighbors, induced by uniaxial stress in the [111] direction. Solid (dashed) circles are the positions of atoms under zero (finite) stress. One neighbor shifts closer to the impurity. The other three neighbors shift slightly further from the impurity and further from each other. The impurity displaces in the $[\bar{1}\,\bar{1}\,\bar{1}]$ direction (see text). The stress causes p-like T_2 impurity levels to split: one state, a_1, is directed along the stress direction and two states, e (not shown), are directed perpendicular to the stress.

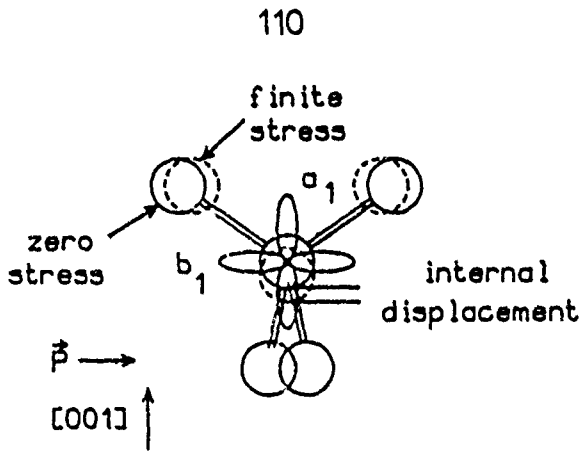


FIG. 2. Schematic illustration of change of atomic positions of an impurity and its neighbors, induced by uniaxial stress in the [110] direction. Solid (dashed) circles are the positions of atoms under zero (finite) stress. Two of the neighbors shift closer to the impurity by the same distance and the other two neighbors shift very little. The two neighbors that shift also shift closer to each other. The impurity displaces in the $[00\bar{1}]$ direction (see text). The stress causes p-like T_2 impurity levels to split: one state, b_2, is directed along the stress direction and two states are directed perpendicular to the stress—a_1 is directed along the [001] direction and b_1 is (not shown) directed along the $[1\bar{1}0]$ direction.

ductor. Following the ideas of Slater and Koster,[16] Harrison,[17] and Vogl *et al.*,[12] we use a ten-band, nearest-neighbor model to treat the electronic structure, with an sp^3s^* orbital basis centered on each atomic site. (The sp^3 basis is needed to reproduce the chemistry of the covalent bonding; the additional orbital, s^*, is required to yield indirect-gap band structures, such as those of Si and GaP.) In the zero-stress case the model is the same as that of Vogl *et al.*, and has the property that the diagonal (on-site) matrix elements depend on the atomic energies, but not on the bond lengths. Thus we assume that stress does not alter the Hamiltonian's on-site matrix elements, but that the stress-induced bond-angle distortions and bond-length changes can be incorporated into the off-diagonal matrix elements of the Vogl Hamiltonian by using the Slater-Koster definitions[16] of these matrix elements and a generalization of Harrison's rule[17] for their dependence on bond length: for the matrix element between s and p_x states on neighboring states we have

$$V_{s,p_x} = V^0_{s,p_x}[\cos(\theta)/\cos(\theta_0)](d_0/d)^{\eta_{s,p}} , \tag{4}$$

with similar expressions for the dependences of the other off-diagonal matrix elements on d and θ. Here d_0 and d are the zero- and finite-stress bond lengths; θ_0 and θ are the zero- and finite-stress bond angles relative to the crystal axes, and V^0_{s,p_x} is the zero-stress matrix element. The coefficients $\eta_{\alpha,\beta}$ were fitted previously[8,9] to the hydrostatic-pressure dependences of the energy-band gap and are listed in Table II. With these matrix elements it is possible to construct the Hamiltonian H_0 for the stressed semiconductor.

TABLE II. Exponents η for the bond-length dependence of the nearest-neighbor matrix elements [see Eq. (2)] and elastic constants c_{11}, c_{12}, and c_{44}. Exponents are taken from Refs. 8 and 9 and elastic constants from Ref. 17, unless noted otherwise.

	$\eta_{s,s}$	$\eta_{s,p}$	$\eta_{x,x}$	$\eta_{x,y}$	$\eta_{s^*,p}$	c_{11}	c_{12}	c_{44}
AlP	2.386	1.637	1.521	1.247	2.486	(14.12	6.25	7.05)[a]
AlAs	3.205	1.656	2.398	1.706	3.214	(9.85	4.44	4.95)[b]
AlSb	2.553	4.249	1.192	3.272	4.469	8.94	4.43	4.16
GaP	3.697	2.804	1.630	2.795	.841	14.12	6.25	7.05
GaAs	4.144	2.341	2.596	2.220	2.665	11.81	5.32	5.92
GaSb	2.719	2.923	2.119	1.172	3.891	8.84	4.03	4.32
InP	3.100	4.443	3.049	2.366	1.207	10.22	5.76	4.60
InAs	2.539	2.812	3.757	2.825	3.014	8.33	4.53	3.96
InSb	4.012	2.987	2.533	2.751	3.134	6.67	3.65	3.02
Si	3.000	1.600	3.825	2.600	3.327	16.57	6.39	7.96
Ge	4.400	2.400	2.300	2.500	3.982	12.89	4.83	6.71

[a]Elastic constants for AlP are taken to be the same as those for GaP. We are unaware of any experimental values for AlP.

[b]Elastic constants for AlAs are taken from C. Coivard, R. Merlin, M. V. Klein, and A. C. Gossard, Phys. Rev. Lett. 45, 298 (1980).

The defect levels are calculated by solving Dyson's equation,

$$\det(1-G_0V)=0\,, \tag{5}$$

where G_0 is the host Green's function and V is the defect potential. The host Green's-function operator is

$$G_0(E)=(E-H_0)^{-1}\,. \tag{6}$$

The defect potential V is a diagonal matrix with its nonzero diagonal matrix elements V_s and V_p determined using the prescription of Hjalmarson *et al.*[1] The impurity levels are computed for both zero stress and finite stress. Stress derivatives are then computed at a small pressure (P=0.5 kbar).[18]

While we could plot E versus V_s or V_p and (as originally predicted by Hjalmarson *et al.*[1]) dE/dP versus the appropriate V, thereby predicting both deep levels and their pressure derivatives, it is now well established that the theory produces very reliable relationships such as E versus dE/dP, whereas its absolute predictions of energy levels (E versus V) are less reliable.[2] Therefore we plot dE/dP versus E, while labeling the lines of each figure with the impurities whose defect potentials V are predicted to produce deep levels at the energy E. The reason that the relationships are so reliable is that they can be calculated without knowing the defect potential V precisely. A difference of only a few eV in the defect potential can cause a few-tenths-of-an-eV difference in an energy level in the gap, but the uncertainty in V does not affect the relationship between dE/dP and E. Deep-level wave functions are typically antibonding and hostlike,[2] and so, although defect potentials that produce deep levels in the band gap may differ by large amounts, say several eV, the energy levels and wave functions of those deep levels are much closer to one another—indicating that a theory of dE/dP as a function of E would be free of the relatively large uncertainties (associated with the approximate value of the defect potential V) to be found in a theory of dE/dP as a function of V.

III. RESULTS

The predicted uniaxial-stress derivatives dE/dP of deep levels in the band gap of AlAs, for stress applied in the [100], [110], and [111] directions, are shown in Figs. 4, 5, and 6 respectively. The pressure coefficients dE/dP are plotted as functions of defect energy level (relative to the valence-band edge) for all s- and p-bonded impurities with levels in the gap. These results are representative of the III-V compound semiconductors.

Note that the signs and the magnitudes of the splittings and the shifts depend rather sensitively on the energy of the level in the gap, on the site of the impurity, and on the direction of the applied stress.

No simple ordering of pressure derivatives is to be expected. This is because the pressure derivatives of deep levels are sensitive to shifts in pressure of both conduction- and valence-band densities of states—as pointed out by Hong *et al.*[9] for the case of hydrostatic pressure (see below.)

A. Symmetry

Uniaxial stress shifts the s-like A_1 levels and splits the p-like T_2 levels, regardless of whether the stress is applied along the [100], [110], or [111] axes. For stress along the [110] direction, the threefold degeneracy of a T_2 state is completely removed, whereas in the [100] and [111] directions a threefold-degenerate T_2 state is split into a twofold-degenerate level and a nondegenerate level (see below.) This splitting, if it is observed, will distinguish a

T_2 level from an A_1 level.[19] However, the absence of an observable splitting is not conclusive evidence for the A_1 character of a level—because the splitting may be too small to be resolved.

There do not appear to be any general rules stating, for example, that T_2 states associated with anion-site impurities split more than cation-site impurities. Therefore, while the splittings may indicate the symmetry of a deep level, quantitative analyses of the observed splittings may be necessary to determine the site of the parent impurity.

B. A_1-derived states

For uniaxial stress in any of the directions [100], [110], or [111], or for hydrostatic pressure,[8,9] the A_1 states behave similarly for all the semiconductors: these states exhibit rather small pressure derivatives of magnitude ~ 1 meV/kbar, that may be either positive or negative (depending on the site of the impurity and the energy of the level), and are almost the same for all deep levels in the band gap. Hence, to a good approximation, the A_1

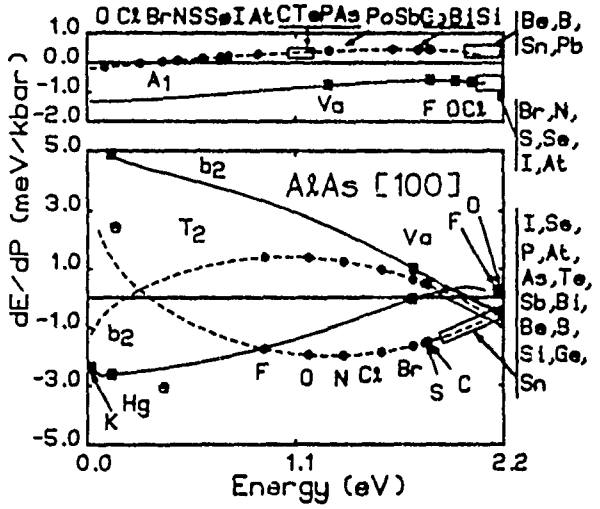


FIG. 4. Predicted pressure derivatives dE/dP (in meV/kbar) vs deep trap energy E (in eV) for s- and p-bonded substitutional point defects in AlAs with uniaxial stress along the [100] direction. The upper panel is for the A_1-symmetric (s-like) states and the lower panel is for the T_2-symmetric (p-like) states. The predictions for the cation-site defects are shown as dashed lines, the anion-site predictions by solid lines. The impurities predicted to produce deep levels of energy E are denoted by circles (cation site) and squares (anion site), although allowance should be made for these predictions being uncertain by a few tenths of eV. A box is used to denote several impurity levels lying close in energy (impurities corresponding to the box are indicated). Deep levels associated with the following impurities are predicted to lie in the conduction band but within the theoretical uncertainty of the band gap: A_1 cation site—Ga, In, and Tl; T_2 cation site—Pb, Ga, In, and Tl; A_1 anion site—C, Te, P, and Po; T_2 anion site—N and Cl. Similarly, in the valence band near its maximum, one might find the following: A_1 cation-site—F and vacancy (Va); T_2 cation site—vacancy, Hg, K, Na, Cd, and Li; A_1 anion site—K, Na, and Li; T_2 anion site—K, Na, Cd, Li, Zn, and Mg.

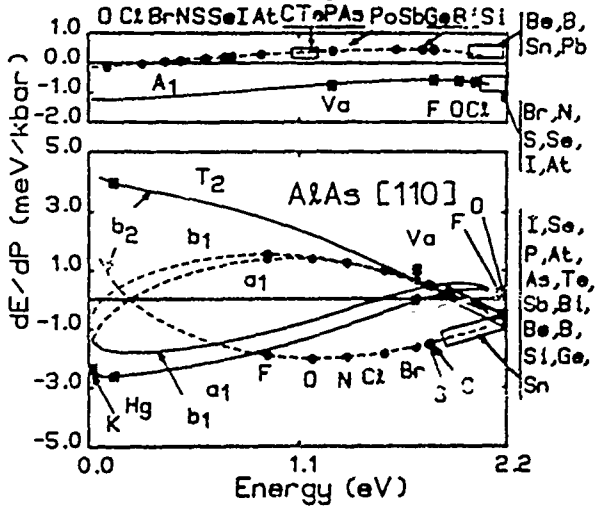


FIG. 5. Predicted pressure derivatives dE/dP (in meV/kbar) vs deep trap energy E (in eV) for s- and p-bonded substitutional point defects in AlAs with uniaxial stress along the [110] direction. The upper panel is for the A_1-symmetric (s-like) states and the lower panel is for the T_2-symmetric (p-like) states. The predictions for the cation-site defects are shown by dashed lines, the anion-site predictions by solid lines. The impurities predicted to produce deep levels of energy E are denoted by circles (cation site) and squares (anion site), although allowance should be made for these predictions being uncertain by a few tenths of eV. A box is used to denote several impurity levels lying close in energy (impurities corresponding to the box are indicated).

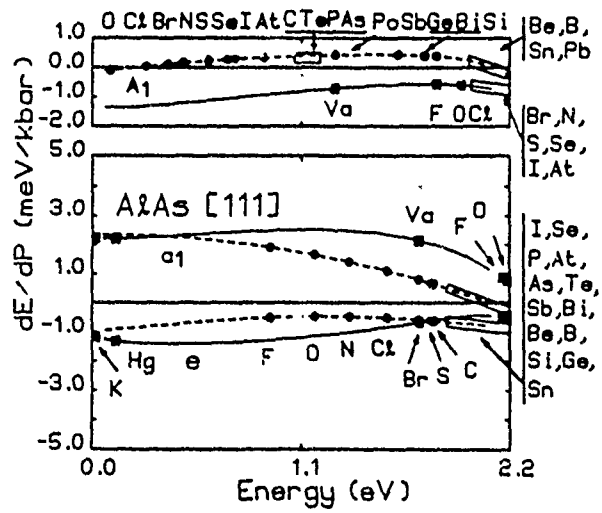


FIG. 6. Predicted pressure derivatives dE/dP (in meV/kbar) vs deep trap energy E (in eV) for s- and p-bonded substitutional point defects in AlAs with uniaxial stress along the [111] direction. The upper panel is for the A_1-symmetric (s-like) states and the lower panel is for the T_2-symmetric (p-like) states. The predictions for the cation-site defects are shown by dashed lines, the anion-site predictions by solid lines. The impurities predicted to produce deep levels of energy E are denoted by circles (cation site) and squares (anion site), although allowance should be made for these predictions being uncertain by a few tenths of eV. A box is used to denote several impurity levels lying close in energy (impurities corresponding to the box are indicated).

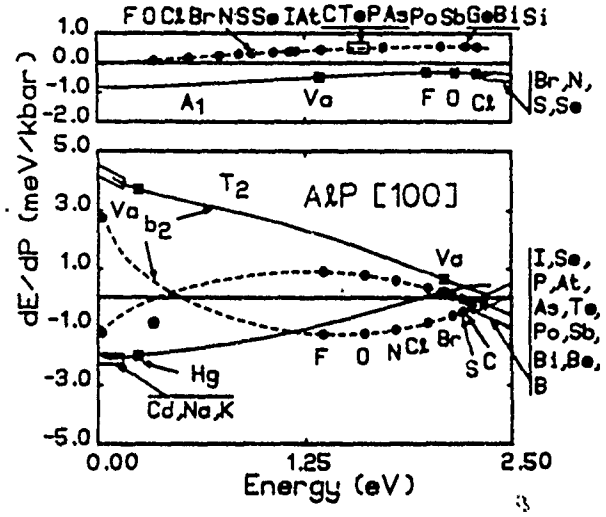


FIG. 7. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to AlP along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—Be, B, Sn, Pb, Ga, In, and Tl; T_2 cation site—Si, Ge, Sn, Pb, Ga, In, and Tl; A_1 anion site—I, At, C, Te, As, and Po; T_2 anion site—F, O, N, and Cl. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—vacancy, K, Na, and Li; T_2 cation site—Hg, K, Na, Cd, and Li; A_1 anion site—K, Na, and Li; T_2 anion site—K, Na, Cd, Li, Zn, and Mg.

deep levels in the gap appear to be almost "attached" to the average valence-band maximum:[18] when stress is applied, they move in energy an amount comparable (to within ±1 meV/kbar) with the shift of the average valence-band maximum.

In general, dE/dP for an A_1 state of a given deep-level energy E in the gap is more positive if the impurity responsible for that level occupies a cation site rather than an anion site. Hong *et al.*[9] have shown how this behavior can be understood for InSb: the pressure derivatives of deep levels in the gap of InSb, a material similar to AlAs, result from competition between the conduction-band states repelling the level downward in energy and the valence-band states pushing the level upward in energy. The valence band is anionlike in character and shifts

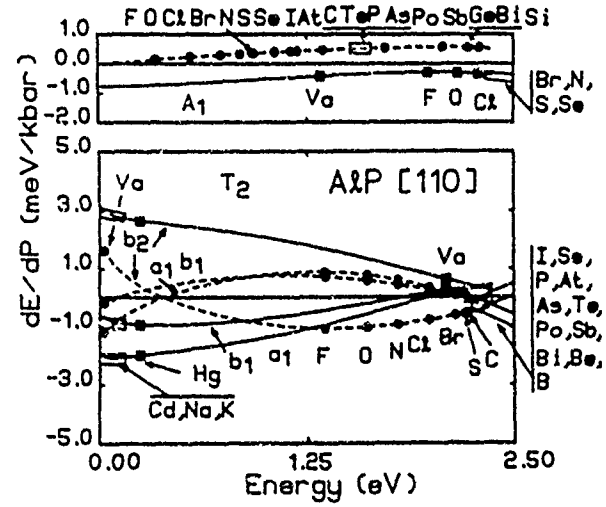


FIG. 8. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to AlP along the [110] direction as in Fig. 5.

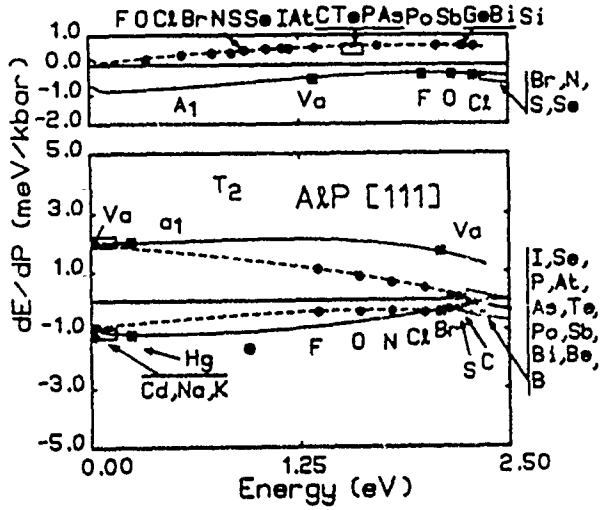


FIG. 9. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to AlP along the [111] direction as in Fig. 6.

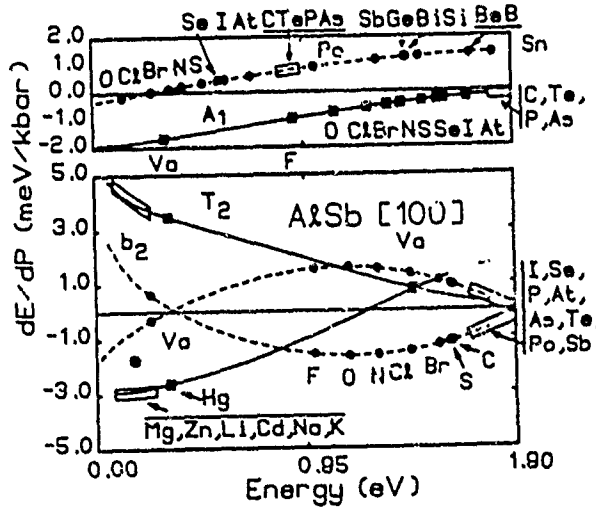


FIG. 10. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to AlSb along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie on the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—Pb, Ga, In, and Tl; T_2 cation site—Bi, Be, B, Si, Ge, Sn, Pb, Ga, In, and Tl; A_1 anion site—Po, Sb, Ge, Bi, Si, Be, and B; T_2 anion site—F, O, N, and Cl. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—F, vacancy, K, Na, and Li; T_2 cation site—Hg, K, Na, Cd, and Li; A_1 anion site—K, Na, and Li; T_2 anion site—Ga, In, and Tl.

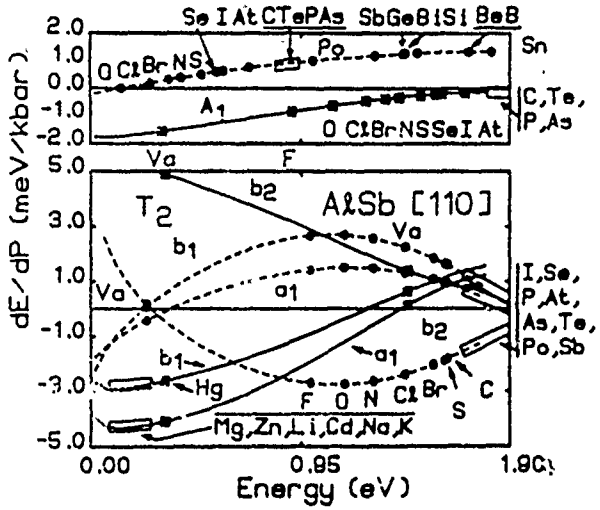


FIG. 11. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to AlSb along the [110] direction as in Fig. 5.

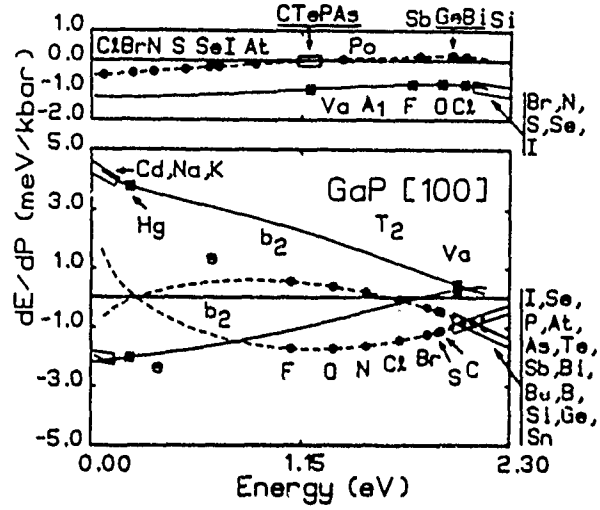


FIG. 13. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaP along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—Be, B, Sn, Pb, Al, In, and Tl; T_2 cation site—Pb, Al, In, Tl, Mg, and Zn; A_1 anion site—At, C, Te, As, Po, Sb, Ge, Bi, and Si; T_2 anion site—N, O, Cl, and Br. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—O, F, vacancy; T_2 cation site—vacancy, Hg, K, Na, Cd, and Li; A_1 anion site—K, Na, and Li; T_2 anion site— Li, Zn, and Mg.

to lower energies under hydrostatic pressure, while the conduction band is cationlike in character and shifts to higher energies. An impurity on an anion site with a deep level in the gap will therefore have an anionlike response to pressure: a more negative pressure derivative than a cation-site defect with the same deep-level energy.

C. T_2-derived states

Typically (but not always) dE/dP is several times larger for the T_2-derived states than for the A_1-derived states: In general, the results found for AlAs are similar both to those found previously for GaAs and GaP (Ref. 2) and to those found here for the other semiconductors.

1. [100] direction

The point-group symmetry of a substitutional impurity in a zinc-blende semiconductor is T_d for zero stress. For uniaxial stress applied in the [100] direction this symmetry is reduced to D_{2d}.[8,19] The s-like A_1 irreducible representation of T_d corresponds to an a_1 representation of D_{2d}.[20] The p-like T_2 representation reduces (see Fig. 1)

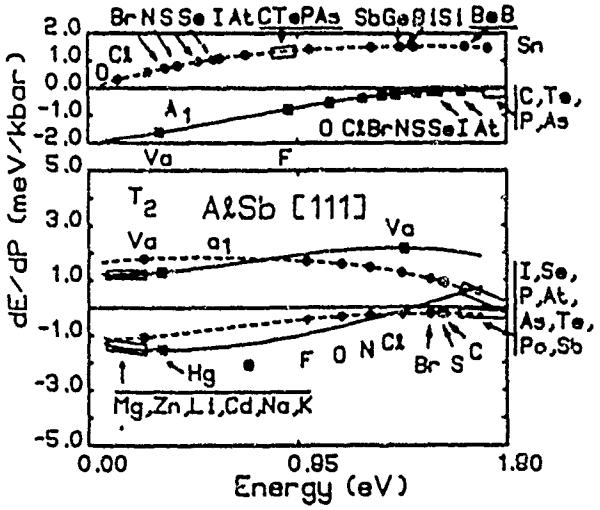


FIG. 12. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to AlSb along the [111] direction as in Fig. 6.

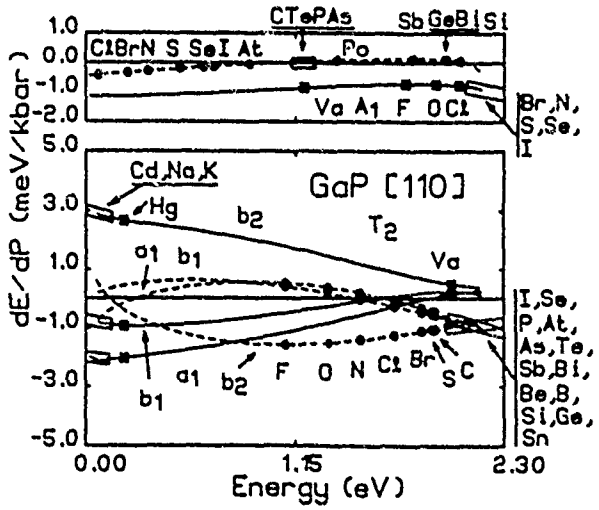


FIG. 14. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaP along the [110] direction as in Fig. 5.

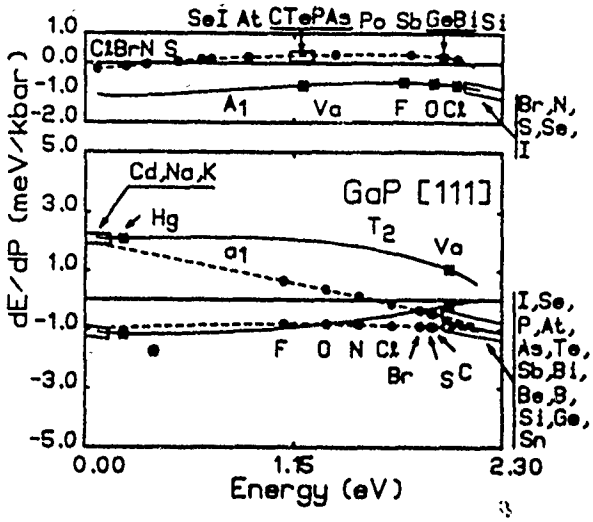


FIG. 15. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaP along the [111] direction as in Fig. 6.

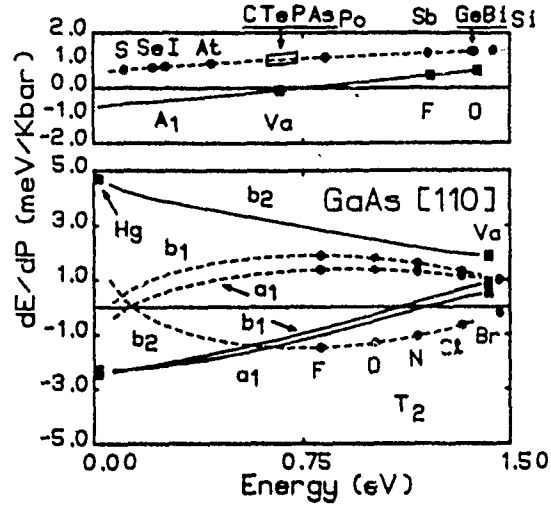


FIG. 17. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaAs along the [110] direction as in Fig. 5.

to a nondegenerate b_2 representation (a p-state polarized along the [100] direction) and a doubly degenerate e representation (p states polarized along the perpendicular directions to the [100]).[19] Typically the b_2 states, being polarized along the stress axis, are more sensitive to the uniaxial stress than the e states, and are also typically several times more sensitive to stress than the A_1-derived a_1 states. (This does not mean, however, that experiments finding one level with a larger derivative dE/dP than another can automatically assign the level with the large derivative to T_2 symmetry, because there are energies for which the A_1 derivative is larger; a more careful analysis is needed.)

Typically, but not always, the e level and the b_2 level split from one another, with one level rising and one level falling in energy. The b_2 level typically has a derivative larger in magnitude than the e level. This behavior is a consequence of the following physics: hydrostatic pressure physically is equivalent to stresses applied simultaneously in the [100], [010], and [001] directions. Thus these levels follow the simple sum rule[18]

$$\left|\frac{dE}{dP}(b_2)+2\frac{dE}{dP}(e)\right|=\frac{dE}{dP}(\text{hydrostatic})\,. \tag{7}$$

Concentrating on the fact that the uncertainty in the

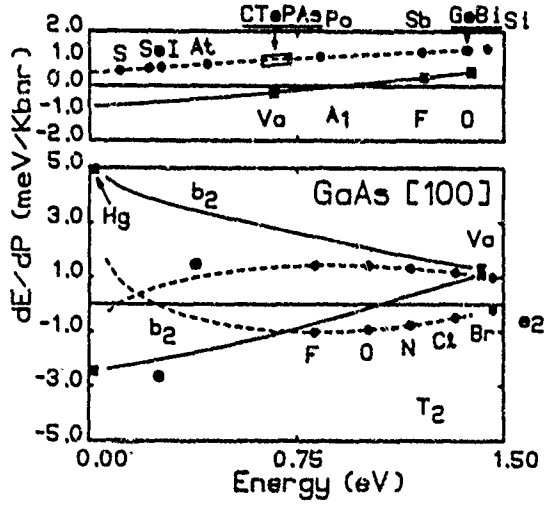


FIG. 16. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaAs along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty in the band gap: A_1 cation site—Be, B, Sn, Pb, Al, In, and Tl; T_2 cation site—S, C, I, Se, P, At, As, and Te; A_1 anion site—Cl, Br, and N; T_2 anion site—F, O, and N. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—N, Br, Cl, O, F; T_2 cation site—vacancy and Hg; A_1 anion site—K, Na, Li, Hg, and Cd; T_2 anion site—K, Na Cd, and Li.

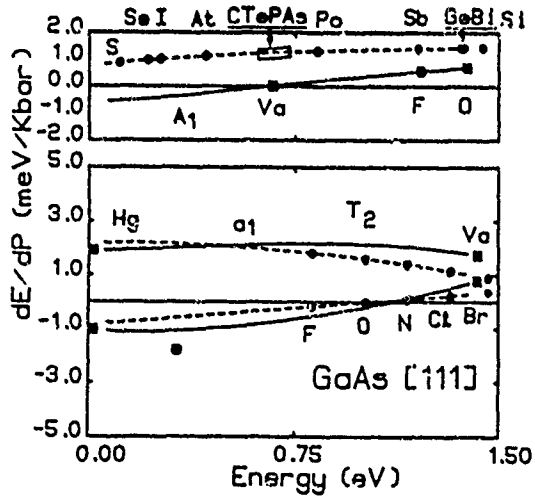


FIG. 18. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaAs along the [111] direction as in Fig. 6.

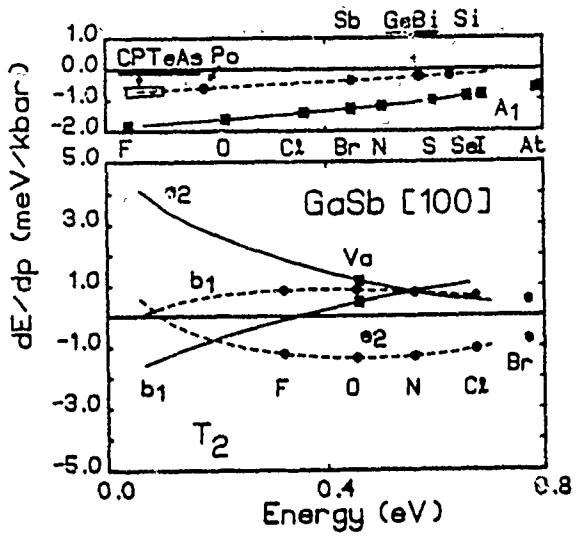


FIG. 19. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaSb along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—Be, B, Sn, Pb, Al, In, and Tl; T_2 cation site—S, C, I, Se, P, At, As, and Te; A_1 anion site—C, Te, P, As, and Po; T_2 anion site—F, O, N, and Cl. Similarly, in the valence band nears its maximum, one might find the following: A_1 cation site—At, I, Se, and S; T_2 cation site—vacancy, Hg, K, Na, and Cd; A_1 anion site—vacancy, K, Na, Li, Hg, and Cd; T_2 anion site—Hg, K, Na, and Cd.

theory is of order 1 meV/kbar (see the Appendix), one can see from Fig. 4 that measurements of the magnitude dE/dP for a deep level helps determine both the symmetry of the level and the site of the associated impurity, in cases such that the lines of Fig. 4 are separated by considerably more than this uncertainty. Even when the lines nearly coincide, for example, in the case of a level ~0.3 eV above the valence-band maximum in AlAs, a [100] stress would have the same effect on the T_2-derived cation-site b_2 and e states as the A_1-derived cation-site a_1 state. Thus, this impurity could be assigned to the cation site; however, since T_2 levels may not split very much for this impurity (see Fig. 4), the symmetry of the state could not be determined by application of [100] stress.

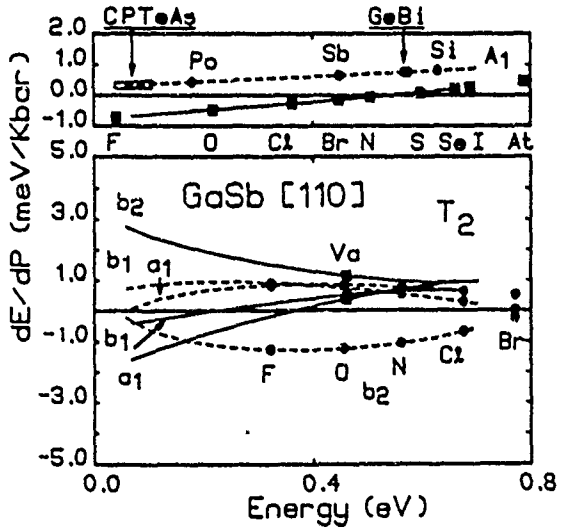


FIG. 20. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaSb along the [110] direction as in Fig. 5.

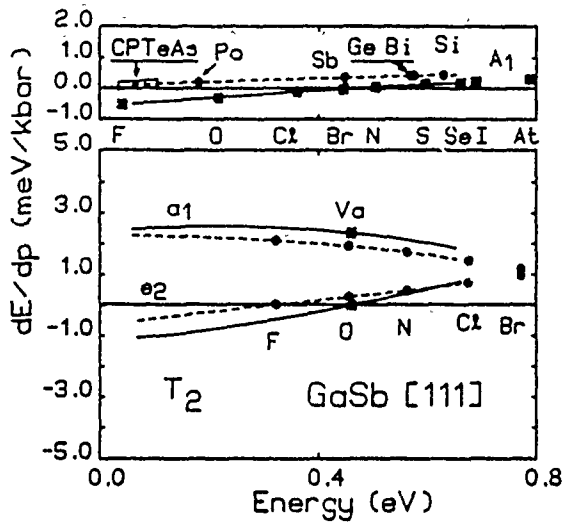


FIG. 21. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to GaSb along the [111] direction as in Fig. 6.

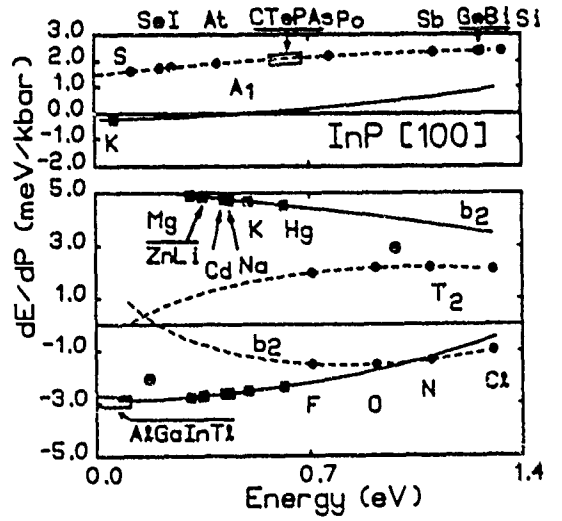


FIG. 22. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InP along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—Be, B, Sn, and Pb; T_2 cation site—Br, S, C, and I; A_1 anion site—vacancy and F; T_2 anion site—vacancy and F. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—O, Cl, Br, and N; T_2 cation site—vacancy and Hg; A_1 anion site—Zn, Cd, Hg, Li, and Na; T_2 anion site—Pb, Sn, Ge, and Si.

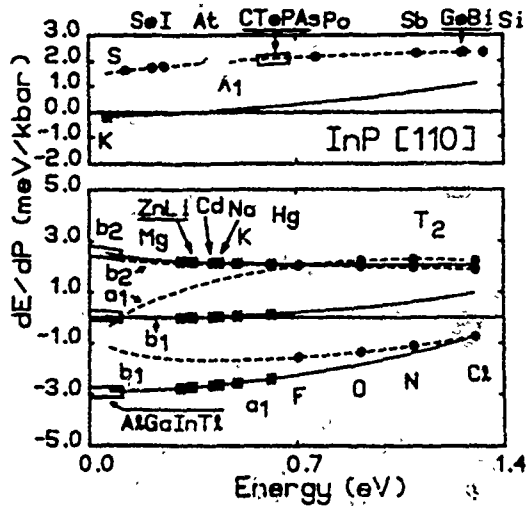


FIG. 23. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InP along the [110] direction as in Fig. 5.

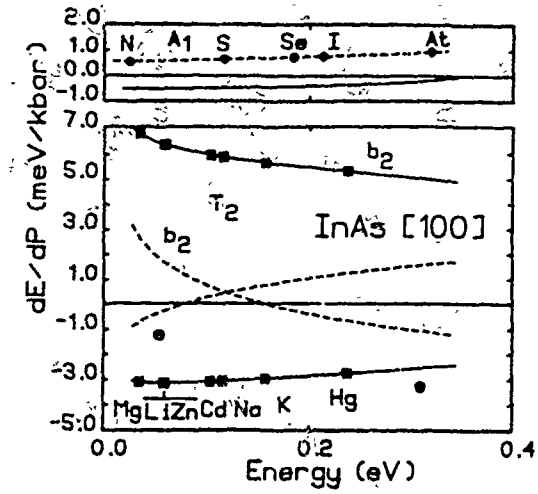


FIG. 25. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InAs along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—C, Te, P, and Sb; T_2 cation site—F, O, and N; A_1 anion site—F, O, and N; T_2 anion site—vacancy and F. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—Br, Cl, O and F; T_2 cation site—vacancy and Hg; A_1 anion site—vacancy and K; T_2 anion site—Ga, Al, In, and Tl.

2. [110] direction

For stress applied in the [110] direction, the orbital degeneracy of the T_2 levels is completely removed (in contrast to the cases of stress in the [100] and [111] directions, which leave one level doubly degenerate). The point group of a [110] stressed zinc-blende semiconductor is C_{2v},[20] and the p-like T_2 levels split (see Fig. 2) into levels of symmetry a_1 (a p state polarized perpendicular to the [110] axis and parallel to the [001] axis), b_1 (a p state polarized along the $[1\bar{1}0]$ direction), and b_2 (a p state polarized along the [110]-stress direction). The curves dE/dP versus E for [110] uniaxial stress are similar to those for [100] stress, with one important exception: the [110] stress completely splits the T_2 levels and removes their degeneracies. This splitting can be used as a signature of a T_2 level; thus [110]-uniaxial-stress measurements are likely to provide the most information about the symmetry of a deep level.[19]

The similarity between the [110]-stress derivatives and the [100]-stress derivatives is due to the similarity of the strains caused by the two types of stress. Applied stress along either direction causes two neighbors of the impurity to shift closer to each other with angular distortions and some compression of the bonds (see Figs. 1 and 2).

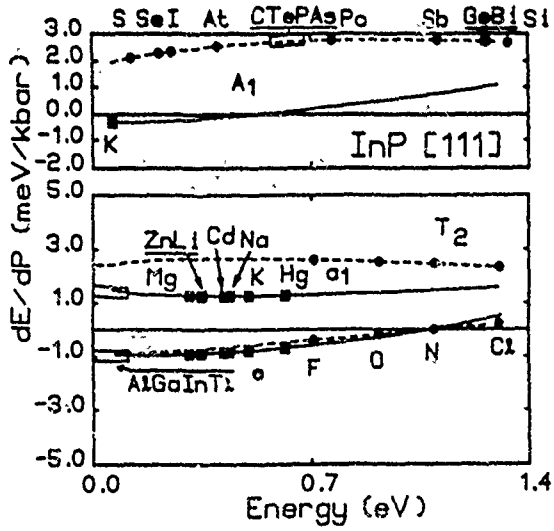


FIG. 24. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InP along the [111] direction as in Fig. 6.

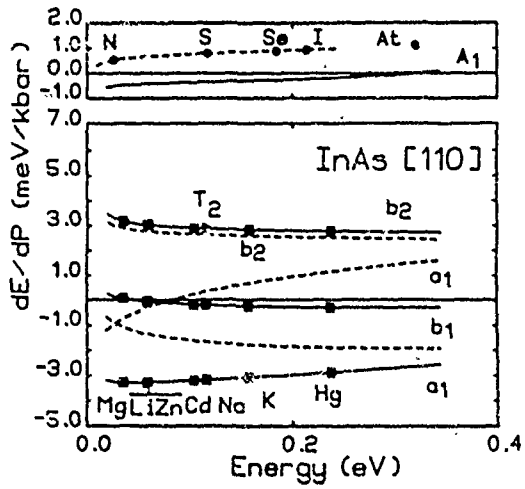


FIG. 26. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InAs along the [110] direction as in Fig. 5.

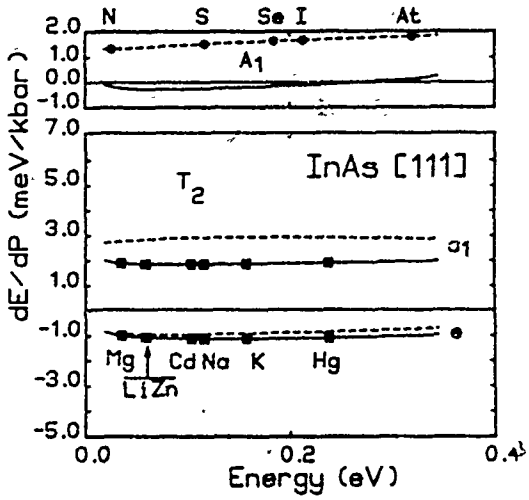


FIG. 27. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InAs along the [111] direction as in Fig. 6.

The states directed along the stress, the b_2 state in the case of [110] applied stress and the b_2 state in the case [100] applied stress, have quite similar stress derivatives, as do the states directed perpendicular to the applied stress.

3. *[111] direction*

For stress along the [111] direction, a T_2 level splits into a doubly degenerate e state (p states polarized perpendicular to the [111] direction) and an a_1 state (a p state oriented along the direction of the applied stress). The point group is C_{3v}.[12] The situation for the pressure derivatives when the uniaxial stress is applied along the [111] direction is different from that of the [100] or [110] directions. Within the uncertainty of the theory, the curves of Fig. 6 never intersect one another (for impurities on the same site) and so all T_2 levels in the gap are predicted to split under [111] stress—for either the cation- or anion-site impurities. Stress applied in the [111] direction can always distinguish between an s-like A_1 impurity state and a p-like T_2 impurity state,[19] because the T_2 states have sizable splittings. However, such a stress can almost never give the site of the associated impurity, because the states of the same symmetry but associated with different sites have similar values of

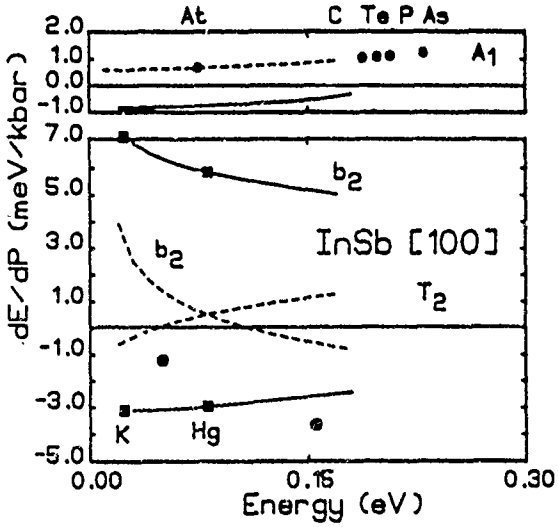


FIG. 28. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InSb along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1 cation site—Po, Sb, Ge, Bi, and Si; T_2 cation site—vacancy and F; A_1 anion site—vacancy and F; T_2 anion site—vacancy and F. Similarly, in the valence band near its maximum, one might find the following: A_1 cation site—N, S, Se, and I; T_2 cation site—Hg, K, Na, and Cd; A_1 anion site—Li, Na, and K; T_2 anion site—Zn, Li, Cd, and Na.

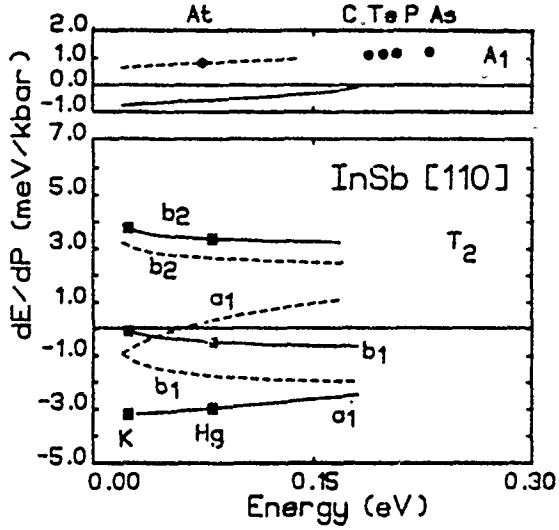


FIG. 29. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InSb along the [110] direction as in Fig. 5.

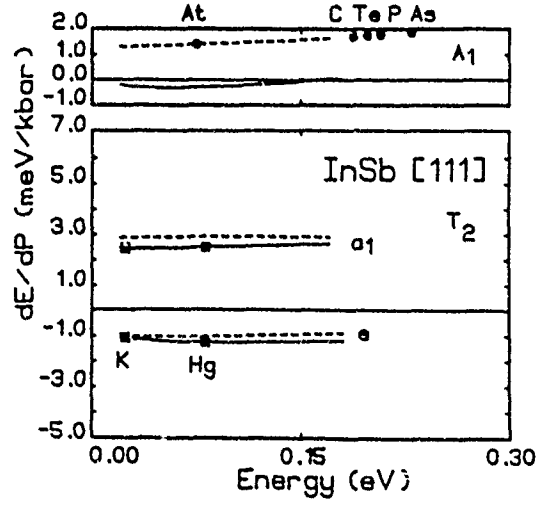


FIG. 30. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to InSb along the [111] direction as in Fig. 6.

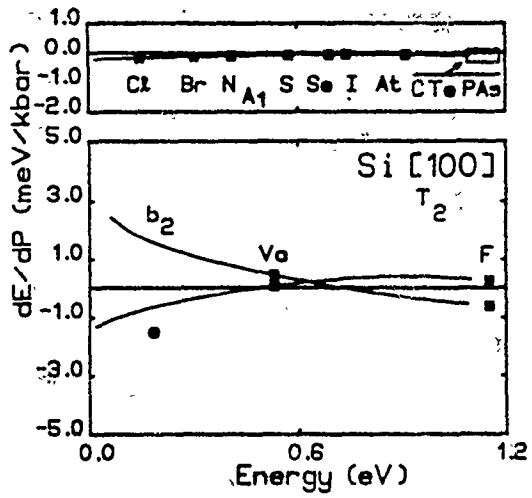


FIG. 31. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to Si along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1—Po, Sb, Ge, and Bi; T_2—O, N, and Cl. Similarly in the valence band near its maximum, one might find the following: A_1—F and O; T_2—K and Hg.

dE/dP to within 1 meV/kbar, the uncertainty of the theory.

The [111] stress derivatives dE/dP are quite different from the derivatives for stress applied in the [100] or [110] directions. The strains between nearest neighbors are different under [111] applied stress: the stress directly compresses one of the nearest-neighbor bonds and the three remaining bonds have angular distortions and expand slightly (see Fig. 3). Because, to a good approximation, one bond is compressed while the other three are only bent, a large splitting results between the a_1 states

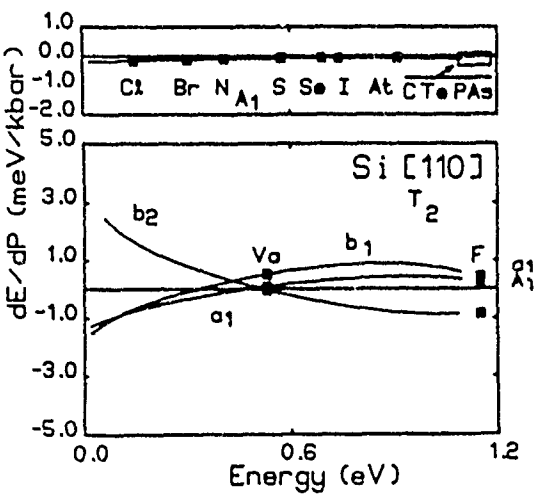


FIG. 32. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to Si along the [110] direction as in Fig. 5.

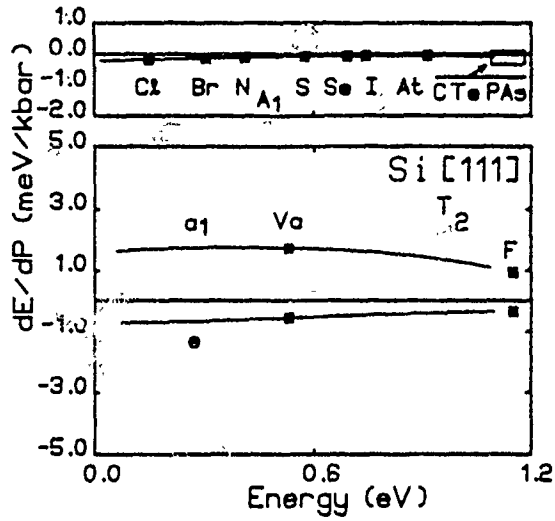


FIG. 33. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to Si along the [111] direction as in Fig. 6.

directed along the stress and the e states perpendicular to the stress.

IV. OTHER MATERIALS

The predicted [100]-, [110]-, and [111]-stress dependences of deep substitutional impurity levels in the semiconductors AlP, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb, Si, and Ge are displayed in Figs. 7–36.[21] In general, the trends found for AlAs are also found in the defect levels for the remaining zinc-blende III-V compound semiconductors.

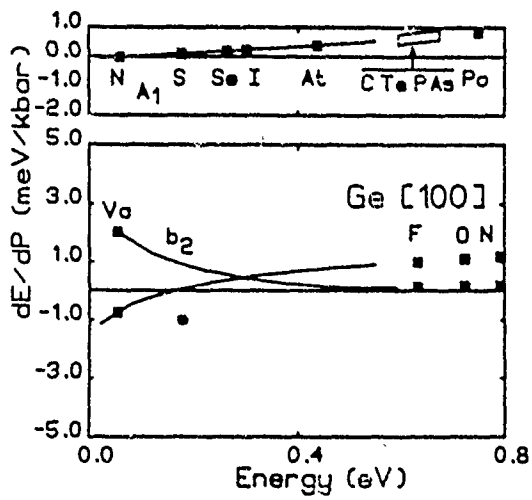


FIG 34. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to Ge along the [100] direction as in Fig. 4. Deep levels associated with the following impurities are predicted to lie in the conduction band, but within the theoretical uncertainty of the band gap: A_1—Sb, Bi, and Si; T_2—Cl, Br, S, C, and I. Similarly, in the valence band near its maximum, one might find the following: A_1—O, Cl, Br; T_2—K and Hg.

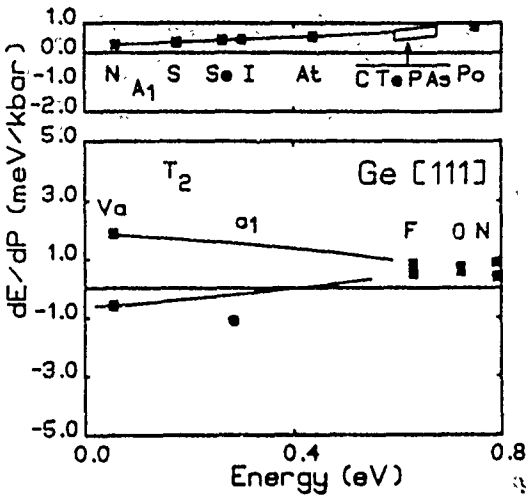


FIG. 35. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to Ge along the [110] direction as in Fig. 5.

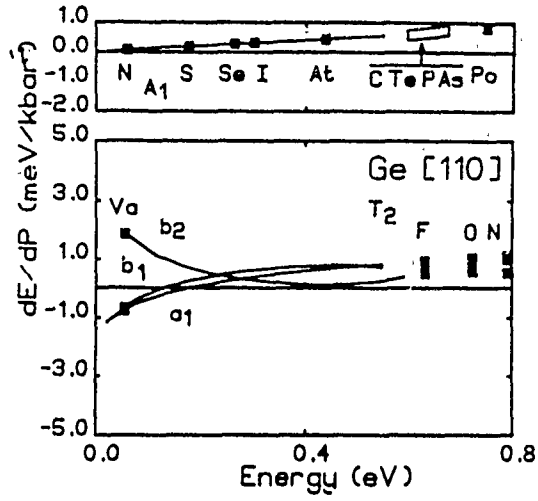


FIG. 36. Predicted pressure derivatives dE/dP vs energy E for uniaxial stress applied to Ge along the [111] direction as in Fig. 6.

V. CONCLUSION

We have examined the uniaxial-stress dependences of deep substitutional impurity levels in the semiconductors AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, InSb, Si, and Ge, for stress applied in the [100], [110], and [111] directions. Experiments involving stress in the [110] direction should be superior for determining the symmetry of a deep level and the site of its parent defect: under [110] stress the p-like T_2 levels should always exhibit significant splitting, with two sublevels most often moving to higher energy if the defect occupies the cation site or with two levels normally falling to lower energy if the defect occupies the anion site.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support, and we thank R.-D. Hong for stimulating conversations.

APPENDIX: THEORETICAL UNCERTAINTY

We estimate the uncertainty in the theory for dE/dP to be ~ 1 meV/kbar. The primary source of uncertainty can be traced to the Green's function, and hence to the Hamiltonian matrix and its dependences on bond lengths and angles. The contribution by the bond-length dependence to the uncertainty has been thoroughly studied by Ren *et al.*[8] and Hong *et al.*,[9] who varied the Hamiltonian matrix elements over the entire range of reasonable values and computed dE/dp for hydrostatic pressure p. They deduced an uncertainty of 0.5 meV/kbar, an estimate that has proven to be rather conservative in analyses of data. The angular dependences of the matrix elements are fixed by symmetry, and so their contributions to the uncertainty come exclusively from the fact that the physics of a particular nearest-neighbor matrix element, for example, inadvertently contains some second-neighbor physics. Such contributions also occur with the hydrostatic pressure dependences, and so should be considerably smaller than 0.5 meV/kbar. An upper bound on the combined uncertainty is thus ~ 1 meV/kbar.

---

*Present address: Amoco Research Center (F-4), P.O. Box 400, Naperville, IL 60566.

†Permanent address: University of Science and Technology of China, Hefei, People's Republic of China.

[1]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980). See also W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[2]S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).

[3]For a review of EXAFS, see B. A. Bunker, J. Vac. Sci. Technol. A 5, 3003 (1987).

[4]D. Feher, Phys. Rev. 114, 1219 (1959).

[5]J. M. Luttinger and W. Kohn, Phys. Rev. 97, 969 (1955).

[6]D. J. Wolford, and J. A. Bradley, Solid State Commun. 53, 1069 (1985).

[7]D. J. Wolford, J. A. Bradley, K. Fry, J. Thompson, and H. E. King, in *Gallium Arsenide and Related Compounds*, Inst. Phys. Conf. Ser. No. 65, edited by G. E. Stillman (IOP, London, 1982), p. 477.

[8]S. Y. Ren, J. D. Dow, and D. J. Wolford, Phys. Rev. B 25, 7661 (1982).

[9]R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Phys. Rev. B 38, 12 549 (1988).

[10]P. Vogl, H. P. Hjalmarson, and John D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[11]See, for example, L. D. Landau, and E. M. Lifshitz, *Theory of Elasticity* (Pergamon, New York, 1986), pp. 32–37.

[12]L. Kleinman, Phys. Rev. **128**, 2641 (1962).

[13]The other components of Eqs. (1) and (3) are found using cyclic permutations of the coordinate labels in the respective equations ($x \rightarrow y \rightarrow z \rightarrow x$).

[14]A. Següller and H. R. Neyer, Phys. Kondens. Mater. **4**, 63 (1965).

[15]Values for Kleinman's parameter have been published (Ref. 12) for some of the materials studied in this paper. We use the average value of these parameters, 0.6, for all the materials.

[16]J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

[17]W. A. Harrison and S. Ciraci, Phys. Rev. B **10**, 1516 (1974); W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

[18]The stress derivatives are the change in energy of a deep impurity level with respect to a reference energy, for an applied unit stress. We do not take the valence-band maximum as the reference energy as is perhaps implied in Figs. 4–36, but rather the average value of the three stressed p-like levels which form the valence-band maximum for zero applied stress. With this reference energy, the average value of the stress derivatives for the T_2 levels, under stress applied in the [100] direction, equals one-third of the value of the derivative for the T_2 levels under hydrostatic pressure (see text). Under hydrostatic pressure an atom at the zero stress position $\mathbf{r}=(x_0,y_0,x_0)$ moves to $\mathbf{r}=(x_0,y_0,x_0)\{1-[P/(c_{11}+2c_{12})]\}$. Similar rules for stress in the [110] and [111] directions cannot be made, because [110] and [111] stress involves shear stresses not present in hydrostatic pressure.

[19]The theory does not account for d-like impurity states which also split under uniaxial stress. In zinc-blende and diamond materials, d states are split by the crystal field into three T_2 and two E states. Under [100] stress the T_2 states split into one b_2 and two e states and the E states do not split. Under [110] stress the T_2 states split into one a_1, one b_1, and one b_2 state and the E states split into one a_1 and one a_2 state. Under [111] stress the T_2 states split into one a_1 and two e states and the E states do not split.

[20]A. A. Kaplyanskii, Opt. Spektrosk. **16**, 1031 (1964) [Opt. Spectrosc. (USSR) **16**, 557 (1964)].

[21]A theory for the effects of uniaxial stress along the [100] directions on deep levels in GaAs and GaP was published previously, and the results are included here for completeness. See, S.-Y. Ren, D. Mao, and M. Li, Chin. Phys. Lett. **3**, 313 (1986).

Special points for superlattices and strained bulk semiconductors

Shang Yuan Ren
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
*and Department of Physics, University of Science and Technology of China, Hefei, People's Republic of China**

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 17 December 1987)

Special k points for performing integrals over the Brillouin zone of [001] and [111] superlattices are obtained and discussed. If the superlattice period is taken properly, the number of special points required in order to reach suitable convergence can be greatly reduced. Twelve and ten special points will give the same accuracy for $N\times M$ [001] and [111] superlattices, respectively, as Chadi's and Cohen's ten special points for bulk semiconductors, provided we have $N+M=4n$, with n an integer. These special points can also be used to calculate the corresponding integrals for strained bulk semiconductors.

I. INTRODUCTION

In many theoretical investigations involving the electronic structures of solids, one often needs to calculate integrals over the first Brillouin zone. Baldereschi[1] and Chadi and Cohen[2] suggested that such integrations can be accurately approximated by summing over a rather small number of special k points in the Brillouin zone, with different weights for each point. (Elaborations of the special point method have been given by Monkhorst, Pack, Chadi, and Cunningham.[3,4]) For example, ten special points give very satisfactory results for the Green's functions of bulk cubic semiconductors.[5]

In this paper we extend the special points method to [001] and [111] lattice-matched $N\times M$ superlattice such as $(GaAs)_N(AlAs)_M$, determining the special points for cases such that $N+M=4n$, where N and M are the numbers of two-atom-thick layers of each slab of a superlattice (e.g., $Na_L/2$ and $Ma_L/2$ are the thicknesses of the GaAs and AlAs slabs in a [001] GaAs-AlAs superlattice period, where a_L is the lattice constant of either bulk material), and n is an integer. The number of special points needed to obtain reasonable accuracy of such integrals can be greatly reduced if such a requirement is satisfied. In this paper we assume that the two components of the superlattice are perfectly lattice matched. For [001] superlattices the three-dimensional lattice translation vectors can be taken to be the following: $(a_L/2)(1,1,0)$, $(a_L/2)(1,-1,0)$, and $(a_L/2)(0,0,N+M)$ for $N+M$ even, or $(a_L/2)(1,1,0)$, $(a_L/2)(1,-1,0)$, and $(a_L/2)(0,1,N+M)$ for $N+M$ odd. For [111] superlattices the corresponding translation vectors are taken to be $(a_L/2)(1,-1,0)$, $(a_L/2)(0,1,-1)$, and $(a_L/2)(N+M, N+M,0)$. We choose the unit of wave vector k and the reciprocal lattice vectors G as $2\pi/a_L$, and k_1, k_2, and k_3 are the three components of the reduced wave vector k in the x, y, and z directions

$$\mathbf{k}=(k_1,k_2,k_3)(2\pi/a_L)\ .$$

Since the theoretical formalism is the same for any superlattice consisting of cubic semiconductors, for simplicity of presentation, we discuss only zinc-blende GaAs/AlAs superlattices.

II. GENERAL APPROACH

We follow the general approach of Chadi and Cohen for generating special points. They begin with one or more wave vectors and, by subjecting these wave vectors to symmetry operations, generate the special points. For zinc-blende crystals with a face-centered-cubic Bravais lattice, they generated ten special points k, based on the starting points $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, $(\frac{1}{4},\frac{1}{4},\frac{1}{4})$, and $(\frac{1}{5},\frac{1}{5},\frac{1}{5})$. We see the equivalent special points for superlattices composed of zinc-blende layers.

In generating the special points for zinc-blende materials, Chadi and Cohen actually used three types of symmetry: (i) time-reversal invariance (k and $-$k are equivalent); (ii) point-group symmetry (k and Tk are equivalent, where T is an element of the point group T_d for zinc-blende structures[6]); and (iii) translational symmetry (k and k+G are equivalent, where G is a reciprocal-lattice vector).

For a superlattice, time-reversal invariance still applies, but the point-group symmetry is lower (C_{2v} for a [001] superlattice and C_{3v} for a [111] superlattice, instead of T_d for the bulk), and the translational symmetry is different: Certainly the reciprocal-lattice vectors in the direction of growth are different and some reciprocal-lattice vectors in perpendicular directions might be changed also for some superlattices.

At first glance, the reduced symmetry appears to greatly increase the number of special points needed to obtain the same accuracy as obtained by the ten special points of Chadi and Cohen. For example, because of the lower symmetry of a [001] superlattice, its starting points corresponding to the three bulk generators $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, $(\frac{1}{4},\frac{1}{4},\frac{1}{4})$,

 ©1988 The American Physical Society

and $(\frac{1}{8},\frac{1}{8},\frac{1}{8})$ are $(\pm\frac{1}{2},\pm\frac{1}{2},\pm\frac{1}{2})$, $(\pm\frac{1}{4},\pm\frac{1}{4},\pm\frac{1}{4})$, and $(\pm\frac{1}{8},\pm\frac{1}{8},\pm\frac{1}{8})$, and produce, in principle, $8^3=512$ special points instead of the ten points of Chadi and Cohen. Here we show how the number of special points can be reduced from 512 to 12 for the [001] superlattice and to 10 for the [111] superlattice—provided one restricts one's attention to superlattices of special periods, namely $N\times M$ superlattices, such that $N+M=4n$, with n an integer.

III. [001] SUPERLATTICES AND STRAINS

A. Superlattices

For [001] superlattices, the C_{2v} symmetry dictates that the starting points corresponding to the bulk $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ generator be $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, $(\frac{1}{2},-\frac{1}{2},\frac{1}{2})$, and $(\frac{1}{2},\frac{1}{2},-\frac{1}{2})$, with weights $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$, respectively. Combining either $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, or $(\frac{1}{2},\frac{1}{2},-\frac{1}{2})$ with $(\pm\frac{1}{4},\pm\frac{1}{4},\pm\frac{1}{4})$ for an arbitrary [001] superlattice, using the method of Chadi and Cohen, we find six special points: Two have weights of $\frac{1}{16}$ and four have weights of $\frac{1}{32}$. Recombining these six points with $(\pm\frac{1}{8},\pm\frac{1}{8},\pm\frac{1}{8})$ will give 40 points: 24 with weight $\frac{1}{128}$ and 16 with weight $\frac{1}{256}$. Repeating this procedure, but starting with $(\frac{1}{2},-\frac{1}{2},\frac{1}{2})$, yields 64 points, each with weight $\frac{1}{128}$. This gives a total of $40+64+40=144$ superlattice special points corresponding to the 10 special points of bulk zinc blende. It is often impractical to use so many special points.

For [001] superlattices with special periods, the 144 special points can be reduced considerably by invoking translational symmetry. For example, consider a GaAs-AlAs $N_{\mathrm{GaAs}}\times M_{\mathrm{AlAs}}$ [001] superlattice such that $N_{\mathrm{GaAs}}+M_{\mathrm{AlAs}}=4n$, where n is an integer. (Here, for example, N_{GaAs} denotes the number of GaAs molecular layers per slab of GaAs. Hence a 1×1 superlattice consists of alternating layers of GaAs and AlAs, and a 2×3 superlattice alternates two GaAs layers and three AlAs layers.) In this case, the reciprocal-lattice vectors of the superlattice are $\mathrm{G}_1=(1,1,0)$, $\mathrm{G}_2=(1,-1,0)$, and $\mathrm{G}_3=(0,0,1/2n)$. Because of the symmetry of this special period, any two of the 144 special points (k_1, k_2, k_3) that have the same values of k_1 and k_2 coalesce into a single special point—reducing the number of distinct points to 20. Examination of the remaining symmetries reduces this set to 12 special points k, each with weight α, denoted by (k,α):

$(\frac{7}{8},\frac{3}{8},\frac{1}{8};\frac{1}{8})$, $(\frac{5}{8},\frac{3}{8},\frac{1}{8};\frac{1}{16})$, $(\frac{7}{8},\frac{1}{8},\frac{1}{8};\frac{1}{16})$,

$(\frac{3}{8},\frac{1}{8},\frac{1}{8};\frac{1}{8})$, $(\frac{3}{8},\frac{3}{8},\frac{1}{8};\frac{1}{16})$, $(\frac{1}{8},\frac{1}{8},\frac{1}{8};\frac{1}{16})$,

$(-\frac{7}{8},\frac{3}{8},-\frac{1}{8};\frac{1}{8})$, $(-\frac{5}{8},\frac{3}{8},-\frac{1}{8};\frac{1}{16})$,

$(-\frac{7}{8},\frac{1}{8},-\frac{1}{8};\frac{1}{16})$, $(-\frac{3}{8},\frac{1}{8},-\frac{1}{8};\frac{1}{8})$,

$(-\frac{3}{8},\frac{3}{8},-\frac{1}{8};\frac{1}{16})$, and $(-\frac{1}{8},\frac{1}{8},-\frac{1}{8};\frac{1}{16})$.

These special points have been used to compute the Green's functions of superlattices, for energies in the fundamental band gaps.[7]

B. Strained bulk zinc blende

One can consider bulk GaAs to be a degenerate form of a GaAs-GaAs [001] superlattice. In this case, the superlattice has D_{2d} symmetry rather than C_{2v} symmetry (because the former AlAs of the GaAs-AlAs superlattice is now GaAs). Either the first six or the last six of the special points above (with their weights doubled) could be used for the D_{2d} symmetry, leading to results with the same accuracy as the ten special points of Chadi and Cohen. Of course, by viewing bulk GaAs as a 2×2 GaAs-GaAs superlattice, the Hamiltonian matrix is four times as large as for bulk GaAs, and so evaluating six special points for sums involving functions of the larger Hamiltonian will be more laborious than evaluating ten for bulk GaAs. Therefore it does not make sense to treat bulk GaAs as a GaAs-GaAs superlattice, but *bulk GaAs strained along the [001] direction has D_{2d} symmetry*, and so either the first six or the last six of the above 12 special points (with their weights doubled) can be used to evaluate integrals over the Brillouin zone for [001]-strained bulk GaAs.

IV. [111] SUPERLATTICES AND STRAINS

A. Superlattices

For [111] superlattices, the point group is C_{3v}, and the generators of special points are $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ and $(-\frac{1}{2},\frac{1}{2},\frac{1}{2})$, with weights of $\frac{1}{4}$ and $\frac{1}{4}$, respectively. Combining $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ with $(\pm\frac{1}{4},\pm\frac{1}{4},\pm\frac{1}{4})$ and then $(\pm\frac{1}{8},\pm\frac{1}{8},\pm\frac{1}{8})$, we obtain ten distinct special points, with weights of $\frac{3}{64}$, $\frac{1}{128}$ or $\frac{1}{256}$. Starting with $(-\frac{1}{2},\frac{1}{2},\frac{1}{2})$, we obtain 40 special points with weights of $\frac{1}{128}$ or $\frac{3}{256}$, bringing the total number of special points to 50. This set can be reduced by considering $N_{\mathrm{GaAs}}+M_{\mathrm{AlAs}}=4n$, in which case $(1/4n,1/4n,1/4n)$ is a reciprocal-lattice vector. By adding this reciprocal-lattice vector, or its negative, to each of the 50 general special points, we reduce the number of distinct points to 20: eight lie in a plane on the superlattice Brillouin-zone boundary which is perpendicular to the [111] direction and passing through the point $(\frac{1}{8},\frac{1}{8},\frac{1}{8})$, and 12 lie on a parallel plane within the Brillouin zone. Using translational symmetry with reciprocal-lattice vectors $(1,-1,0)$, $(0,1,-1)$, and $(-1,0,1)$ together with C_{3v} point-group symmetry, these 20 points reduce to the following ten distinct special points k with weights α, $(\mathrm{k};\alpha)$:

$(\frac{3}{8},\frac{1}{8},-\frac{1}{8};\frac{3}{32})$, $(\frac{5}{8},-\frac{1}{8},-\frac{1}{8};\frac{3}{32})$, $(\frac{1}{8},\frac{1}{8},\frac{1}{8};\frac{1}{64})$,

$(\frac{7}{8},-\frac{3}{8},-\frac{1}{8};\frac{3}{16})$, $(\frac{9}{8},-\frac{3}{8},-\frac{3}{8};\frac{3}{32})$,

$(\frac{9}{8},-\frac{5}{8},-\frac{1}{8};\frac{3}{16})$, $(\frac{7}{8},-\frac{5}{8},\frac{1}{8};\frac{3}{32})$,

$(\frac{5}{8},-\frac{3}{8},\frac{1}{8};\frac{3}{32})$, $(\frac{1}{8},\frac{1}{8},-\frac{1}{8};\frac{3}{32})$, and $(-\frac{5}{8},\frac{3}{8},\frac{3}{8};\frac{3}{64})$.

B. Strain

The GaAs-GaAs [111] superlattice has C_{3v} symmetry, and so the above ten special points can be used for treating Brillouin-zone sums for [111]-strained bulk GaAs.

V. CONCLUSION

Special points for superlattices can be computed using the method of Chadi and Cohen. By limiting oneself to special-period superlattices, such as those for which $N_{GaAs}+M_{AlAs}=4n$, one finds that relatively few special points will provide accurate integrals over the superlattice Brillouin zone. We have presented 12 special points for $N\times M$ [001] superlattices and ten for $N\times M$ [111] superlattices, valid for $N+M=4n$, where n is an integer. However, our general approach can be used for generating special points that will produce more accuracy or for superlattices that do not satisfy the condition $N+M=4n$. These same special points obtained for $N+M=4n$ may be used for Brillouin-zone sums for strained bulk zinc-blende material.

ACKNOWLEDGMENT

We are grateful to the U.S. Office of Naval Research for their generous support of this research (Contract No. N00014-84-K-0352).

---

*Permanent address.

[1]A. Baldereschi, Phys. Rev. B 7, 5212 (1973).

[2]D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).

[3]An alternate approach for determining special points has been given by Monkhorst and Pack. See, for example, H. J. Monkhorst and J. D. Pack, Phys. Rev. B 13, 5188 (1976); D. J. Chadi, *ibid.* 16, 1746 (1977); J. D. Pack and H. J. Monkhorst, *ibid.* 16, 1748 (1977).

[4]Special points for the two-dimensional Brillouin zone have been determined by S. L. Cunningham, Phys. Rev. B 10, 4988 (1974).

[5]See, for example, S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).

[6]We use the notation of M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).

[7]See, for example, S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B (to be published).

المجلة العربية للعلوم والهندسة

SULFUR NEAREST-NEIGHBOR DEFECT PAIRS IN Si

Gyungock Kim, John D. Dow*

and

Seongbok Lee†

Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556, U.S.A.

THE ARABIAN JOURNAL FOR SCIENCE AND ENGINEERING

REPRINT

Published by KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS, DHAHRAN 31261, SAUDI ARABIA

SULFUR NEAREST-NEIGHBOR DEFECT PAIRS IN Si

Gyungock Kim, John D. Dow*

and

Seongbok Lee†

Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556, U.S.A.

الخلاصـــة :

نطور في هذا البحث منهجاً رياضياً لحساب انفصام وضع الشحنة في المستويات العميقة والمرتبطة بأزواج الخلل في شبه الموصلات . ونطبق هذا المنهج على أزواج الكبريت التبادلية ذات القرب الأكثر في بلورات السيليكون . ونستنتج ، من حساب متوافق ضمنياً ، أن الانفصام بين المستويات العميقة للذرات المتجاورة الأكثر قرباً (كبريت ، كبريت) $(S,S)^0$ و (كبريت ، كبريت) + $(S,S)^+$ هو (٠,١٩) الكترون فولط ev وذلك بتطابق جيد مع القيمة التجريبية وهي (٠,١٨) الكترون فولط ev . وكذلك تتطابق نسب المركبات التنسورية tensor العالية الدقة مع القيم التجريبية المتوفرة . وبهذا تدعم نتائجنا نظرية (النصف ربط) في المستويات العميقة للمركبات المزدوجة للنحاس مع الكبريت .

ABSTRACT

A formalism is developed for computing charge-state splittings of deep levels associated with paired defects in semiconductors and is applied to nearest-neighbor substitutional sulfur pairs in silicon. Self-consistent calculations predict a 0.19 eV splitting between the deep levels of nearest-neighbor $(S,S)^0$ and $(S,S)^+$ in Si, in good agreement with the experimental value of 0.18 eV. Computed ratios of hyperfine tensor components also agree with available data. Our results lend support to the meso-bonding theory of paired-chalcogen deep levels.

*To whom correspondence should be addressed.

†Present Address:
Department of Physics
University of Texas at Austin
Austin, Texas 78712, U.S.A.

G. Kim, J. D. Dow, and S. Lee

SULFUR NEAREST-NEIGHBOR DEFECT PAIRS IN Si

1. INTRODUCTION

Sulfur and other chalcogen impurities in Si are known to occupy substitutional sites and, when present in abundance, form nearest-neighbor pairs with deep levels in the fundamental band gap whose characters are currently controversial [1–4]. The deep level in the fundamental band gap for an isolated-S impurity in Si is known to have A_1 or s-like character [5], and is pulled down from the conduction band by the strong (~7 eV) electron-attractive defect potential of a S atom (relative to the Si atom it replaces). It is somewhat surprising, then, to realize that a pair of S atoms, (S,S) produces a deep level at higher energy than an isolated-S atom, because the additional S defect potential should pull levels down in energy. Sankey *et al.* [2] explained this fact by showing that the A_1 deep level of isolated-S is pulled down in energy so much that it lies resonant with the valence band: The (S,S) deep level in the gap is derived not from the s-like A_1 level but from the p-like T_2 level of the isolated S defect, which for isolated-S lies in the conduction band, and is pulled down into the gap by the defect potential of the second S atom.

In explaining this fact, Sankey *et al.* made an assertion concerning the character of the "molecular" wavefunction for the (S,S) deep level in the gap: it is meso-bonding, namely an a_1-symmetric bonding linear combination of the antibonding T_2-symmetric wavefunctions of the two isolated-S defects. (The T_2 state is resonant with the conduction band for isolated-S, not in the fundamental band gap.) This implies that the a_1-symmetric (totally) antibonding linear combination of the isolated-S A_1-symmetric antibonding deep level wavefunctions lies at lower-energy. Further evidence supporting the Sankey viewpoint has been presented by Hu *et al.* [4] who also concluded that (S,S) is meso-bonding.

However, Wörner *et al.* [3] have interpreted data for the nearest-neighbor $(Se,Se)^+$ deep level in the gap as evidence for a totally antibonding state: an antibonding linear combination of antibonding isolated-Se A_1-symmetric deep level wavefunctions. (Of course, the characters of the (S,S) and the (Se,Se) deep levels should be similar, according to all of the theories.) Thus there is a clear disagreement concerning the character of the main chalcogen-pair deep level in the gap, between, on the one hand, Sankey *et al.* [2] and Hu *et al.* [4] who favor meso-bonding, and, on the other, Wörner *et al.* [3] who propose a totally antibonding state.

In this paper we present calculations of the charge-state splittings of deep levels in the gap of Si associated with chalcogen nearest-neighbor substitutional defects. Our goal is to determine if those splittings can be understood by using a theory in which the deep levels are meso-bonding in character—and we find that they can. Thus our results lend more support to the viewpoint of Sankey *et al.*

2. CHARGE-STATE SPLITTINGS

The charge-state splitting of a deep level is a many-body effect which is zero in ordinary one-electron theories. It is the difference ΔE between the ionization energies E of the neutral and charged defects $(S,S)^0$ and $(S,S)^+$.

To understand the origin of charge-state splittings, first consider the charge-state splitting of atomic He. Neutral He has two 1s electrons in nearly-hydrogenic orbitals described by effective charges $Z_{\rm eff} = 27/16 = 1.7$ [6]. But He^+ has one 1s electron in an orbital with an effective charge of 2. The difference in ionization energies of a 1s electron in He and He^- is (assuming Koopmans' theorem [7, 8]) approximately the 1s orbital energy. Thus, for atomic He and for most atoms, the charge-state splittings ΔE are of order 20 eV and are Coulomb energies:

$$\Delta E \sim 20\,\text{eV} \sim \langle e^2/r \rangle. \qquad (1)$$

where $\langle e^2/r \rangle$ denotes an appropriate Coulomb integral [6], and is of order $e^2/r_{\rm av}$, where $r_{\rm av}$ is the average distance between, for example, the two 1s electrons in He.

For deep levels in the fundamental band gap of a semiconductor that are associated with substitutional point-defects such as S in Si, the charge-state splitting is also a Coulomb integral of the approximate order of magnitude

$$\Delta E \sim e^2/\varepsilon r_{\rm av} \sim 0.3\,\text{eV}. \qquad (2)$$

This is about two-orders of magnitude smaller than in atoms because: (*i*) the dielectric constant ε is of order 12 in typical semiconductors (*versus* unity for atoms); and (*ii*) the average distance $r_{\rm av}$ is considerably larger (a lattice constant rather than a Bohr radius) due to the antibonding character of the deep level wavefunctions [9]. In He the Coulomb forces cause the two 1s electrons to circumnavigate the

atom 180° out of phase (classically), and so their average separation is of the order of the Bohr radius. But the wavefunction of the S deep level in Si is now known to be antibonding in character [4, 5, 9–13], and so is not peaked on the S atom but rather on adjacent Si atoms. Hence r_{av} is the average separation of electrons on different Si atoms neighboring the S defect—of order a lattice constant, almost an order of magnitude larger than a Bohr radius.

The charge-state splitting of a meso-bonding (S,S) deep level in Si should be somewhat smaller than for isolated-S, both because the average inter-electron distance r_{av} is larger for the "molecular" defect than for the "atomic" defect and because the "molecular" deep level has wavefunctions that are T_2 or p-like in character, with even less amplitude in the S cells than the A_1 or s-like states of the isolated-S deep level. Thus, if the meso-bonding character of nearest-neighbor paired-chalcogen deep levels in Si is correct, the theory should be able to correctly determine the magnitude of the charge-state splitting.

3. MODEL

To evaluate the charge-state splitting of a chalcogen pair in Si, we employ the Hjalmarson *et al.* theory of deep impurity levels [14] in combination with the Haldane–Anderson scheme for introducing Coulomb effects [15, 16]. This approach has been used successfully by Sankey *et al.* [10] to discuss deep levels caused by interstitial impurities in Si, by Vogl and Baranowski [17] to treat deep levels of transition-metal impurities in semiconductors, and by Lee *et al.* [9] to determine charge-state splittings of deep levels associated with substitutional point defects in semiconductors. We first repeat the theory of Lee *et al.*, which is a self-consistent tight-binding model of deep levels, and verify the physics of the model for isolated impurities (while rectifying some numerical errors [18]), and then extend the theory to paired defects.

4. THEORY

4.1. General

The Hamiltonian H_0 of the perfect Si crystal is taken to be the one-electron tight-binding Hamiltonian of Vogl *et al.* [19]. Expressed in terms of the tight-binding basis states $|i, b, \mathbf{k}\rangle$ obtained from the localized orbitals $|i, b, \mathrm{S}\rangle$ centered on the b-th site in the unit cell at S:

$$|i, b, \mathbf{k}\rangle = N^{-1/2} \Sigma_S \exp(i\mathbf{k}\cdot\mathrm{S} + i\mathbf{k}\cdot\mathrm{v}_b)\, |i, b, \mathrm{S}\rangle \,, \qquad (3)$$

the Hamiltonian in k-space is:

$$H_0(\mathbf{k}) = \Sigma_{i,b} |i, b, \mathbf{k}\rangle\, E(i, b)\, \langle i, b, \mathbf{k}| + \Sigma_{i,j} [|\mathrm{i, a, k}\rangle\, T(ia, jc)\, \langle j, c, \mathbf{k}| + \mathrm{h.c.}] \,. \qquad (4)$$

Here h.c. means Hermitian conjugate, i and j run over the basis orbitals, s, p_x, p_y, p_z, and s^*, and b labels one of the two sites in the unit cell of Si (we shall refer to the sites as anion (=a) and cation (=c) sites to emphasize that the theory for diamond-structure Si can be applied to zincblende-structure semiconductors as well). The parameters $E(i, b)$ and $T(ia, jc)$ are tabulated in reference [19]. The eigenstates of H_0 are the Bloch waves $|n, \mathbf{k}\rangle$ of Si, and the eigenvalues $E_n(\mathbf{k})$ are the host band structure.

The perturbed Hamiltonian H describes the (S,S) defect, and the defect potential is $V \equiv H - H_0$. The Hamiltonians H and H_0 are similar in form, with the localized basis functions for H being (formally) Löwdin orbitals. The defect levels E are the solutions of the secular equation

$$\det(1 - G_0(E)V) = 0 \,, \qquad (5)$$

where $G_0(E)$ is the perfect-crystal Green's function operator

$$G_0(E) = (E - H_0)^{-1} . \qquad (6)$$

In cases such that E is degenerate with a host energy band, the boundary conditions are satisfied if we replace E with $E + i\varepsilon$, where ε is a positive infinitesimal. The advantage of the Green's function formalism is that the dimension of the secular equation matrix is small: the size of the defect matrix V rather than the size of the crystal. Determining the defect energy levels in this formalism reduces to: (*i*) constructing the host Green's function $G_0(E)$; (*ii*) obtaining the defect potential V; and (*iii*) solving the secular equation for E.

4.2. Green's Function

The host Green's function is

$$G_0(E) = \Sigma_{n,\mathbf{k}} [|n, \mathbf{k}\rangle\, \langle n, \mathbf{k}|] / [E - E_n(\mathbf{k})] \,. \qquad (7)$$

4.3. Defect Potential

4.3.1. Symmetry

The defect potential V is assumed to be a mean-field one-electron potential whose value depends on the charge-state of the (S,S) defect. The solution of the secular equation is facilitated if the range of this

potential is minimal. Therefore, we neglect: (*i*) the long-ranged non-central-cell Coulombic parts of the defect potential; (*ii*) lattice relaxation around the impurities; and (*iii*) matrix elements of the potential associated with s^* orbitals. These approximations have been discussed in other contexts [14], are standard approximations for treating deep levels, and lead to errors of a few tenths of an eV in predicted absolute deep level energies [14, 19]. Since our goal is to understand the physics of charge-state splittings, which are differences in deep level energies, the errors introduced by these approximations for each of the deep levels should, to a large (but not complete) extent, cancel in the difference. Concentrating on the form of the first-order perturbation correction to the energy and using the fact that most of the electronic charge in a deep level is distant from the defect [5], we estimate the anticipated uncertainty in the charge-state splitting as the uncertainty in the absolute energy position of a deep level multiplied by the fractional difference in the central-cell charge densities (See Tables 3 and 4, below.)—of order 0.01 eV for (S,S) in Si. Thus the theory of charge-state splitting should be qualitatively predictive, chemically correct, and quantitatively accurate.

Thus we take the defect potential (in a localized Löwdin orbital basis of s, p_x, p_y, and p_z orbitals) to involve only the two S sites and the six neighboring Si sites directly bonded to S atoms. Hence the defect matrix is a 32×32 matrix. The omission of a more-distant Si site at R introduces a small error of order $e^2|\psi(R)|^2 v/\varepsilon R$ in the absolute deep level energy and $e^2\delta|\psi(R)|^2 v/\varepsilon R$ or ~0.01 eV in the (S,S) charge-state splitting, where v is the atomic volume and $\delta|\psi(R)|^2$ is the difference in wavefunctions of the charged and neutral defects, and is small. (See Tables 3 and 4, below.) The point-group symmetry of a substitutional isolated-S defect in Si (or in a zincblende semiconductor) is tetrahedral: T_d. The group of a (S,S) pair oriented along the (111) crystal axis is C_{3v}, with the vertical reflection plane bisecting the (111) bond. (Note that the symmetry is reduced to C_3 for a (Se,S) pair, for example.) The irreducible representations of C_{3v} are a_1 (σ-like), a_2 (rotation about the S–S bond), and e (π-like). By selecting as basis functions the linear combinations of Löwdin orbitals that transform according to the irreducible representations of C_{3v} (See Table 1), the secular determinant matrix assumes the form of a direct sum matrix $10a_1+2a_2+10e$, and the 32×32 secular determinant factors into one 10×10 determinant, one 2×2 determinant, and one 20×20 determinant, each with the form

$$\det(1-G_oV)=0\,, \tag{8}$$

but involving only matrix elements between the basis functions of the irreducible representations. For example, the 2×2 equation for a_2-symmetric deep levels is

$$\begin{vmatrix} 1-\langle 1|G_o|1\rangle\langle 1|V|1\rangle & -\langle 1|G_o|2\rangle\langle 2|V|2\rangle \\ -\langle 2|G_o|1\rangle\langle 1|V|1\rangle & 1-\langle 2|G_o|2\rangle\langle 2|V|2\rangle \end{vmatrix}=0 \tag{9}$$

where $|1\rangle$ and $|2\rangle$ are the basis functions for the a_2 representation listed in Table 1, and we have

$$\langle 1|V|1\rangle = (1/6)\{V(p_x,R_2)+V(p_y,R_3)+V(p_x,R_4) + V(p_y,R_2)+V(p_z,R_3)+V(p_z,R_4)\} \tag{10a}$$

$$\langle 2|V|2\rangle = (1/6)\{V(p_x,R_7)+V(p_y,R_5)+V(p_x,R_6) + V(p_y,R_7)+V(p_z,R_5)+V(p_z,R_6)\} \tag{10b}$$

Here we have $R_0=(0,0,0)$ and $R_1=(a_L/2)(1,1,1)$; these are the positions of the substitutional sulfur atoms in the silicon crystal. The neighboring Si atoms are at the positions $R_2=(a_L/2)(-1,-1,1)$, $R_3=(a_L/2)(1,-1,-1)$, $R_4=(a_L/2)(-1,1,-1)$, $R_5=(a_L/2)(0,2,2)$, $R_6=(a_L/2)(2,0,2)$, and $R_7=(a_L/2)(2,2,0)$, where a_L is the lattice constant of the perfect silicon crystal. We have also assumed, following Hjalmarson *et al.* [14], that there is no lattice relaxation and hence that the defect potential is diagonal. (Therefore we take $V(p_\mu,R_i)=V(p,(Si))$ for all μ and $1<i<8$.)

Similarly, for a_1 (σ-like) levels we have:

$$\det D=0\,, \tag{11}$$

where,

$$D_{i,j}=(1-G_oV)_{i,j}=\delta_{i,j}-\sum_{l=1}^{10}\langle i|G_o|l\rangle\langle l|V|j\rangle \tag{12}$$

Here δ is a Kronecker delta, and $|i\rangle$, with $i=1,2,\ldots 10$, are the basis functions for the a_1 irreducible representation (Table 1). If we adopt Hjalmarson's approximations for the potential, we find

$$D_{i,j}=\delta_{i,j}-\langle i|G_o|j\rangle\langle i|V|i\rangle\,. \tag{13}$$

Table 1. Basis Functions of the Irreducible Representations for Paired Sulfur Impurities.

Irreducible Representations of C_{3v}	Basis Functions
a_1	$\|s,R_0\rangle$,
	$\|s,R_1\rangle$,
	$(1/\sqrt{3})\{\|s,R_2\rangle + \|s,R_3\rangle + \|s,R_4\rangle\}$,
	$(1/\sqrt{3})\{\|s,R_5\rangle + \|s,R_6\rangle + \|s,R_7\rangle\}$,
	$(1/\sqrt{3})\{\|p_z,R_0\rangle + \|p_x,R_0\rangle + \|p_y,R_0\rangle\}$,
	$(1/\sqrt{3})\{\|p_z,R_1\rangle + \|p_x,R_1\rangle + \|p_y,R_1\rangle\}$,
	$(1/\sqrt{3})\{\|p_z,R_2\rangle + \|p_x,R_3\rangle + \|p_y,R_4\rangle\}$,
	$(1/\sqrt{3})\{\|p_z,R_7\rangle + \|p_x,R_5\rangle + \|p_y,R_6\rangle\}$,
	$(1/\sqrt{6})\{\|p_x,R_2\rangle + \|p_y,R_3\rangle + \|p_x,R_4\rangle + \|p_y,R_2\rangle + \|p_z,R_3\rangle + \|p_z,R_4\rangle\}$, and
	$(1/\sqrt{6})\{\|p_x,R_7\rangle + \|p_y,R_5\rangle + \|p_x,R_6\rangle + \|p_y,R_7\rangle + \|p_z,R_5\rangle + \|p_z,R_6\rangle\}$
a_2	$(1/\sqrt{6})\{\|p_x,R_2\rangle + \|p_y,R_3\rangle - \|p_x,R_4\rangle - \|p_y,R_2\rangle - \|p_z,R_3\rangle + \|p_z,R_4\rangle\}$, and
	$(1/\sqrt{6})\{\|p_x,R_7\rangle + \|p_y,R_5\rangle - \|p_x,R_6\rangle - \|p_y,R_7\rangle - \|p_z,R_5\rangle + \|p_z,R_6\rangle\}$
e	$(1/\sqrt{6})\{2\|p_x,R_0\rangle - \|p_y,R_0\rangle - \|p_z,R_0\rangle\}$, and
	$(1/\sqrt{6})\{2\|p_x,R_1\rangle - \|p_y,R_1\rangle - \|p_z,R_1\rangle\}$;
	$(1/\sqrt{2})\{\|p_y,R_0\rangle - \|p_z,R_0\rangle\}$, and
	$(1/\sqrt{2})\{\|p_y,R_1\rangle - \|p_z,R_1\rangle\}$;
	$(1/\sqrt{6})\{\|s,R_2\rangle - 2\|s,R_3\rangle + \|s,R_4\rangle\}$, and
	$(1/\sqrt{6})\{\|s,R_7\rangle - 2\|s,R_5\rangle + \|s,R_6\rangle\}$;
	$(1/\sqrt{2})\{\|s,R_2\rangle - \|s,R_4\rangle\}$, and
	$(1/\sqrt{2})\{\|s,R_7\rangle - \|s,R_6\rangle\}$;
	$(1/\sqrt{12})\{2\|p_x,R_2\rangle - \|p_y,R_3\rangle + 2\|p_x,R_4\rangle - \|p_y,R_2\rangle - \|p_z,R_3\rangle - \|p_z,R_4\rangle\}$, and
	$(1/\sqrt{12})\{2\|p_x,R_7\rangle - \|p_y,R_5\rangle + 2\|p_x,R_6\rangle - \|p_y,R_7\rangle - \|p_z,R_5\rangle - \|p_z,R_6\rangle\}$;
	$(1/\sqrt{12})\{2\|p_x,R_2\rangle - \|p_y,R_3\rangle - 2\|p_x,R_4\rangle + \|p_y,R_2\rangle + \|p_z,R_3\rangle - \|p_z,R_4\rangle\}$, and
	$(1/\sqrt{12})\{2\|p_x,R_7\rangle - \|p_y,R_5\rangle - 2\|p_x,R_6\rangle + \|p_y,R_7\rangle + \|p_z,R_5\rangle - \|p_z,R_6\rangle\}$;
	$(1/\sqrt{6})\{\|p_z,R_2\rangle - 2\|p_x,R_3\rangle + \|p_y,R_4\rangle\}$, and
	$(1/\sqrt{6})\{\|p_z,R_7\rangle - 2\|p_x,R_5\rangle + \|p_y,R_6\rangle\}$;
	$(1/\sqrt{2})\{\|p_z,R_2\rangle - \|p_y,R_4\rangle\}$ and $(1/\sqrt{2})\{\|p_z,R_7\rangle - \|p_y,R_6\rangle\}$;
	$(1/2)\{\|p_y,R_2\rangle + \|p_y,R_3\rangle - \|p_z,R_3\rangle - \|p_z,R_4\rangle\}$, and
	$(1/2)\{\|p_y,R_7\rangle + \|p_y,R_5\rangle - \|p_z,R_5\rangle - \|p_z,R_6\rangle\}$; and
	$(1/2)\{-\|p_y,R_2\rangle + \|p_y,R_3\rangle + \|p_z,R_3\rangle - \|p_z,R_4\rangle\}$, and
	$(1/2)\{-\|p_y,R_7\rangle + \|p_y,R_5\rangle + \|p_z,R_5\rangle - \|p_z,R_6\rangle\}$

Here we have

$$\langle 1|V|1\rangle = V(s,R_0) = V(s,(S)) = \langle 2|V|2\rangle = V(s,R_1),$$

$$\langle 3|V|3\rangle = (1/3)\{V(s,R_2) + V(s,R_3) + V(s,R_4)\} = V(s,(Si)) = \langle 4|V|4\rangle,$$

$$\langle 5|V|5\rangle = (1/3)\{V(p_z,R_0) + V(p_x,R_0) + V(p_y,R_0)\} = V(p,(S)) = \langle 6|V|6\rangle,$$

and, for $n = 7$, 8, 9, and 10,

$$\langle n|V|n\rangle = (1/3)\{V(p_z,R_2) + V(p_x,R_3) + V(p_y,R_4)\} = V(p,(Si)). \quad (14)$$

Here we have $V(s,(S)) = V(s,R_0)$, $V(p,(S)) = V(p_\mu,R_0)$, $V(s,(Si)) = V(s,R_2)$, and $V(p,(Si)) = V(p_\mu,R_2)$, where μ is x, y, or z.

For π-like e-states we have a similar secular determinant to Equation (13) involving the basis states of Table 1, with

$$\langle n|V|n\rangle = (1/6)\{V(p_x, R_o) + V(p_y, R_o) + V(p_z, R_o)\} = V(p,(S)),$$

for $n = 1$, 2, 3, and 4, and with

$$\langle l|V|l\rangle = (1/6)\{V(s,R_2) + 4V(s,R_3) + V(s,R_4)\} = V(s,(Si)),$$

for $l = 5$, 6, 7, and 8, and

$$\langle m|V|m\rangle = (1/12)\{4V(p_x,R_2) + V(p_y,R_3) + 4V(p_x,R_4) + V(p_y,R_2) + V(p_z,R_3) V(p_z,R_4)\} = V(p,(Si)),$$

for $m = 9, 10,\ldots, 20.$ (15)

Similar results hold for the isolated-S defect, with the basis functions of Table 2 (See Appendix.)

4.3.2. Dependence on Mean-Field

The effective mean-field potential felt by an electron in a deep level depends on the charge-state of the level, much as in atomic physics the potential and orbital energy depend on the atomic charge state. For example, in He^+ the 1s electron's potential is

Table 2. Basis Functions of the Irreducible Representations for an Isolated Sulfur Impurity in Si at the Origin. The Notation is the Same as for Table 1.

Irreducible Representations of T_d	Basis Functions
A_1	$\|s,R_o\rangle$, $(1/2)\{\|s,R_1\rangle + \|s,R_2\rangle + \|s,R_3\rangle, + \|s,R_4\rangle\}$, and $(1/\sqrt{12})\{\|p_x,R_1\rangle + \|p_y,R_1\rangle + \|p_z,R_1\rangle - \|p_x,R_2\rangle - \|p_y,R_2\rangle + \|p_z,R_2\rangle + \|p_x,R_3\rangle - \|p_y,R_3\rangle - \|p_z,R_3\rangle - \|p_x,R_4\rangle + \|p_y,R_4\rangle - \|p_z,R_4\rangle\}$
E	$(1/\sqrt{24})\{2\|p_x,R_1\rangle - \|p_y,R_1\rangle - \|p_z,R_1\rangle - 2\|p_x,R_2\rangle + \|p_y,R_2\rangle - \|p_z,R_2\rangle + 2\|p_x,R_3\rangle + \|p_y,R_3\rangle + \|p_z,R_3\rangle - 2\|p_x,R_4\rangle - \|p_y,R_4\rangle + \|p_z,R_4\rangle\}$ and $(1/\sqrt{8})\{\|p_y,R_1\rangle - \|p_z,R_1\rangle - \|p_y,R_2\rangle - \|p_z,R_2\rangle - \|p_y,R_3\rangle + \|p_z,R_3\rangle + \|p_y,R_4\rangle + \|p_z,R_4\rangle\}$
T_1	$(1/\sqrt{8})\{\|p_y,R_1\rangle - \|p_z,R_1\rangle + \|p_y,R_2\rangle + \|p_z,R_2\rangle - \|p_y,R_3\rangle + \|p_z,R_3\rangle - \|p_y,R_4\rangle - \|p_z,R_4\rangle\}$ $(1/\sqrt{8})\{\|p_x,R_1\rangle - \|p_z,R_1\rangle + \|p_x,R_2\rangle + \|p_z,R_2\rangle - \|p_x,R_3\rangle + \|p_z,R_3\rangle - \|p_x,R_4\rangle - \|p_z,R_4\rangle\}$ $(1/\sqrt{8})\{\|p_x,R_1\rangle - \|p_y,R_1\rangle + \|p_x,R_2\rangle + \|p_y,R_2\rangle - \|p_x,R_3\rangle + \|p_y,R_3\rangle - \|p_x,R_4\rangle - \|p_y,R_4\rangle\}$
T_2	$\|p_x,R_o\rangle$, $\|p_y,R_o\rangle$, and $\|p_z,R_o\rangle$; $(1/2)\{\|s,R_1\rangle - \|s,R_2\rangle + \|s,R_3\rangle - \|s,R_4\rangle\}$, $(1/2)\{\|s,R_1\rangle - \|s,R_2\rangle - \|s,R_3\rangle + \|s,R_4\rangle\}$, and $(1/2)\{\|s,R_1\rangle + \|s,R_2\rangle - \|s,R_3\rangle - \|s,R_4\rangle\}$; $(1/2)\{\|p_x,R_1\rangle + \|p_x,R_2\rangle + \|p_x,R_3\rangle + \|p_x,R_4\rangle\}$, $(1/2)\{\|p_y,R_1\rangle + \|p_y,R_2\rangle + \|p_y,R_3\rangle + \|p_y,R_4\rangle\}$, and $(1/2)\{\|p_z,R_1\rangle + \|p_z,R_2\rangle + \|p_z,R_3\rangle + \|p_z,R_4\rangle\}$; and $(1/\sqrt{8})\{\|p_y,R_1\rangle + \|p_y,R_2\rangle - \|p_y,R_3\rangle - \|p_y,R_4\rangle + \|p_z,R_1\rangle - \|p_z,R_2\rangle - \|p_z,R_3\rangle + \|p_z,R_4\rangle\}$, $(1/8)\{\|p_x,R_1\rangle + \|p_x,R_2\rangle - \|p_x,R_3\rangle - \|p_x,R_4\rangle + \|p_z,R_1\rangle - \|p_z,R_2\rangle + \|p_z,R_3\rangle - \|p_z,R_4\rangle\}$, and $(1/\sqrt{8})\{\|p_x,R_1\rangle - \|p_x,R_2\rangle - \|p_x,R_3\rangle + \|p_x,R_4\rangle + \|p_y,R_1\rangle - \|p_y,R_2\rangle + \|p_y,R_3\rangle - \|p_y,R_4\rangle\}$

$-2e^2/r$, but in He° the effective (Hartree) potential is $-Z_{\rm eff}e^2/r$, where $Z_{\rm eff}$ is 1.7 [8]. Haldane and Anderson [16] have proposed a simple scheme for treating such Coulomb-induced charge-state effects.

The atomic orbital energies are written

$$E(\mathrm{s},\sigma) = E_s^\circ + U_{ss}\,\Sigma'_{\sigma'}\, n(\mathrm{s},\sigma') + U_{sp}\,\Sigma_{\mu,\sigma'}\, n(\mathrm{p}_\mu,\sigma') \qquad (16a)$$

and

$$E(\mathrm{p}_\nu,\sigma) = E_p^\circ + U_{pp}\,\Sigma'_{\mu,\sigma'}\, n(\mathrm{p}_\mu,\sigma') + U_{sp}\,\Sigma_{\sigma'}\, n(\mathrm{s},\sigma')\,, \qquad (16b)$$

where σ is a spin index (up or down), ν and μ run over x, y, and z, and the prime on the summations means that terms with $\sigma=\sigma'$ and $\nu=\mu$ are excluded. The occupation numbers n are either unity or zero in atoms. The Coulomb integrals U_{ss}, U_{sp}, and U_{pp}, together with the orbital energies E_s° and E_p°, are determined by fitting atomic spectra. (This has been done for S and Si by Sankey *et al.* [10].)

In the Haldane–Anderson scheme, the occupation numbers for atoms in the solid are allowed to assume non-integral values (reflecting the fact that the charge is not localized on a single atom). Thus there are occupation numbers $n(i,\sigma;\mathrm{R})$ for each site, and expressions for the orbital energies analogous to Equations (16*a*) and (16*b*) at each site. (Note that R labels the site, not the unit cell, and so can refer to either an anion or a cation site.)

The occupation numbers $n(i,\sigma;\mathrm{R})$ in the defect-perturbed solid are

$$n(i,\sigma;\mathrm{R}) = \int_{-\infty}^{\infty} f(E)\langle i,\sigma;\mathrm{R}|\delta(E-H)|i,\sigma;\mathrm{R}\rangle\,\mathrm{d}E, \qquad (17)$$

where $f(E)$ is the Fermi–Dirac function (which, at zero temperature is unity for valence band states and occupied deep levels in the gap, and zero otherwise).

The spectral density operator $\delta(E-H)$ is related to the perturbed and unperturbed Green's function operators G and G_o as follows:

$$\delta(E-H) = (-1/\pi)\,\mathrm{Im}\,G = (-1/\pi)\,\mathrm{Im}\,[G_o(1-G_oV)^{-1}]\,. \qquad (18)$$

Thus at zero temperature we have

$$n(i,\sigma;\mathrm{R}) = (-1/\pi)\int_{-\infty}^{E_v}\mathrm{Im}\,\langle i,\sigma,\mathrm{R}|G(E)|i,\sigma,\mathrm{R}\rangle + \Sigma_j|\langle i,\sigma;\mathrm{R}|\Psi_j\rangle|^2 \qquad (19)$$

where E_v is the valence band maximum and the summation runs over the occupied bound state: in the band gap. The bound states are normalized according to the condition [20]

$$\langle\Psi_j|V(\mathrm{d}/\mathrm{d}E)\,[G_o(E)]\,V|\Psi_j\rangle = -1\,. \qquad (20)$$

Following Hjalmarson *et al.* [14], the defect potential $V\equiv H-H_o$ is diagonal in a Löwdin-orbital basis for the perturbed crystal, with the values

$$\langle i,\sigma;\mathrm{R}|V|i,\sigma;\mathrm{R}\rangle = \beta_i[E(i,\sigma;\mathrm{R})-W(i,\sigma;\mathrm{R})] \qquad (21)$$

where $\beta_i=0.8$ for $i=$ s and $\beta_i=0.6$ for $i=\mathrm{p}_x$, p_y, or p_z [14]. W is the value of E for Si. Thus the defect potential depends implicitly on the occupation numbers n through Equations (21) and (16), and the occupation numbers depend on the defect potential through Equations (17), (18), and (19). The difference between $(\mathrm{S,S})^+$ and $(\mathrm{S,S})^\circ$ comes ultimately from the difference in occupation numbers n, which are self-consistently determined. We solve this problem iteratively and find a self-consistent solution to the equations. The resulting occupation numbers are given in Table 3. As expected, (*i*) they add to (almost) six electrons per S atom and four per Si, and (*ii*) they are almost the same for $(\mathrm{S,S})^+$ and $(\mathrm{S,S})^\circ$—as was the case for isolated S^+ and S° [9] (See Table 4.).

5. RESULTS

5.1. Charge-State Splitting

The principal result of the calculation is a charge-state splitting for (S,S) of 0.19 eV that agrees remarkably well with the experimental value of 0.18 eV [12, 13]. For the isolated substitutional sulfur impurity,

Table 3. Computed Occupation Numbers $n(i,\sigma;\mathrm{R})$ for the (S,S) Defect in Si.

Occupation Number	$(\mathrm{S,S})^\circ$	$(\mathrm{S,S})^+$
$\Sigma_\sigma\, n(\mathrm{s}\sigma;\mathrm{R}_o)$	2.01	1.96
$\Sigma_{\mu,\sigma}\, n(\mathrm{p}_\mu\sigma;\mathrm{R}_o)$	3.86	3.86
$\Sigma_\sigma\, n(\mathrm{s}\sigma;\mathrm{R}_2)$	1.26	1.26
$\Sigma_{\mu,\sigma}\, n(\mathrm{p}_\mu\sigma;\mathrm{R}_2)$	2.65	2.60

Table 4. Computed Occupation Numbers for the Isolated-S Defect in Si.

Occupation Number	$(\mathrm{S})^\circ$	$(\mathrm{S})^+$
$\Sigma_\sigma\, n(\mathrm{s}\sigma;\mathrm{R}_o)$	1.97	1.92
$\Sigma_{\mu,\sigma}\, n(\mathrm{p}_\mu\sigma;\mathrm{R}_o)$	3.84	3.87
$\Sigma_\sigma\, n(\mathrm{s}\sigma;\mathrm{R}_1)$	1.25	1.30
$\Sigma_{\mu,\sigma}\, n(\mathrm{p}_\mu\sigma;\mathrm{R}_1)$	2.64	2.57

we find a charge-state splitting of 0.27 eV, compared with the experimental value of 0.30 eV [12, 13]. The smaller charge-state splitting for the gap deep level of (S,S) than the gap level of S is due largely to the p-like character of the "molecular" level (which is derived from the isolated-S T_2 resonance in the conduction band, not from the isolated-S A_1 deep level in the gap). Because of this p-like character, the deep level of $(S,S)^+$ has only about 5% of its electron's charge on either S site, whereas the s-like A_1 deep level of S^+ has about 10% of an electron in the central cell. (See Tables 5 and 6.) About 54% of the charge lies on the six neighboring Si atoms, with the rest being further away from the defect.

5.2. Charge Densities and Hyperfine Tensor

Our charge densities are essentially the same as those of Sankey *et al.* [1, 2] for $(S,S)^+$ and Ren *et al.* [5] for $(S)^+$—and so agree with the data [3, 11, 13, 21, 22], as those theories do. For sulfur pairs Sankey's wavefunction coefficient γ_s [2, 21, 22] is 0.21 and 0.20 for $(S,S)^0$ and $(S,S)^+$, respectively; we find 0.18 for both charge states [23].

The hyperfine tensor **A** depends on the contact interaction with the nuclei and hence samples the deep level's charge density at the relevant sites. Thus the various measured ratios of hyperfine tensor elements are related to charge densities at the defect's sites, as discussed in detail by Sankey *et al.* [1, 2]. We find

$$(A_\parallel - A_\perp)/(A_\parallel + 2A_\perp) = 0.016 ,$$

compared with Sankey's values of 0.015. (These small values explain why the experimental hyperfine tensor has appeared to be isotropic for such an anisotropic defect.) We are unaware of any measurement of this ratio for $(S,S)^+$, but the corresponding ratio measured for $(Se,Se)^+$ is 0.01 [3]. The comparison of theoretical results for (S,S) with data for (Se,Se) is appropriate, despite one's initial impulse to the contrary, because the wavefunctions of S and Se deep levels in Si are known to be almost the same [2, 5], due to the antibonding character of the deep levels: these levels are Si dangling-bond-like [14].

For the ratio of S-site hyperfine tensor components of the "molecule" $(S,S)^+$ and the "atom" $(S)^+$, we find

$$A_\parallel[(S,S)^+]/A_\parallel[(S)^+] = 0.35 ,$$

in agreement with the experimental value of 0.37 [3]. (The same result is obtained for Se and S [3], lending support to our argument that the ratios of hyperfine tensor components are almost independent of chalcogen.)

Table 5. Computed Electronic Charge Distribution of the a_1-Symmetric Meso-bonding Deep Level State in the Fundamental Band Gap of Si Associated with Nearest-Neighbor Substitutional Sulfur-Pair Impurities $(S,S)^+$. The Index μ Runs Over x, y, and z; the Index R_j Runs Over R_0 and R_1, That is, Over the Sulfur Sites; and $R_{j'}$ Runs Over R_2, R_3, R_4, R_5, R_6, and R_7: the First-Neighbor Silicon Atoms to Each S Impurity.

State	Electron Charge (%)
$\Sigma_j\lvert\langle\Psi\lvert s,R_j(S)\rangle\rvert^2/2$	3.4
$\Sigma_{j,\mu}\lvert\langle\Psi\lvert p_\mu,R_j(S)\rangle\rvert^2/2$	1.9
$\Sigma_{j'}\lvert\langle\Psi\lvert s,R_{j'}(Si)\rangle\rvert^2/6$	1.3
$\Sigma_{j',\mu}\lvert\langle\Psi\lvert p_\mu,R_{j'}(Si)\rangle\rvert^2/6$	7.7

Table 6. Computed Electronic Charge Distribution of the A_1 Deep Level State in the Si Band Gap Associated with a Single Substitutional Sulfur Impurity $(S)^+$. The Index μ is x, y, and z.

State	Electron Charge (%)
$\lvert\langle\Psi\lvert s,R_0(S)\rangle\rvert^2$	9.6
$\Sigma_\mu\lvert\langle\Psi\lvert p_\mu,R_0(S)\rangle\rvert^2$	0.0
$\Sigma_j\lvert\langle\Psi\lvert s,R_j(Si)\rangle\rvert^2/4$	1.7
$\Sigma_{j,\mu}\lvert\langle\Psi\lvert p_\mu,R_j(Si)\rangle\rvert^2/4$	13.8

We find the ratio of perpendicular to parallel principal hyperfine tensor components to be

$$A_\perp/A_\parallel = 0.95$$

for $(S,S)^+$. This is in good agreement with the value 0.97 that Wörner and Shirmer [3] found for $(Se,Se)^+$. (The corresponding ratio for $(S,S)^+$ could not be determined experimentally.)

Thus our results describe the data very well and so lend support to the meso-bonding picture of the (S,S) deep levels in the gap of Si, while confirming the antibonding character of the A_1 isolated-S deep level.

5.3. Absolute Level Positions

The Vogl model of electronic structure has a feature that is reflected in the absolute energies of the calculated S and (S,S) deep levels, namely the p-states are artificially depressed in energy due to the strong s^*-p coupling, necessary to produce the correct indirect fundamental gap while simultaneously restricting the basis set to only five orbitals per

atom. As a result, the calculated absolute deep level energies of both S and (S,S) are several tenths of an eV below the corresponding experimental values [24]. This problem is analogous to a more serious difficulty of local density theory, which yields band gaps too small by typically a factor of two [20, 25, 26]. In the Vogl model of Si, the strength of the s^*-p coupling is adjusted to depress the indirect conduction-band minimum to its observed energy. This has the effect of also depressing p-like deep levels in Si, sometimes by too much. One could, of course, correct for this by including more and more orbitals in the basis set; but the costs of such a correction, in terms of increased computational labor and reduced physical transparency, normally outweigh the benefits. It is easier, and almost as accurate, to simply mentally adjust the p-like deep level predictions of the Vogl model upwards in energy by a few tenths of an eV; normally it does not make sense to formally correct the theory. Nevertheless we can assess the seriousness of this problem by increasing the energy of the s^* orbital until the neutral isolated-S deep level coincides with the experimental level at 0.86 eV (with respect to the valence band maximum) and then determine if the charge-state splittings of S and (S,S), as well as the absolute energy of the $(S,S)^0$ level, agree with the data. In this case we find that the charge-state splitting of S is 0.32 eV (compared with the experimental value of 0.30 eV [12, 13] and the value 0.27 eV obtained with an unadjusted s^* energy), the charge-state splitting of (S,S) is 0.15 eV (*versus* experiment with 0.18 eV [12, 13] and unadjusted theory of 0.19 eV), and an absolute $(S,S)^0$ level of 0.97 eV, in agreement with the observed energy 0.98 eV. Furthermore the totally antibonding A_1-derived a_1-symmetric deep levels of $(S,S)^0$ and $(S,S)^+$ lie in the valence band, not in the gap [27]. The calculated charge-state splitting is almost independent of the approximations of the sp^3s^* Vogl model.

6. SUMMARY

We conclude that the basic physical picture of a meso-bonding (S,S) deep level in the fundamental band gap of Si is valid, that the charge-state splitting of (S,S) is smaller than that of isolated-S because of the p-like meso-bonding character of (S,S) *versus* the s-like A_1 character of the S deep level, and that the theory predicts the charge-state splittings, charge-densities, and hyperfine tensors with rather good accuracy: for example, the uncertainty in the charge-state splitting predictions associated with quirks of the model is of order 0.01 eV.

ACKNOWLEDGEMENTS

The authors wish to thank E. Yamaguchi for helpful discussions. We are grateful to the U.S. Office of Naval Research for generous support (Contract No. N00014-84-0352), and one of us (J.D.D.) thanks the Minnesota Supercomputer Institute for support.

APPENDIX: ISOLATED SULFUR

For a single substitutional S impurity in Si, the defect potential involves the impurity site and the four neighboring sites. With the relevant Löwdin orbitals being s, p_x, p_y, and p_z, this leads to a 20×20 defect matrix, and tetrahedral (T_d) point-group symmetry. The twenty basis functions form linear combinations (See Table 2) that reduce the secular matrix to the direct sum $3A_1+E+T_1+4T_2$.

Our calculated results for S in Si differ in numerical detail from those of reference [9], although the physics of reference [9] is still correct. The present numerical results supersede those of reference [9], and are different as detailed in reference [18]. The resulting occupation numbers are given in Table 4. The wavefunction coefficients γ_2 [2] for $(S)^0$ and $(S)^+$ are 0.30 and 0.31, respectively. These compare favorably with the value for $(S)^+$ of 0.32 obtained by Ren *et al.* [5]. Thus we have 9.6% (Table 6) of the deep level's charge in the central cell, whereas the experiments find $\simeq 10\%$ [13, 22]. Moreover, the nearest-neighbor hybrid is 89% p-like in the present theory (See Table 6: [13.8/(1.7+13.8)]), compared with $\simeq 91\%$ deduced from ESR and ENDOR data by Grimmeiss *et al.* [13] and Niklas and Spaeth [22].

REFERENCES

[1] O. F. Sankey and J. D. Dow, "Electronic Energy Level of Substitutional Defect Pairs in Si", *Physical Review B*, 26 (1982), p. 3243.

[2] O. F. Sankey and J. D. Dow, "Theory of Electron Spin Resonance Measurements of Chalcogen Pairs in Si", *Solid State Communications*, 51 (1984), p. 705.

[3] R. Wörner and O. F. Schirmer, "ESR of Selenium Pairs (Se_2^-) in Silicon", *Solid State Communications*, 51 (1984), p. 665.

[4] W. M. Hu, Y. M. Gu, and S. Y. Ren, "Electronic Structures of Chalcogenide Pairs in Si", *Acta Physica Sinica*, 35 (1986), p. 1583.

[5] S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, "Charge Densities and Wavefunctions of Chalcogenide Deep Impurities in Si", *Physical Review B*, 26 (1982), p. 951.

[6] See, for example, L. I. Schiff, *Quantum Mechanics, Third Edn.* New York: McGraw-Hill, 1968, p. 259.

[7] T. Koopmans, "Über Die Zuordnung Von Wellenfunktionen Und Eigenwerten Zu Den Einzelnen Elektronen Eines Atoms", *Physica*, 1 (1934), p. 104.

[8] See, for example, J. Callaway, *Quantum Theory of Solids*, New York: Academic Press, 1974, p. 333.

[9] S. Lee, J. D. Dow, and O. F. Sankey, "Theory of Charge-State Splittings of Deep Levels", *Physical Review B*, 31 (1985), p. 3910.

[10] O. F. Sankey and J. D. Dow, "Theory of Tetrahedral-Site Interstitial s- and p-Bonded Impurities in Si", *Physical Review B*, 27 (1983), p. 7641.

[11] G. W. Ludwig, "Paramagnetic Resonance Study of a Deep Donor in Silicon", *Physical Review A*, 137 (1965), p. 1520.

[12] S. D. Brotherton, M. J. King, and G. J. Parker, "The Electrical Properties of Sulphur in Silicon", *Journal of Applied Physics*, 52 (1981), p. 4649. See also reference [1] for citations of the experimental literature.

[13] H. G. Grimmeiss, E. Janzen, H. Eunen, O. Shirmer, J. Schneider, R. Wörner, C. Holm, E. Sirrl, and P. Wagner, "Tellurium Donors in Silicon", *Physical Review B*, 24 (1981), p. 4571; H. G. Grimmeiss, E. Janzen, and K. Larsson, "Multivalley Spin Splitting of 1s States for Sulfur, Selenium, and Tellurium Donors in Silicon", *Physical Review B*, 25 (1981), p. 2627, and references therein.

[14] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, "Theory of Substitutional Deep Traps in Covalent Semiconductors", *Physical Review Letters*, 44 (1980), p. 810; see also W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, "The Nitrogen Isoelectronic Trap in $GaAs_{1-x}P_x$. II. Model Calculation of the Electronic States N_Γ and N_X at $T = 0$", *Physical Review B*, 16 (1977), p. 1597, for a discussion of the concepts that form the foundation for this work.

[15] P. W. Anderson, "Localized Magnetic States in Metals", *Physical Review B*, 124 (1961), p. 41.

[16] F. D. M. Haldane and P. W. Anderson, "Simple Model of Multiple Charge States of Transition-Metal Impurities in Semiconductors", *Physical Review B*, 13 (1976), p. 2553.

[17] P. Vogl and J. Baranowski, "Electronic Structure of 3d-Transition Metal Impurities in Semiconductors", *Acta Physica Polonica*, A67 (1985), p. 133.

[18] We find that the theoretical A_1 deep level of $(S)^0$ lies at 0.28 eV with respect to the valence band maximum, and that the corresponding level of $(S)^-$ is at 0.01 eV, with a charge-state splitting of 0.27 eV. These results supersede the values of 1.15 eV for $(S)^0$ and 0.92 eV for $(S)^-$ (with a charge-state splitting of 0.23 eV) of reference [9]. Here we assumed $\beta_s = 0.8$, $\beta_p = 0.6$, and a band gap of Si of 1.2 eV.

[19] P. Vogl, H. P. Hjalmarson, and J. D. Dow, "A Semi-Empirical Tight-Binding Theory of the Electronic Structure of Semiconductors", *Journal of Physics and Chemistry of Solids*, 44 (1983), p. 365. In the present paper, we use the notation T, instead of the original notation V for the off-diagonal matrix elements.

[20] G. A. Baraff and M. Schlüter, "New Self-Consistent Approach to the Electronic Structure of Localized Defects in Solids", *Physical Review Letters*, 19 (1979), p. 4965; G. A. Baraff and M. Schlüter, "Calculation of the Total Energy of Charged Point Defects Using the Green's-Function Technique", *Physical Review B*, 30 (1984), p. 1853.

[21] G. D. Watkins and J. W. Corbett, "Defects in Irradiated Silicon: Electron Paramagnetic Resonance and Electron-Nuclear Double-Resonance of the Si-E Center", *Physical Review A*, 134 (1964), p. 1359.

[22] J. R. Niklas and J. M. Spaeth, "ENDOR Investigation of Tellurium Donors in Silicon", *Solid State Communications*, 46 (1983), p. 121.

[23] Sankey (reference [2]) finds s-state and p-state charge densities on the S site to be 4.1% and 2.2%, respectively, in comparison with the values 3.4% and 1.9% of Table 5.

[24] The theoretical level for $(S,S)^+$ is at 0.25 eV and the experimental level is at 0.80 eV [12].

[25] M. S. Hybertsen and S. G. Louie, "Electron Correlation in Semiconductors and Insulators: Band Gaps and Quasiparticle Energies", *Physical Review B*, 34 (1986), p. 5390.

[26] S. T. Pantelides, "The Electronic Structure of Impurities and Other Point Defects in Semiconductors", *Review of Modern Physics*, 50 (1978), p. 797; J. Bernholc, N. Lipari, S. Pantelides, and M. Scheffler, "Electronic Structure of Deep sp-Bonded Substitutional Impurities in Silicon", *Physical Review B*, 26 (1982), p. 5706; J. Bernholc, N. Lipari, and S. Pantelides, "Self-Consistent Method for Point Defects in Semiconductors: Application to the Vacancy in Silicon", *Physical Review Letters*, 41 (1978), p. 895.

[27] Reference [2] finds the totally antibonding deep level of (S,S) in the gap.

Paper Received 13 February 1988; Revised 24 October 1988.

ELECTRONIC AND STRUCTURAL PROPERTIES OF ORDERED III–V ALLOYS

Kathie E. Newman, Jun Shen, Dan Teng, Bing-Lin Gu,* Shang-Yuan Ren,* and John D. Dow

Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
*Physics Department, Tsinghua University, Beijing, PRC

ABSTRACT

The strain energy of ordered phases of ternary III-V alloys has been estimated. A chalcopyrite structure is found to always have a lower strain energy than 1×1 superlattices oriented along the (0,0,1) or (1,1,1) directions. Ga-As and In-As bond lengths in ordered compounds $GaInAs_2$ are found to be in good agreement with the Mikkelsen-Boyce EXAFS measurements for the alloy. The effect of both ordering and strain on the bandstructure is also considered.

1. INTRODUCTION

Recent reports of new ordered structures in the ternary III-V alloys $Al_{1-x}Ga_xAs$, $GaAs_{1-x}Sb_x$ and $Ga_{1-x}In_xAs$ for $x \simeq \frac{1}{2}$ and $x \simeq \frac{1}{4}$[1] have kindled considerable interest in the semiconductor community. Remembering that a zinc-blende structure is composed of two face-centered cubic (fcc) sublattices, and ignoring the atoms located on the "passive" non-ordering sublattice, one finds that the newly discovered $x = \frac{1}{2}$ semiconductor structures, termed the (0,0,1) 1×1 superlattice, chalcopyrite, and (1,1,1) 1×1 superlattice (see Fig. 1), are all related to the fcc special-k point structures of metals: (0,0,1), $(0,\frac{1}{2},1)$, and (1,1,1), space groups P4/mmm, $I4_1/amd$, and $R\bar{3}m$.[2]

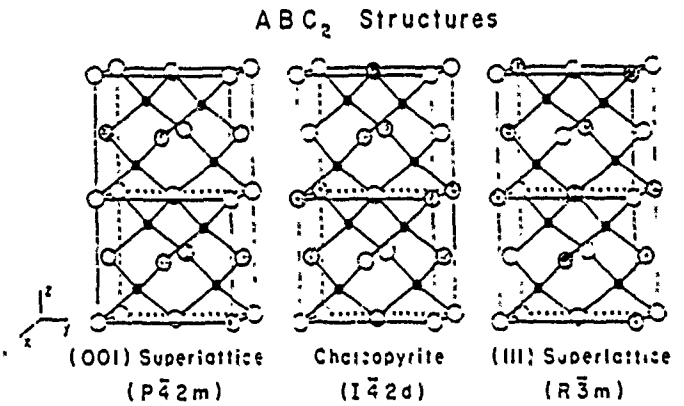


FIGURE 1. Special-k point structures for ordered compounds ABC_2: k = (0,0,1), 1×1 superlattice oriented along the z direction; k = $(0,\frac{1}{2},1)$, chalcopyrite; and k = (1,1,1), 1×1 superlattice oriented along the (1,1,1) direction. Also shown are the respective space-group designations of each crystal type. Note, each structure is shown undistorted from the parent zinc-blende form. Cations A and B are shown as large open and shaded circles, respectively, and the anions C are shown as smaller filled circles.

G. M. Stocks and A. Gonis (eds.), Alloy Phase Stability, 621–625
© 1989 by Kluwer Academic Publishers.

The structure reported for x = ¼, famatinite, shown in Fig. 2, is also a (1 ¼ 0) special-k structure.

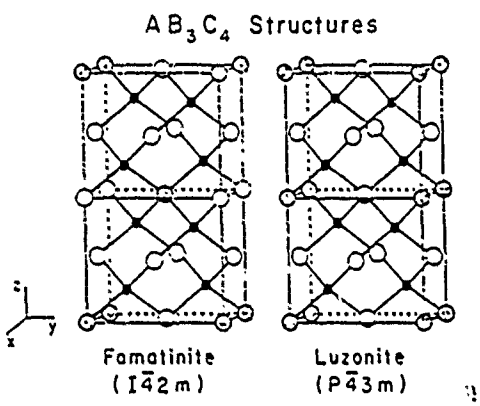


FIGURE 2. Special-k point structures for AB_3C_4 ordered compounds: k = (0,0,1), luzonite, and k = (0,¼,1), famatinite.

Not known presently is the stability of the phases found. All experimental reports have used epitaxial growth, suggesting strongly the possibility of strain stabilization. Mbaye et al.,[3] who have calculated a phase diagram in a tetrahedron approximation, used strain arguments to explain why ordered phases are not normally seen.

The existence of sizable strains in the materials due to the tetrahedral sp^3 bond is clearly the feature that distinguishes the semiconductor ordering problem from that in metals. This paper investigates the effects of strain on the new ordered compounds. Strain energies for the five special-k structures are found. Finally, the effects of strain and order on the bandstructure of the new ordered forms is investigated, using as examples the (0,0,1)-oriented superlattice and chalcopyrite ordered structures.

2. THEORY

2.1 The Effects of Strain on Structure

We estimate the strain energy of a semiconductor from a simple phenomenological formula due to Harrison:[4]

$$E_{strain} = \frac{1}{2} C_0 (\delta d)^2 + \frac{1}{2} C_1 (\delta\theta_{ij})^2 \qquad (1)$$

This formula has two contributions: A bond-stretching term, present because there is a natural bond length between any pair of atoms, and a bond-bending term, indicating that the energy should increase if angles deviate from those in a perfect tetrahedron. The parameters C_0 and C_1 are fit to elastic constants.

Our calculations start by specifying the atomic positions in the ordered structure. For each type of structure, there is an obvious symmetry-allowed distortion that best accommodates the bond-stretching and bond-bending forces. For example, for chalcopyrite, see Fig. 1, the positions of the S atoms in the basis are given in terms of just three parameters, the dimensions a and 2c of the body-centered tetragonal (bct) unit cell, with $c \simeq a \simeq a_0$, with a_0 the dimension of the parent zinc-blende cube, and an internal coordinate variable p ($p \simeq \frac{1}{4}a_0$). Similarly, the superlattice oriented along the (0,0,1) direction has as its parameters the dimensions of the distorted cube, a^2c, and an internal coordinate h ($h \simeq \frac{1}{4}a_0$) that is the distance above the z = 0 plane of an anion C. [The superlattice oriented along the (1,1,1) direction is rhombohedral in symmetry and is characterized by 5 parameters.] Then, for each atom in the basis of the structure, we simply tabulate all contributions to Eq. (1), and then minimize the strain energy with respect to the free parameters. By evaluating Eq. (1) at its minima, we have found both the positions of all atoms as well as the size of the strain energy.

2.2 Bandstructure Calculations

In our calculations of bandstructure, we employ a modified empirical tight-binding Hamiltonian that includes the 5 atomic orbitals s, p_x, p_y, p_z and s^*. Such a Hamiltonian is known to be accurate for the valence bands and the conduction band edge of zinc-blende semiconductors. The parameters of this Hamiltonian have been determined by fitting to experimental data for all III-V zinc-blende compounds by Vogl et al.[6]

Calculations use as input the positions of the atoms in each type of unit cell and follow the standard Slater-Koster method, e.g., strain forces a modification of the direction cosines used in the tight-binding Hamiltonian of the zinc-blende structure.[4] The required tight-binding parameters are found from the parent III-V compounds, e.g., we use Harrison's Law,[4] that the product of nearest-neighbor matrix elements V with the square of the bond length, Vd^2, should be nearly constant. Finally, we have included the effects of valence-band offsets on the ordered-compound's bandstructure,[7] and have concluded that these corrections are small.

3. RESULTS

We have computed the strain energies and effects of strain on five types of ternary III-V ordered compounds.[8] Here we simply quote results for a prototypical semiconductor alloy, $Ga_{1-x}In_xAs$.

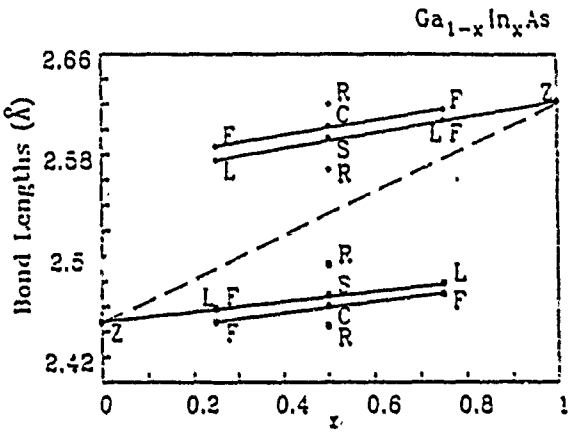


FIGURE 3. Calculated bond lengths in the $Ga_{1-x}In_xAs$ family of compounds.

In Fig. 3 we show results for the Ga-As bond lengths (squares) and In-As bond lengths (diamonds) in the $Ga_{1-x}In_xAs$ family of compounds. For x = 0 or 1, the bond length is that of the zinc-blende (Z) compounds GaAs or InAs. For x = ¼, luzonite compounds (L) contain three Ga-As bonds for every In-As bond, while famatinite compounds (F) have two long and one short Ga-As bonds for each In-As bond. A similar result holds for x = 3/4. For x = ½, the (0,0,1)-oriented superlattice (S) and chalcopyrite compounds have equal numbers of Ga-As and In-As bonds, while the rhomobohedral (R) (1,1,1)-oriented superlattice contains three long and one short Ga-As bond and three short and one long In-As bond.

The dashed line shown in Fig. 3 indicates Vegard's Law expected for the average bond length measured in the alloy by x-ray diffraction. Mikkelsen and Boyce[9] have shown using EXAFS (extended x-ray absorption fine structure) measurements that there are two distinct bond lengths in $Ga_{1-x}In_xAs$ alloys, one for each type of bond. The solid lines in Fig. 3 that pass through the parent zinc-blende compounds are in good agreement with the slopes of the Mikkelsen-Boyce data for the alloy. These lines pass through compounds with either simple cubic or simple tetragonal symmetry. [Parallel to those lines are lines passing through the body-centered tetragonal compounds, famatinite, and chalcopyrite.] We interpret this agreement as follows. The alloy is to be thought of as a mixture of all possible orientations of tetrahedra $AsGa_nIn_{4-n}$. Portions of the alloy may show incipient order of the types shown in Figs.

1 and 2. The EXAFS data are statistical averages of the configurations found in the alloy; a line near the middle of the distribution of possible bond lengths is found. More importantly, the fact that the average results for the strain calculations alone reproduce well the EXAFS data implies that charge-transfer effects are probably only a small correction for the III-V family of compounds.[10]

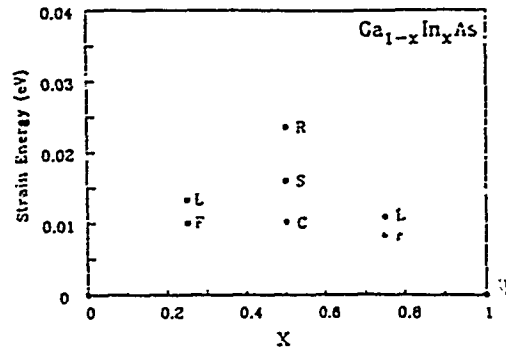


FIGURE 4. Calculated strain energies of the $Ga_{1-x}In_xAs$ family of compounds.

Shown in Fig. 4 are the calculated strain energies of the ordered structures. For x = 1/4 and 3/4, we find that the famatinite structure has the lowest energy, while for x = ½, the chalcopyrite structure is lowest in energy, followed then by the (0,0,1) superlattice (S) and then by the rhombohedral-symmetry (R) superlattice oriented along the (1,1,1) direction. These results are consistent with the calculated bond lengths shown in Fig. 3. For example, for rhombohedral symmetry, there are two types of Ga-As (In-As bonds). One bond, found in a tetrahedron $GaIn_3As$, which is near the parent bond length, and the other one, found with a 3-fold degeneracy, in a tetrahedron $GaInAs_3$, at a distance quite far from the parent bond length. It is easily seen from Eq. (1) that those structures with large bond-length deviations from the parent zinc-blende compounds will tend to have a large strain energy. As a general trend, we find that structures with a body-centered tetragonal Bravais lattice most easily accommodate strain.

Turning finally to electronic properties, in Figs. 4 and 5 we show the low-temperature bandstructures of two $GaInAs_2$ ordered structures, the (0,0,1) superlattice and chalcopyrite structure. The band gap of the (0,0,1) superlattice is at 0.87 eV, compared with 0.81 eV in the alloy for x = 0.47,[11] and 0.83 eV in chalcopyrite. Notatable is the strain splitting of the top of the valence band. Because the local strain field of the tetrahedron has opposite orientations in the two structures, with the z direction singled out in the (0,0,1) superlattice, and the xy plane in chalcopyrite (see Fig. 1), the heavy hole is higher in the (0,0,1) superlattice than in chalcopyrite. Obviously, these two structures will have distinguishable experimental spectra.

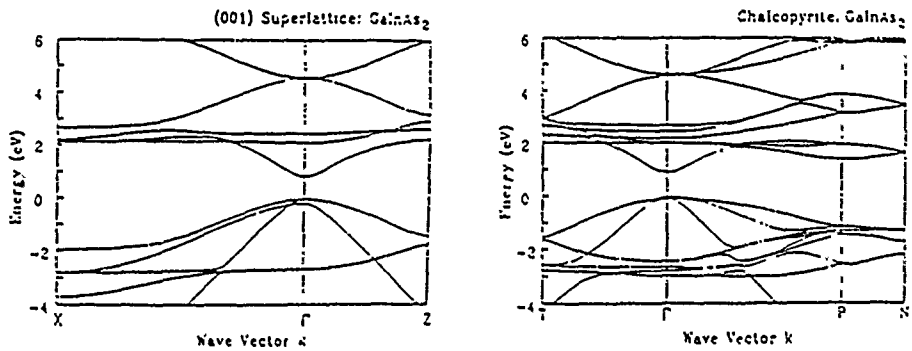


FIGURE 5. Band structures of the (0,0,1)-oriented superlattice and of the chalcopyrite forms of $GaInAs_2$ compounds.

4. CONCLUSIONS

We have investigated the relative stability for III-V compounds of the three possible ordered $x = \frac{1}{2}$ structures: the (0,0,1)- and (1,1,1)-oriented 1×1 superlattices and chalcopyrite. Without using a second-neighbor interaction on an fcc lattice,[2] we find that the strain energy due to the bonding of anions to cations alone is enough to distinguish the three phases. Our calculations predict that the chalcopyrite structure should have a lower energy than either type of superlattice. Similarly, we find that for $x = \frac{1}{4}$, the famatinite structure has lower strain energy than the luzonite structure.

Our strain calculations also allow us to predict structural parameters such as bond lengths for the new ordered III-V compounds. This in turn allows us to investigate the influence of order and strain on the bandstructure of these structures. We hope our calculations will aid experimentalists in the search for additional new ordered semiconductor compounds.

ACKNOWLEDGMENTS — We thank ONR and AFOSR for their financial support.

REFERENCES

1. T. S. Kuan, T. F. Kuech, W. I. Wang, and E. L. Wilkie, Phys. Rev. Lett. 54, 201 (1985); H. R. Jen, M. J. Cherng, and G. B. Stringfellow, Appl. Phys. Lett. 48, 1603 (1986); H. Nakayama and H. Fujita, in *Gallium Arsenide and Related Compounds -- 1985*, edited by M. Fujimoto, IOP Conference Proceedings No. 79 (Institute of Physics, Bristol and London, 1986), p. 289; and M. A. Shahid, S. Mahajan, D. E. Laughlin, and H. M. Cos, Phys. Rev. Lett. 58, 2567 (1987).
2. For a review, see D. de Fontaine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1979). Vol. 34, p. 73.
3. See A. A. Mbaye, L. G. Ferreira, and A. Zunger, Phys. Rev. Lett. 58, 49 (1987) and references quoted therein.
4. W. A. Harrison, *Electronic Structure and the Properties of Solids* (W. H. Freeman, San Francisco, 1980), p. 76, 196, and 481.
5. U. Kaufman and J. Schneider, *Festkörperprobleme XIV*, edited by H. J. Queisser (Pergamon, 1974) p. 229.
6. P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
7. E. A. Kraut, J. Vac. Sci. Technol. B2, 488 (1984).
8. D. Teng, J. Shen, K. E. Newman, B.-L. Gu, S. Y. Ren, and J. D. Dow, unpublished.
9. J. C. Mikkelsen and J. B. Boyce, Phys. Rev. Lett. 49, 1412 (1982).
10. K. C. Hass and D. Vanderbilt, J. Vac. Sci. Technol. A5, 3019 (1987).
11. K. Alavi, R. L. Aggarwal, and S. H. Groves, Phys. Rev. B21, 1311 (1980); measurement is at 4.2 K.

From: PROPERTIES OF IMPURITY STATES IN SUPERLATTICE SEMICONDUCTORS
Edited by C. Y. Fong, Inder P. Batra, and S. Ciraci
(Plenum Publishing Corporation, 1988)

DEEP IMPURITY LEVELS IN SEMICONDUCTORS, SEMICONDUCTOR ALLOYS, AND SUPERLATTICES

John D. Dow, Shang Yuan Ren, and Jun Shen

Department of Physics
University of Notre Dame
Notre Dame, Indiana 46556 U.S.A.

INTRODUCTION

Impurity levels determine the electrical properties of semiconductors and often strongly influence the optical properties as well. Until rather recently it was widely believed that "shallow impurities," namely those impurities that produce energy levels within ≈0.1 eV of a band edge, were well understood in terms of hydrogenic effective-mass theory [1]. However "deep impurities" were regarded as more mysterious, having levels more than 0.1 eV deep in the gap; and several theoretical attempts were made to understand why their binding energies were so large. While specific deep levels were explained rather well by the early theories, most notably the pioneering work of Lannoo and Lenglart on the deep level in the gap associated with the vacancy in Si [2], numerous attempts to explain why the binding energy of a particular level might be large (making the level deep), rather than small, continued until recently, when it was realized that this basic picture of impurity levels was incomplete [3,4].

DISTINCTION BETWEEN DEEP LEVELS AND DEEP IMPURITIES

Now it is recognized that every impurity whose valence differs from that of the host atom it replaces produces both deep and shallow levels, but that the "deep" levels often are not deep energetically. The definitions of deep and shallow levels have been changed [3]: now a deep level is one that originates from the central-cell potential of the defect, and a shallow level originates from the long-ranged Coulomb potential due to the impurity-host valence difference. Impurities that are s-p bonded normally produce four deep levels in the vicinity of the fundamental band gap: one s-like and three p-like. More often than not, the deep levels do not lie in the band gap but are resonant with the host bands. Thus, in terms of the old picture, these resonant levels have negative "binding energies." The four deep levels are normally anti-bonding and host-like in character. As a result, several different deep impurities produce levels with essentially the same wavefunctions [5,6].

In contrast to the deep levels, whose wavefunctions are relatively localized and have multi-band character, shallow levels have wavefunctions that are extended in real space and of single-band character (localized near

a band extremum in $\vec{k}$-space). The shallow levels are hydrogenic, originate from the long-ranged Coulomb potential, are associated with the impurity-host valence difference, and have binding energies of reduced Rydbergs, i.e., scaled from 13.6 eV by factors of effective mass and the inverse square of the dielectric constant (resulting in binding energies of typically tens of meV). Shallow levels can be obtained by solving an effective-mass Schrödinger equation for the envelope wavefunction $\phi(r)$, which is (assuming isotropic effective mass m* and dielectric constant ϵ)

$$[(-\hbar^2/2m^*)\nabla^2 - (e^2/\epsilon r)]\phi = (E-E_0)\phi, \qquad (1)$$

where E_0 is the band extremum to which the level is "attached."

The attachment to a near-by band edge is the experimental signature of a shallow level, and can be verified by measuring the energy of the level and the band edge, either versus alloy composition x in an alloy such as $Al_xGa_{1-x}As$ or $GaAs_{1-x}P_x$, or versus pressure. In contrast to shallow levels, deep levels are not attached to band edges, and their energies (with respect to the valence band maximum) tend to vary more-or-less linearly with alloy composition.

DEPENDENCE ON ALLOY COMPOSITION

Fig. 1 summarizes data [7] for the alloy-composition dependence of levels associated with anion-site impurities in $GaAs_{1-x}P_x$ alloys. The levels associated with S, shallow donor levels, follow the conduction band edge as the composition varies, and are "attached" to it. Se levels exhibit the same behavior as S. The N and O levels vary approximately linearly with alloy composition and are unattached to any band edge. These data, when first obtained, were rather perplexing, because O, S, and Se were expected to behave the same, since all came from Column-VI of the Periodic Table, and all were presumed to occupy a Column-V anion site. By the old definition, the oxygen state was clearly a deep level, lying more than 0.1 eV from the conduction band edge. The N level, however, had the same qualitative dependence on alloy composition as the O state (indicating it should be classified as deep), yet it was close enough to the band edge in GaP to be classified as shallow by the old definition of a shallow level. Moreover it was energetically deep in the gap for $GaAs_{0.5}P_{0.5}$, but invisible (possibly in the conduction band) for GaAs.

The N level data in $GaAs_{1-x}P_x$ provided an important clue for understanding the physics of deep impurities, because N is isoelectronic to As and P, and so the substitutional impurity has no long-ranged Coulomb potential: its level could only be explained in terms of the central-cell defect potential (and, possibly also the modest strain field around the impurity) [4]. The natural explanation for the data of Fig. 1 is that the central-cell potential produces the oxygen deep level and that the N level is similar to the oxygen level, and should also be termed "deep." The shallow S (and Se) levels are caused by the Coulomb potential associated with the valence difference between the Column-VI impurities and the Column-V host. S and Se must also have deep levels similar to the oxygen deep level, but their deep levels must be resonant with the conduction band of $GaAs_{1-x}P_x$ for all alloy compositions x, and hence are invisible (as with the case for N when $x<0.2$). Thus the picture has emerged that the central-cell potential produces deep levels, and the Coulomb potential yields shallow levels (which ideally are infinite in number, because a Coulomb potential has an infinite number of bound states). For s- and p- bonded substitutional impurities, we expect one s-like and three p-like deep levels associated with every band, based on Rayleigh's interlacing theorems [8]; this means that in the vicinity of the

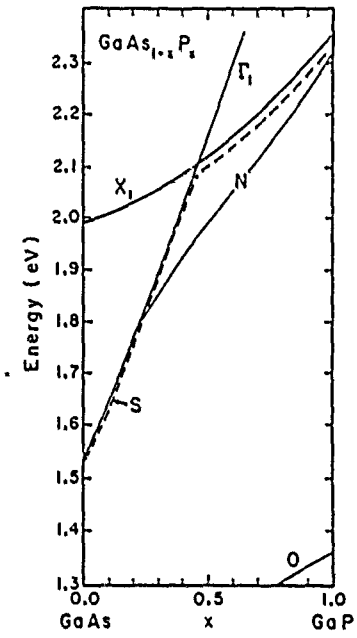


Fig. 1. Schematic illustration of the dependences of the shallow S (or Se) level (dashed) and the deep N and O levels (solid) on alloy composition x in $GaAs_{1-x}P_x$, after Ref. [7]. The energies of the band edges Γ_1 and X_1 are also shown. The zero of energy is the valence band maximum of the alloy.

band gap of a zinc-blende semiconductor we should find one s-like (A_1-symmetric) deep level and one triply degenerate p-like (T_2-symmetric) level for each impurity. The oxygen and nitrogen levels of Fig. 1 are the A_1-symmetric deep levels. In a three-dimensional material, there is no guarantee that any of the deep levels due to the central-cell potential will lie in the fundamental band gap, and so, as in the case of S and Se in $GaAs_{1-x}P_x$, the deep levels may all lie outside of the band gap or else, as in the case of oxygen, one (or more) may lie within the gap. (Fig. 2)

The character of an impurity as "shallow" or "deep" is determined by whether one or more deep levels lies within the fundamental band gap. For example, oxygen in GaP is a deep impurity because its s-like A_1 deep level lies in the band gap, and is occupied by one electron when the oxygen is neutral and in its ground state. This one-electron level can hold two electrons of opposite spin, but when oxygen is neutral, it holds only one, (the extra electron due to the valence difference between O and P). In the case of S or Se, the central-cell potential is weaker than that of O, and the corresponding deep level is degenerate with the conduction band rather than in the gap. The resonant level cannot bind the extra electron, which spills out and decays to the conduction band edge (via the electron-phonon interaction); once at the conduction band minimum, the electron is trapped by the long-ranged Coulomb potential in a shallow impurity level -- and, because the ground state of the S or Se impurity in this host has an electron in the shallow level, the S or Se impurity is termed a shallow impurity.

DEEP-SHALLOW TRANSITIONS

It is possible to change the character of an impurity from deep to shallow by perturbing the host until a deep level passes from the band gap into the host bands. Straightforward ways to achieve such a change are to apply pressure or to vary the alloy composition of the host. N changes its character in $GaAs_{1-x}P_x$ as a function of alloy composition x: it yields a deep level in the gap for $x>0.2$ and so is a deep impurity for such alloy compositions, but for $x<0.2$, neutral N is neither a shallow impurity (because it has no long-ranged Coulomb potential) nor a deep impurity (because its deep levels are in the conduction band), and so we term it an "inert impurity." Thus, in $GaAs_{1-x}P_x$ for $x\approx0.22$, N undergoes a "deep-inert transition." See Fig. 1. A classic example of a similar "deep-shallow transition" is Si on a cation site in $Al_xGa_{1-x}As$, which has a deep-level behavior similar to that of N in $GaAs_{1-x}P_x$. (See Fig. 3.) As a result, neutral Si is a deep impurity, with one electron occupying its A_1-symmetric deep level for $x\geq0.3$, but is a shallow impurity for $x<0.2$ because the deep level is resonant with the conduction band and the electron is autoionized, falls to the conduction band edge, and is trapped by the Coulomb potential of the defect. This Si defect is thought to be the DX center [9-11], or at least a component of it [11,12]. It is noteworthy that when cation-site Si is a shallow impurity (i.e., its deep levels are resonant with the conduction band), it dopes the material n-type. But when it is a deep impurity (with its A_1 level in the gap) the neutral impurity can trap either an electron or a hole in this level, and so tends to make the material semi-insulating.

DEEP-SHALLOW TRANSITIONS IN SUPERLATTICES

Deep-shallow transitions occur in random alloys such as $Al_xGa_{1-x}As$ because the deep levels are much less sensitive to changes of alloy composition x than the band edges. Only slightly overstating the point, we may say that the deep levels are almost constant in absolute energy, and that varying x can cause the conduction or valence band edge to move through the deep level. Since deep levels are normally only observed when they lie within the fundamental band gap, varying alloy composition changes the window of

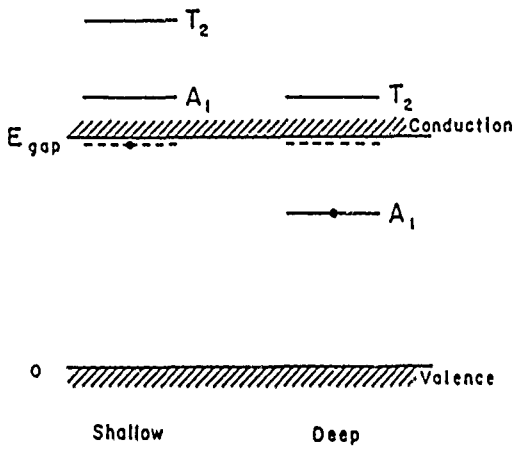


Fig. 2. Illustration of the difference between a shallow impurity such as S on a P site in GaP and a deep impurity such as O. Shallow levels are dashed and deep levels are solid. If the deep levels lie outside of the band gap, the extra electron occupies the lowest shallow level at zero temperature and the impurity is a shallow donor.

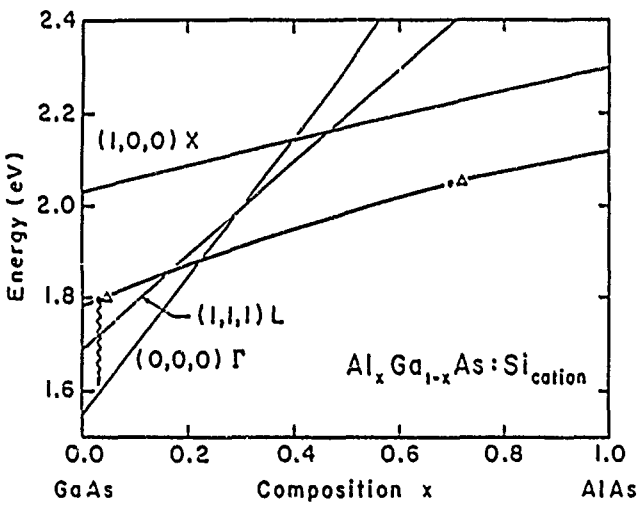


Fig. 3. Chemical trends with alloy composition x in the energies (in eV) of principal band gaps at Γ, L, and X, with respect to the valence band maximum of the alloy, in the alloy $Al_xGa_{1-x}As$, as deduced from the Vogl model [15]. Also shown is the predicted energy of the A_1-symmetric cation-site deep level of Si (heavy line), similar to the predictions of Hjalmarson [11]. The Vogl model is known to obtain very little band bending. Moreover the L minimum for x=0.45 is known to be at a bit too low an energy in this model. When the deep level of neutral Si lies below the conduction band minimum, it is occupied by one electron (solid circle) and one hole (open triangle). When this level is resonant with the conduction band, the electron spills out and falls (wavy line) to the conduction band minimum, where it is trapped (at zero temperature) in a shallow donor level (not shown).

observability of deep levels, namely the band gap. Thus selecting the alloy composition so that a given impurity produces a shallow level rather than a deep one is a form of "band-gap engineering."

Similar band-gap engineering can be achieved by altering the atomic ordering in the semiconductor by, for example, making the host a superlattice rather than a random alloy. For example, Si on a Ga site in the center of a GaAs quantum well in a $N_1 \times N_2$ GaAs/$A\ell_xGa_{1-x}As$ superlattice can be made to undergo a shallow-deep transition by reducing the thickness of the quantum well until the conduction band maximum of the superlattice passes above the Si deep level (See Fig. 4.). The dependence on alloy composition of a Ga-site deep level in a 2×2 GaAs/$A\ell_xGa_{1-x}As$ superlattice is given in Fig. 5, and shows that even near the center of a GaAs layer Si can become a deep trap if the layer is sufficiently thin or two-dimensional.

Fig. 6 shows the predictions for the cation-site deep level in bulk GaAs, in a 3×10 GaAs/$A\ell_{0.7}Ga_{0.3}As$ superlattice, and in bulk $A\ell_{0.7}Ga_{0.3}As$ for Si at various sites β in the superlattice. In this superlattice the deep level lies in the gap and Si is therefore a deep impurity. In $N_1 \times N_2$ superlattices such that the GaAs layers are not very thin, $N_1 > 6$, the superlattice conduction band edge approaches the bulk GaAs conduction band edge, and therefore the conduction band of the superlattice covers up the Si deep level, making Si on a Ga site a shallow donor in the superlattice.

T_2 LEVELS

The examples we have given thus far all consider only A_1-symmetric bulk deep levels. The T_2 bulk levels are split in a (001) superlattice into a_1, b_1, and b_2 levels, as illustrated in Fig. 7 for the As vacancy. The valence band maximum is also split into a $(p_x \pm p_y)$-like (or b_1- or b_2-like) maximum and a p_z-like (or a_1-like) edge slightly below it. For sites near the β=0 interface, the b_1 deep levels have orbitals and energies similar to the T_2 levels of bulk GaAs and the b_2 levels are $A\ell_xGa_{1-x}As$-like. The splittings are small even for impurities at the interface, of order 0.1 eV, and decrease rapidly as the impurity moves from the interface. They are non-zero even near the center of the layers, however, because the valence band maximum is split, and so the spectral densities employed in solving Eq. (2) below are also split.

Three factors influence the positions of deep levels in the superlattice layers: (i) the band offset, which tends to move the levels down in $A\ell_xGa_{1-x}As$ relative to in GaAs, (ii) quantum confinement, which tends to cause levels in GaAs and the p_z-like a_1 levels in particular to be farther from the center of the gap, and (iii) the relative electropositivity of $A\ell$ with respect to Ga, which tends to move energy levels up when the associated wavefunctions overlap an $A\ell$-rich layer. To determine the balance among these competing effects, a calculation is required.

FORMALISM

Our basic approach to the problem of deep levels in superlattices is based on the theory of Hjalmarson et al. [3]. This theory is relatively easy to implement and learn, and has been summarized in accessible lecture notes [13] for deep levels in the bulk. Nelson et al. [14] have provided an alternative approach to deep levels in superlattices which is especially well-suited to small-period superlattices.

The deep level energies E in a superlattice can be computed similarly by solving the one-electron Schrödinger equation, using the Green's function method to take advantage of the localized nature of the central-cell defect

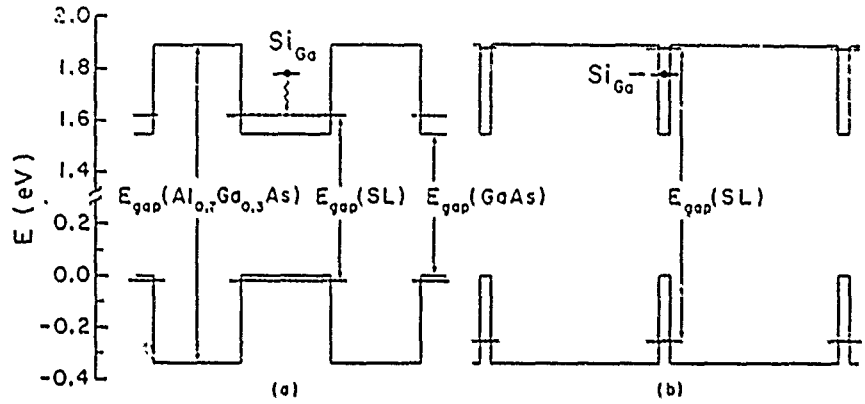


Fig. 4. Schematic illustration of the quantum-well confinement effect on the band gap $E_{gap}(SL)$ of a $N_1 \times N_2$ GaAs/$Al_{0.7}Ga_{0.3}As$ superlattice, after Ref. [12]: (a) $N_1=N_2=18$; and (b) $N_1=2$, $N_2=34$. The band edges of the superlattice are denoted by chained lines. For this alloy composition the superlattice gap is indirect for case (b), with the conduction band edge at $\vec{k}=(2\pi/a_L)(1/2,1/2,0)$. Note the broken energy scale. The zero of energy is the valence band maximum of GaAs.

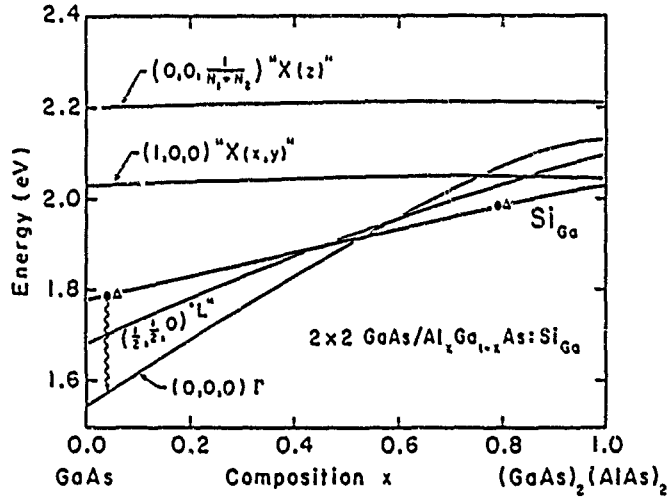


Fig. 5. Predicted chemical trends with alloy composition x in the energies (in eV) of principal band gaps and the Si A_1-symmetric deep level in a $(GaAs)_2(Al_xGa_{1-x}As)_2$ superlattice. Compare with Fig. 3 for the alloy. The superlattice wavevectors of the gaps are $\vec{k}=\vec{0}$, $\vec{k}=(2\pi/a_L)(1/2,1/2,0)$, which has states derived from the L point of the bulk Brillouin zone, and the points derived from the bulk X-point: $(2\pi/a_L)(0,0,[N_1+N_2]^{-1})$, (where $N_1=N_2=2$), and $(2\pi/a_L)(1,0,0)$.

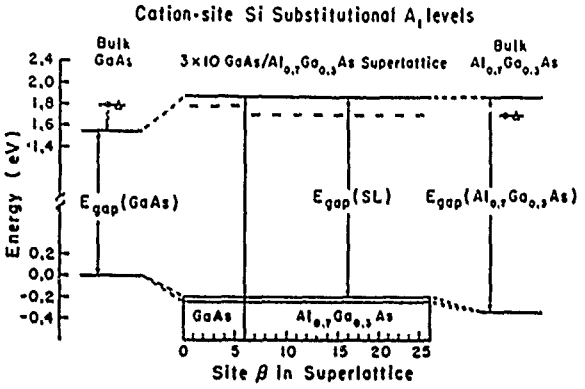


Fig. 6. Predicted A_1-derived deep levels of Si in GaAs, in a 3×10 GaAs/$Al_{0.7}Ga_{0.3}As$ superlattice (as a function of β, the position of the Si in the superlattice), and in bulk $Al_{0.7}Ga_{0.3}As$. Interfaces (which are As sites) correspond to β = 0, 6, and 26. Note that in bulk GaAs Si is a shallow donor, but that in this superlattice and in bulk $Al_{0.7}Ga_{0.3}As$ it is predicted to be a deep impurity.

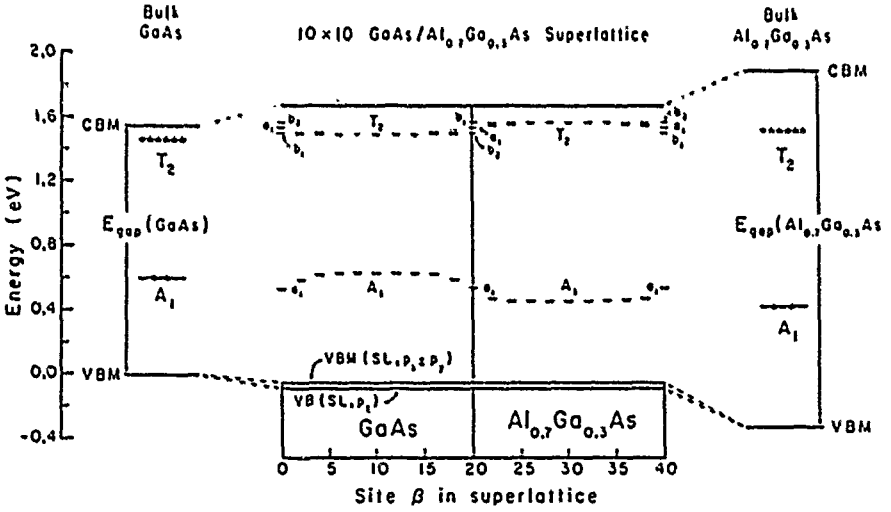


Fig. 7. Predicted energy levels of an As-vacancy in bulk GaAs, in a $(GaAs)_{10}(Al_{0.7}Ga_{0.3}As)_{10}$ superlattice (as a function of β, the position of the vacancy: even values of β correspond to As-sites), and in bulk $Al_{0.7}Ga_{0.3}As$, after Ref. [12]. Note the splitting of the T_2 levels at and near the interfaces which correspond to β = 0, 20, and 40. The electron (hole) occupancies of the deep levels in bulk GaAs and bulk $Al_{0.7}Ga_{0.3}As$ are denoted by solid circles (open triangles).

potential operator V:

$$\det\,(1-G(E)V) = 0 = \det\,(1 - P\int[\delta(E'-H)\,V/(E-E')]\,dE'). \tag{2}$$

Here we have $G(E) = (E-H)^{-1}$, where H is the host Hamiltonian operator, P denotes the principal value integral over all energies, and $\delta(x)$ is Dirac's delta function. For simplicity of presentation, we consider a periodic $N_1 \times N_2$ $GaAs/A\ell_xGa_{1-x}As$ superlattice grown in the (001) direction. This superlattice has N_1 two-atom thick layers of GaAs and N_2 two-atom thick layers of $A\ell_xGa_{1-x}As$ stacked alternately in a periodic structure. We assume that the layers are perfectly lattice-matched. We describe the superlattice Hamiltonian in terms of an s^*sp^3 basis similar to that of Vogl et al. [15]. Our Hamiltonian is a nearest-neighbor tight-binding model, and, in the limit of x = 0, is identical to the Vogl model for GaAs. We treat the superlattice using a superhelix or supercell method: (for the case x = 1) a superhelix or supercell is a helical string with axis aligned along the (001)- or z-direction consisting of $2N_1 + 2N_2$ adjacently bonded atoms As, Ga, As, Ga, As, ... Ga, As, Aℓ, As, Aℓ, As, Aℓ, ... As, Aℓ. (For x ≠ 1, replace Aℓ by the virtual cation $A\ell_xGa_{1-x}$.) The center of this helix is at $\bar{L}$ and each of the atoms of the helix is at position $\bar{L} + \bar{v}_\beta$ (for β=0, 1, 2, ..., $2N_1+2N_2-1$). The superlattice is a stack of superslabs. A superslab of $GaAs/A\ell_xGa_{1-x}As$ consists of all such helices with the same value of L_z and all possible different values of L_x and L_y. The origin of coordinates is taken to be at an As atom and a neighboring cation is at $(1/4,1/4,1/4)a_L$, where a_L is the lattice constant. At each site there are five s^*sp^3 basis orbitals $|n,\bar{L},\bar{v}_\beta)$, where n = s*, s, p_x, p_y, or p_z and the site is specified by $\beta = 0,1,2,\ldots,2N_1+2N_2-1$. From these basis orbitals we form the sp^3 hybrid orbitals at each site $\bar{R} = (\bar{L},\bar{v}_\beta)$. The hybrid orbitals are

$$|h_1,\bar{R}) = [|s,\bar{R}) + \lambda|p_x,\bar{R}) + \lambda|p_y,\bar{R}) + \lambda|p_z,\bar{R})]/2$$

$$|h_2,\bar{R}) = [|s,\bar{R}) + \lambda|p_x,\bar{R}) - \lambda|p_y,\bar{R}) - \lambda|p_z,\bar{R})]/2$$

$$|h_3,\bar{R}) = [|s,\bar{R}) - \lambda|p_x,\bar{R}) + \lambda|p_y,\bar{R}) - \lambda|p_z,\bar{R})]/2$$

and

$$|h_4,\bar{R}) = [|s,\bar{R}) - \lambda|p_x,\bar{R}) - \lambda|p_y,\bar{R}) + \lambda|p_z,\bar{R})]/2 \tag{3}$$

where we have $\lambda = +1$ (-1) for atoms at anion (cation) sites. Introducing the label ν = s*, h_1, h_2, h_3, or h_4, our hybrid basis orbitals are $|\nu,\bar{R})$, and the related tight-binding orbitals [16] are

$$|\nu,\beta,\bar{k}) = N_s^{-1/2}\,\Sigma_{\bar{L}}\,\exp\,(ik\cdot\bar{L}+i\bar{k}\cdot\bar{v}_\beta)\;|\nu,\bar{L},\bar{v}_\beta) \tag{4}$$

where $\bar{k}$ is (in a reduced zone scheme) any wavevector of the mini-zone or (in an extended zone scheme) any wavevector of the zinc-blende Brillouin zone. Here N_s is the number of supercells.

The mini-zone wavevector $\bar{k}$ is a good quantum number. Evaluation of the matrix elements $(\nu,\beta,\bar{k}|H|\nu',\beta',\bar{k})$ leads to a tight-binding Hamiltonian of the block tridiagonal form. For example, the diagonal (in β) 5×5 matrix, $H(\beta,\beta)$ at site β, is

$$H(\beta,\beta)=(\nu,\beta,\vec{k}|H|\nu',\beta,\vec{k}) = \begin{pmatrix} \epsilon_{s*} & 0 & 0 & 0 & 0 \\ 0 & \epsilon_h & T & T & T \\ 0 & T & \epsilon_h & T & T \\ 0 & T & T & \epsilon_h & T \\ 0 & T & T & T & \epsilon_h \end{pmatrix} . \tag{5}$$

where we have the hybrid energy

$$\epsilon_h = (\epsilon_s + 3\epsilon_p)/4 \tag{6}$$

and the hybrid-hybrid interaction

$$T = (\epsilon_s - \epsilon_p)/4 . \tag{7}$$

The energies ϵ_{s*}, ϵ_h, and T in $H(\beta,\beta)$ refer to the atom at the β-th site, and are obtained from the energies w tabulated by Vogl et al. [15]. To account for the observed [17] valence band edge discontinuity of 32% of the direct band gap [18], a constant is added to ϵ_{s*} and ϵ_h for $Al_xGa_{1-x}As$; this constant is adjusted to give the valence band maximum of $Al_xGa_{1-x}As$ below the valence band maximum of GaAs by 32% of the direct band gap difference in the limit $N_1=N_2 \to \infty$.

Expressions for the remaining, off-diagonal matrix elements $(\nu,\beta,\vec{k}|H|\nu',\beta',\vec{k})$ or $H(\beta,\beta')$ are given in Ref. [12]. The Hamiltonian matrix has dimension $5(2N_1+2N_2)$ for each $\vec{k}$. We diagonalize this Hamiltonian numerically for each (special-point) $\vec{k}$, finding its eigenvalues $E_{\gamma,\vec{k}}$ and the projections of the eigenvectors $|\gamma,\vec{k}\rangle$ on the $|\nu,\beta,\vec{k})$ hybrid basis: $(\nu,\beta,\vec{k}|\gamma,\vec{k}\rangle$. Here γ is the band index (and ranges from 1 to 200 for $N_1=N_2=10$) and $\vec{k}$ lies within the mini-Brillouin zone in a reduced zone scheme. With these quantitites, we can evaluate Eq. (2) for the Hjalmarson model of the defect potential matrix V [3]: the matrix is zero except at the defect site and diagonal on that site, $(0, V_s, V_p, V_p, V_p)$, in the Vogl s^*sp^3 local basis centered on each atom. The point-group for a general substitutional defect in a $GaAs/Al_xGa_{1-x}As$ (001) superlattice is C_{2v}, provided the $Al_xGa_{1-x}As$ is treated in a virtual crystal approximation. In the $GaAs/Al_xGa_{1-x}As$ superlattice the A_1 and T_2 deep levels of the bulk GaAs or $Al_xGa_{1-x}As$ produce two a_1 levels (one s-like, derived from the A_1 level and one T_2-derived p_z-like), one b_1 level [(p_x+p_y)-like], and one b_2 level [(p_x-p_y)-like]. Of course, for impurities far from a $GaAs/Al_xGa_{1-x}As$ interface, the s-like a_1 level will have an energy very close to the energy of a bulk A_1 level, and p_z-like a_1 level and the b_1 and b_2 levels will lie close to the bulk T_2 level also.

The secular equation, Eq. (2), is reduced by symmetry to the following three equations:

$$G(b_1;E) = V_p^{-1} \quad \text{for } b_1 \text{ levels,} \tag{8}$$

$$G(b_2;E) = V_p^{-1} \quad \text{for } b_2 \text{ levels,} \tag{9}$$

and

$$\det \begin{pmatrix} G(s,s;E)\,V_s - 1 & G(s,z;E)\,V_p \\ G(z,s;E)\,V_s & G(z,z;E)\,V_p - 1 \end{pmatrix} = 0, \tag{10}$$

for a_1 levels, where we have

$$G(b_1;E)=\Sigma_{\gamma,\vec{k}}\ |(h_1,\beta,\vec{k}|\gamma,\vec{k}\rangle-(h_4,\beta,\vec{k}|\gamma,\vec{k}\rangle|^2/2(E-E_{\gamma,\vec{k}}), \tag{11}$$

$$G(b_2;E)=\Sigma_{\gamma,\vec{k}}\ |(h_2,\beta,\vec{k}|\gamma,\vec{k}\rangle-(h_3,\beta,\vec{k}|\gamma,\vec{k}\rangle|^2/2(E-E_{\gamma,\vec{k}}), \tag{12}$$

$$G(s,s;E) =$$

$$\Sigma_{\gamma,\vec{k}}\ |(h_1,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_2,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_3,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_4,\beta,\vec{k}|\gamma,\vec{k}\rangle|^2/4(E-E_{\gamma,\vec{k}}), \tag{13}$$

$$G(z,z;E) =$$

$$\Sigma_{\gamma,\vec{k}}\ |(h_1,\beta,\vec{k}|\gamma,\vec{k}\rangle-(h_2,\beta,\vec{k}|\gamma,\vec{k}\rangle-(h_3,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_4,\beta,\vec{k}|\gamma,\vec{k}\rangle|^2/4(E-E_{\gamma,\vec{k}}), \tag{14}$$

and

$$G(s,z;E)=\Sigma_{\gamma,\vec{k}}\ [(h_1,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_2,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_3,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_4,\beta,\vec{k}|\gamma,\vec{k}\rangle]$$
$$\times\ [(h_1,\beta,\vec{k}|\gamma,\vec{k}\rangle-(h_2,\beta,\vec{k}|\gamma,\vec{k}\rangle-(h_3,\beta,\vec{k}|\gamma,\vec{k}\rangle+(h_4,\beta,\vec{k}|\gamma,\vec{k}\rangle]^*/[4(E-E_{\gamma,\vec{k}})] \tag{15}$$

Here $G(z,s;E)$ is the Hermitian conjugate of $G(s,z;E)$ and β is the site of the defect in the superlattice.

For each site β the relevant host Green's functions, Eqs. (11) to (15), are evaluated using the special points method [19], and the secular equations (8) to (10) are solved, yielding $E(b_1;V_p)$, $E(b_2;V_p)$, and two values of $E(a_1;V_s,V_p)$. The defect potential matrix elements V_s and V_p are obtained using a slight modification of Hjalmarson's approach [20]. For $N_1=N_2=10$, there are 40 possible sites β, each with four relevant deep levels: two a_1, one b_1, and one b_2; thus there are 160 levels.

SUMMARY

With the basic approach outlined here, one can compute the deep levels associated with substitutional impurities in lattice-matched superlattices and predict the conditions under which deep-shallow and deep-inert transitions are to be expected.

Acknowledgment -- We are grateful to the U.S. Office of Naval Research (Contract No. N00014-84-K-0352) and the U.S. Air Force Office of Scientific Research (Contract No. AF-AFOSR-85-0331) for their generous support.

REFERENCES

[1] W. Kohn, in Solid State Physics (edited by F. Seitz and D. Turnbull, Academic Press, New York, 1957) Vol. 5, p. 258-321; J. M. Luttinger and W. Kohn, Phys. Rev. 97, 969 (1955).
[2] M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969).
[3] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Letters 44, 810 (1980); see also Ref. [4] for the concepts that form the foundation of this work.

[4] W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).
[5] S. Y. Ren, W. M. Hu, O. F. Sankey, J. D. Dow, Phys. Rev. B 26, 951 (1982).
[6] S. Y. Ren, Scientia Sinica 27, 443 (1984).
[7] D. J. Wolford, W. Y. Hsu, J. D. Dow, and B. G. Streetman, J. Lumin. 18/19, 863 (1979).
[8] A. A. Maradudin, E. W. Montroll, and G. H. Weiss, Solid State Phys. Suppl. 3, 132 (1963).
[9] H. P. Hjalmarson and T. J. Drummond, Appl. Phys. Letters 48, 657 (1986); see also Ref. [10].
[10] J. C. M. Henning and J. P. M. Ansems, Mat. Sci. Forum 10-12, 429 (1986); Semicond. Sci. Technol. 2, 1 (1987); A. K. Saxena, Solid State Electron. 25, 127 (1982).
[11] See also, J. W. Farmer, H. P. Hjalmarson, and G. A. Samara, Proc. Mater. Res. Soc., 1987, to be published.
[12] S. Y. Ren, J. D. Dow, and J. Shen, "Deep impurity levels in semiconductor superlattices," to be published. See also Ref. [21].
[13] J. D. Dow, in Highlights of Condensed Matter Theory (Proc. Intl. School of Phys. "Enrico Fermi" Course 89, Varenna, 1983) ed. by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna, Italy, and North Holland, Amsterdam, 1985), pp. 465 et seq.
[14] J. S. Nelson, C. Y. Fong, I. P. Batra, W. E. Pickett, and B. M. Klein, unpublished.
[15] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[16] The tight-binding formalism we use here is hybrid-based. For bulk semiconductors the hybrid-based tight-binding formalism is equivalent to the widely used atomic-orbital based tight-binding formalism. But for superlattices these two formalisms are different: because the atomic-orbital based tight-binding parameters are obtained by fitting to the band structure of bulk GaAs and bulk AlAs, the tight-binding parameters for the As atom at β=0 (which is considered to be an As atom in GaAs) are usually different from the corresponding tight-binding parameters for the atom at β=2N$_1$ (which is considered to be an As atom in AlAs). But both of these As atoms are interfacial As atoms and physically are completely equivalent to each other. The way we correct for this problem is that we consider the hybrids h_1 and h_4 of the As atom at β=0 to be hybrids of GaAs, while h_2 and h_3 are taken to be AlAs hybrids. Similarly, h_2 and h_3 of the As atom at β=2N$_1$ are considered to be hybrids in GaAs, and h_1 and h_4 are AlAs hybrids. We believe that this properly accounts for the correct physics and the nature of interfacial bonds.
[17] D. J. Wolford, T. F. Kuech, J. A. Bradley, M. A. Grell, D. Ninno, and M. Jaros, J. Vac. Sci. Technol. B4, 1043 (1986).
[18] In our model, which uses low-temperature band gaps, we take the valence band offset for $Al_{0.7}Ga_{0.3}As$ to be 0.334 eV below the valence band edge of GaAs.
[19] S. Y. Ren and J. D. Dow, "Special points for superlattices and strained bulk semiconductors," Phys. Rev. B, in press.
[20] The defect potential matrix elements are related to the difference in atomic energies of the impurity and the host atom it replaces:

$$V_\ell = \beta_\ell \, (w_{\text{impurity}}(\ell) - w_{\text{host}}(\ell)) + C_\ell,$$

where we have ℓ = s or p, β_s = 0.8, and β_p = 0.6, and the atomic orbital energies w can be found in Table 3 of Ref. [15]. The constant C_ℓ is zero in Ref. [3], but here we take C_s = 1.434 eV in order to have the Si deep level in GaAs appear 0.234 eV above the conduction band minimum, where

it is observed [12], instead of 0.165 eV below it, thereby compensating for the small theoretical uncertainty.

[21] R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Proc. Materials Research Soc. 77, 545-550 (1987), Interfaces, Superlattices, and Thin Films, ed. J. D. Dow and I. K. Schuller.

Electronic structure and deep impurity levels in [111] GaAs/$Al_xGa_{1-x}As$ semiconductor superlattices

Shang-Yuan Ren[a] and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 3 August 1988; accepted for publication 7 October 1988)

A theory of the electronic structures of periodic $N_1 \times N_2$ GaAs/$Al_xGa_{1-x}As$ superlattices grown along the [111] direction is presented. Deep levels associated with *s*- and *p*-bonded substitutional impurities in these superlattices are also predicted. It is found that: (i) in contrast with [001] superlattices, [111] superlattices are almost always direct band-gap superlattices. (ii) The [111] superlattices exhibit weaker quantum-well confinement than the corresponding [001] superlattices. (iii) As the thickness, t(GaAs), of each GaAs layer is reduced below a critical value ($t_c \simeq 13$ Å or $N_1 \simeq 4$ for $x = 0.7$), common shallow donor impurities such as Si cease donating electrons to the conduction band and instead become deep traps. For [111] superlattices t_c is smaller than the corresponding t_c for [001] superlattices. The fundamental band gap and the band edges of the superlattice, and hence the ionization energies of deep levels, depend strongly on the layer thickness t(GaAs) but only weakly on $t(Al_xGa_{1-x}As)$. The T_2- and A_1-derived deep levels (of the bulk point group T_d) are split and shifted, respectively, near a GaAs/$Al_xGa_{1-x}As$ interface: the *p*-like T_2 level splits into an a_1 (p_u-like) level and a doubly degenerate e (p_π-like) level of the point group for any general superlattice site (C_{3v}), whereas the *s*-like A_1 bulk level becomes an a_1 (*s*-like) level of C_{3v}. The order of magnitude of the shifts and splittings of deep levels at a GaAs/$Al_xGa_{1-x}As$ interface is less than 0.1 eV, depends on x, and becomes very small for impurities more than $\simeq 3$ atomic planes away from an interface. These predictions are based on a periodic superslab calculation for unit superslabs with total thickness t(GaAs) + $t(Al_xGa_{1-x}As)$ as large as 65.3 Å or $N_1 + N_2 = 20$ two-atom-thick layers. The Hamiltonian is a tight-binding model in a hybrid basis that is a generalization of the Vogl model and properly accounts for the nature of interfacial bonds. The deep levels are computed using the Hjalmarson *et al.* theory [Phys. Rev. Lett. 44, 810 (1980)] and the special points method.

I. INTRODUCTION

The theoretical and experimental study of semiconductor superlattices has become a very important area of basic semiconductor physics, device physics, and materials science. GaAs/$Al_xGa_{1-x}As$ superlattices,[1,2] which are commonly grown in the [001] direction, are the ones that have been most thoroughly studied and understood. Impurities play a very important role in determining the physical properties of semiconductors.[3] Recently we published a theoretical study of the electronic structures and substitutional deep impurities in such superlattices,[4,5] which, we believe, is the first systematic theoretical study of deep levels in superlattices. In the last few years, people have begun to investigate superlattices grown in the [111] direction, both experimentally[6,7] and theoretically.[8,9] Here we present a theory of the electronic structures of GaAs/$Al_xGa_{1-x}As$ superlattices grown in the [111] direction,[10] and we predict the energy levels of substitutional deep impurities in these superlattices. We compare our results for [111] superlattices with those for corresponding [001] superlattices.[4]

II. FORMALISM

A. Host Hamiltonian

We treat a periodic GaAs/$Al_xGa_{1-x}As$ superlattice whose layers are perpendicular to the bulk [111] direction, which we define as the direction "3." The bulk directions $[2,-1,-1]$ and $[0,1,-1]$ are defined as the directions "1" and "2." We employ a nearest-neighbor tight-binding Hamiltonian with an s^*sp^3 basis of five orbitals at each site. Our Hamiltonian, in the limit of $x = 0$, is identical to the Vogl model for GaAs.[11] Some differences are introduced because of the superlattice, which we treat using a superhelix or supercell method. The superlattice we consider has N_1 two-atom-thick layers of GaAs and N_2 two-atom-thick layers of $Al_xGa_{1-x}As$ repeated periodically; the GaAs and $Al_xGa_{1-x}As$ are assumed to be perfectly lattice matched. We denote this superlattice either as a $(GaAs)_{N_1}(Al_xGa_{1-x}As)_{N_2}$ superlattice or as a GaAs/$Al_xGa_{1-x}As$ superlattice with N_1 GaAs layers and N_2 $Al_xGa_{1-x}As$ layers, or as an $N_1 \times N_2$ GaAs/$Al_xGa_{1-x}As$ superlattice.

We first define (for the case $x = 1$) a *superhelix* or supercell as a helical string with axis aligned along the [111] direction consisting of $2N_1 + 2N_2$ adjacently bonded atoms As, Ga, As, Ga, As,..., Ga, As, Al, As, Al, As, Al,..., As, Al. (For $x \neq 1$, replace Al by Al_xGa_{1-x}.) The center of this helix is at L and each of the atoms of the helix is at position $\mathbf{L} + \mathbf{v}_\beta$ (for $\beta = 0, 1, 2, \ldots, 2N_1 + 2N_2 - 1$). A *superslab* of GaAs/$Al_xGa_{1-x}As$ consists of all such helices with the same value of L_3 and all possible different values of L_1 and L_2; and the superlattice is a stacked array of these superslabs. If the origin of coordinates is taken to be at an As atom, the x and y axes are oriented such that a neighboring cation is at

[a] Permanent address: Department of Physics, University of Science and Technology of China, Hefei, China.

 0021-8979/89/051987-08$02.40 © 1989 American Institute of Physics

$(\frac{1}{4},\frac{1}{4},\frac{1}{4})a_L$, where a_L is the lattice constant.[12] At each site there are s^*sp^3 basis orbitals $|n,\mathbf{L},\nu_\beta\rangle$, where $n = s^*, s, p_x, p_y$, or p_z and $\beta = 0, 1, 2, \ldots, 2N_1 + 2N_2 - 1$. From these basis orbitals we form the sp^3 hybrid orbitals at each site $\mathbf{R} = (\mathbf{L},\nu_\beta)$.

The hybrid orbitals are

$$|h_1,\mathbf{R}\rangle = [|s,\mathbf{R}\rangle + \lambda|p_x,\mathbf{R}\rangle + \lambda|p_y,\mathbf{R}\rangle + \lambda|p_z,\mathbf{R}\rangle]/2,$$
$$|h_2,\mathbf{R}\rangle = [|s,\mathbf{R}\rangle + \lambda|p_x,\mathbf{R}\rangle - \lambda|p_y,\mathbf{R}\rangle - \lambda|p_z,\mathbf{R}\rangle]/2,$$
$$|h_3,\mathbf{R}\rangle = [|s,\mathbf{R}\rangle - \lambda|p_x,\mathbf{R}\rangle + \lambda|p_y,\mathbf{R}\rangle - \lambda|p_z,\mathbf{R}\rangle]/2,$$

and

$$|h_4,\mathbf{R}\rangle = [|s,\mathbf{R}\rangle - \lambda|p_x,\mathbf{R}\rangle - \lambda|p_y,\mathbf{R}\rangle + \lambda|p_z,\mathbf{R}\rangle]/2, \quad (1)$$

where $\lambda = +1(-1)$ for atoms at anion (cation) sites. Next we introduce the label $\nu = s^*, h_1, h_2, h_3$, or h_4, and our hybrid basis orbitals are $|\nu,\mathbf{R}\rangle$. In terms of these orbitals we form the tight-binding orbitals

$$|\nu,\beta,\mathbf{k}\rangle = N_s^{-1/2}\Sigma_{\mathbf{L}}\exp(i\mathbf{k}\cdot\mathbf{L} + i\mathbf{k}\cdot\nu_\beta)|\nu,\mathbf{L},\nu_\beta\rangle, \quad (2)$$

where $\mathbf{k}$ is any wave vector in the Brillouin zone in the superlattices. Here N_s is the number of supercells.

The wave vector is a good quantum number, and so the tight-binding Hamiltonian is diagonal in $\mathbf{k}$. Evaluation of the matrix elements $\langle\nu,\beta,\mathbf{k}|H|\nu',\beta',\mathbf{k}\rangle$ leads to a tight-binding Hamiltonian of the block tridiagonal form. For different β and β', the first three rows of block matrices are

$$\begin{matrix} H(0,0) & H(0,1) & 0 & \ldots & \ldots & 0 & H(0,2N_1+2N_2-1) \\ H(0,1)^\dagger & H(1,1) & H(1,2) & 0 & 0 & \ldots 0 & 0 \\ 0 & H(1,2)^\dagger & H(2,2) & H(2,3) & 0 & \ldots 0 & 0. \end{matrix} \quad (3)$$

The last row of blocks is

$$H(0,2N_1+2N_2-1)^\dagger \quad 0 \quad 0\ldots 0 \quad H(2N_1+2N_2-2,2N_1+2N_2-1)^\dagger \quad H(2N_1+2N_2-1,2N_1+2N_2-1). \quad (4)$$

Here $H(\beta,\beta')$ depends on $\mathbf{k}$ and is given in terms of various 5×5 matrices for different ν and ν'.

The diagonal (in β) 5×5 matrix, $H(\beta,\beta)$ at site β, is

$$H(\beta,\beta) = \langle\nu,\beta,\mathbf{k}|H|\nu',\beta,\mathbf{k}\rangle = \begin{pmatrix} \epsilon_{s^*} & 0 & 0 & 0 & 0 \\ 0 & \epsilon_h & T & T & T \\ 0 & T & \epsilon_h & T & T \\ 0 & T & T & \epsilon_h & T \\ 0 & T & T & T & \epsilon_h \end{pmatrix}, \quad (5)$$

where

$$\epsilon_h = (\epsilon_s + 3\epsilon_p)/4 \quad (6)$$

is the hybrid energy, and

$$T = (\epsilon_s - \epsilon_p)/4 \quad (7)$$

is the hybrid-hybrid interaction; the energies ϵ_{s^*}, ϵ_h, and T in $H(\beta,\beta)$ refer to the atom at the βth site, and may be obtained from the energies w tabulated by Vogl, Hjalmarson, and Dow.[11] To account for the observed[13] valence-band edge discontinuity of 32% of the direct band gap, a constant is added to ϵ_{s^*} and ϵ_h for $Al_xGa_{1-x}As$; this constant is adjusted to give the valence-band maximum of $Al_xGa_{1-x}As$ below the valence-band maximum of GaAs by 32% of the band-gap difference in the limit $N_1 = N_2 \to \infty$.

The off-diagonal matrix elements $\langle\nu,\beta,\mathbf{k}|H|\nu',\beta',\mathbf{k}\rangle$ or $H(\beta,\beta')$ are best expressed in terms of matrix elements of H between s^*, s, and p orbitals. This is accomplished by the transformation

$$\langle\nu,\beta,\mathbf{k}|H|\nu',\beta',\mathbf{k}\rangle = \Sigma_{n,n'}C(\nu,n;\beta)C(\nu',n';\beta')\langle n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k}\rangle. \quad (8)$$

where we have the s^*sp^3 tight-binding functions,

$$|n,\beta,\mathbf{k}\rangle = N_s^{-1/2}\Sigma_{\mathbf{L}}\exp(i\mathbf{k}\cdot\mathbf{L} + i\mathbf{k}\cdot\nu_\beta)|n,\mathbf{L},\nu_\beta\rangle, \quad (9)$$

and the 5×5 matrices $C(\nu,n;\beta)$ are

$$C(\nu,n;\beta) = \begin{matrix} \\ s^* \\ h_1 \\ h_2 \\ h_3 \\ h_4 \end{matrix}\begin{matrix} \begin{matrix} s^* & s & p_x & p_y & p_z \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & +\lambda/2 & +\lambda/2 & +\lambda/2 \\ 0 & 1/2 & +\lambda/2 & -\lambda/2 & -\lambda/2 \\ 0 & 1/2 & -\lambda/2 & +\lambda/2 & -\lambda/2 \\ 0 & 1/2 & -\lambda/2 & -\lambda/2 & +\lambda/2 \end{pmatrix} \end{matrix}, \quad (10)$$

where $\lambda = +1(-1)$ if β refers to an anion (cation).

There are several distinct cases for which the off-diagonal (in β) matrix elements $\langle n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k}\rangle$ are nonzero (for $\beta \neq \beta'$).

1. Intramaterial matrix elements

If β and β' both refer to nearest-neighbor sites in the same material (either the GaAs or the $Al_xGa_{1-x}As$) we have (assuming β and β' are in material number 1, the GaAs), for example,

$$\langle n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k}\rangle = H_{c1a1}, \quad (11)$$

if β refers to a cation and β' refers to an anion. H_{c1a1} is a 5×5 matrix whose rows and columns are labeled by n, which ranges over the values s^*, s, p_x, p_y, and p_z. Similarly we have matrix elements H_{a1c1}, H_{a2c2}, and H_{c2a2}. These matrix elements are

$$H_{c1a1} = \begin{pmatrix} 0 & 0 & -C_0V_6 & -C_0V_6 & -C_0V_6 \\ 0 & C_0V_1 & -C_0V_4 & -C_0V_4 & -C_0V_4 \\ C_0V_7 & C_0V_5 & C_0V_2 & C_0V_3 & C_0V_3 \\ C_0V_7 & C_0V_5 & C_0V_3 & C_0V_2 & C_0V_3 \\ C_0V_7 & C_0V_5 & C_0V_3 & C_0V_3 & C_0V_2 \end{pmatrix}. \quad (12)$$

Here we have $C_0 = g_0^*$, where we have

$$4g_0 = \exp(i\mathbf{k}\cdot\mathbf{x}_0),$$

and $\mathbf{x}_0 = (a_L/4)(1,1,1)$, with a_L being the lattice constant of GaAs or AlAs. Here k is the wave vector. In addition, we have

$$V_1 = V(s,s), \qquad V_2 = V(x,x),$$

$$V_3 = V(x,y), \qquad V_4 = V(sa,pc),$$

$$V_5 = V(sc,pa), \qquad V_6 = V(s^*a,pc),$$

and

$$V_7 = V(s^*c,pa);$$

in the notation of Vogl and co-workers,[11] and

$$H_{a1c1} = \begin{pmatrix} 0 & 0 & V_6(g_1-g_2-g_3) & V_6(-g_1+g_2-g_3) & V_6(-g_1-g_2+g_3) \\ 0 & V_1(g_1+g_2+g_3) & V_4(g_1-g_2-g_3) & V_4(-g_1+g_2-g_3) & V_4(-g_1-g_2+g_3) \\ -V_7(g_1-g_2-g_3) & -V_5(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) & V_3(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) \\ -V_7(-g_1+g_2-g_3) & -V_5(-g_1+g_2-g_3) & V_3(-g_1-g_2+g_3) & V_2(g_1+g_2+g_3) & V_3(g_1-g_2-g_3) \\ -V_7(-g_1-g_2+g_3) & -V_5(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) & V_3(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) \end{pmatrix} \tag{13}$$

We also have

$$4g_1 = \exp(i\mathbf{k}\cdot\mathbf{x}_1), \quad 4g_2 = \exp(i\mathbf{k}\cdot\mathbf{x}_2),$$

and

$$4g_3 = \exp(i\mathbf{k}\cdot\mathbf{x}_3).$$

Here we have $\mathbf{x}_1 = (a_L/4)(1,-1,-1)$, $\mathbf{x}_2 = (a_L/4)(-1,1,-1)$, and $\mathbf{x}_3 = (a_L/4)(-1,-1,1)$.

All of the matrix elements V are those tabulated by Vogl and co-workers,[11] for material number 1 (viz., GaAs). Identical expressions exist for H_{a2c2} and H_{c2a2}, with $Al_xGa_{1-x}As$ matrix elements. [The $Al_xGa_{1-x}As$ matrix elements are obtained by a virtual-crystal average of the Vogl matrix elements for AlAs and GaAs: x times the corresponding AlAs matrix elements plus $(1-x)$ times the GaAs matrix elements.]

2. Intermaterial matrix elements

At the interface between GaAs and $Al_xGa_{1-x}As$ there will be nonzero matrix elements of H for each bond between nearest neighbors. These are H_{a2c1} and H_{c2a1}:

$$H_{a2c1} = \begin{pmatrix} 0 & 0 & V_6(g_1-g_2-g_3) & V_6(-g_1+g_2-g_3) & V_6(-g_1-g_2+g_3) \\ 0 & V_1(g_1+g_2+g_3) & V_4(g_1-g_2-g_3) & V_4(-g_1+g_2-g_3) & V_4(-g_1-g_2+g_3) \\ -V_7(g_1-g_2-g_3) & -V_5(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) & V_3(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) \\ -V_7(-g_1+g_2-g_3) & -V_5(-g_1+g_2-g_3) & V_3(-g_1-g_2+g_3) & V_2(g_1+g_2+g_3) & V_3(g_1-g_2-g_3) \\ -V_7(-g_1-g_2+g_3) & -V_5(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) & V_3(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) \end{pmatrix} \tag{14}$$

and

$$H_{c2a1} = \begin{pmatrix} 0 & 0 & V_6(C_1-C_2-C_3) & V_6(-C_1+C_2-C_3) & V_6(-C_1-C_2+C_3) \\ 0 & V_1(C_1+C_2+C_3) & V_4(C_1-C_2-C_3) & V_4(-C_1+C_2-C_3) & V_4(-C_1-C_2+C_3) \\ -V_7(C_1-C_2-C_3) & -V_5(C_1-C_2-C_3) & V_2(C_1+C_2+C_3) & V_3(-C_1-C_2+C_3) & V_3(-C_1+C_2-C_3) \\ -V_7(-C_1+C_2-C_3) & -V_5(-C_1+C_2-C_3) & V_3(-C_1-C_2+C_3) & V_2(C_1+C_2+C_3) & V_3(C_1-C_2-C_3) \\ -V_7(-C_1-C_2+C_3) & -V_5(-C_1-C_2+C_3) & V_3(-C_1+C_2-C_3) & V_3(C_1-C_2-C_3) & V_2(C_1+C_2+C_3) \end{pmatrix}. \tag{15}$$

Here the Vogl matrix elements are those for the bond in question: If the cation is Al_xGa_{1-x} and the anion is As, then the matrix element is the $Al_xGa_{1-x}As$ matrix element obtained by a virtual-crystal average of the AlAs and GaAs values.

After the tight-binding Hamiltonian is given, calculation of the electronic structure is straightforward. In this work we study electronic structure and deep impurity levels in superlattices as large as $N_1 + N_2 = 20$; that is, in 40-atoms-thick superslabs. The dimension of the Hamiltonian matrix at each value of k is $5(2N_1 + 2N_2)$, because there are five orbitals per site. We diagonalize this Hamiltonian numerically for each (special-point) k in the Brillouin zone (Fig. 1), finding its eigenvalues $E_{\gamma,\mathbf{k}}$ and, if necessary, the projections of the eigenvectors $|\gamma,\mathbf{k}\rangle$ on the $|\nu,\beta,\mathbf{k}\rangle$ hybrid basis: $(\nu,\beta,\mathbf{k}|\gamma,\mathbf{k})$. Here γ is the band index (and ranges from 1 to 200 for $N_1 = N_2 = 10$) and k is the wave vector: it

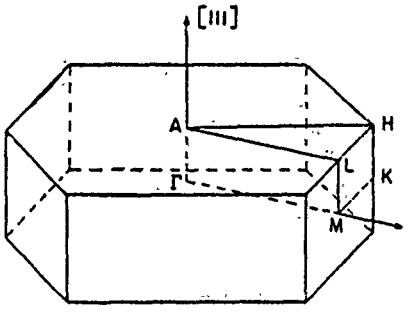


FIG. 1. The Brillouin zone of [111] superlattices. Γ is at (0,0,0), A is at $(1,1,1)/(2N_1+2N_2)$, M is at $(1,-2,1)/3$ or equivalent points, K is at $(0,2,-2)/3$ or equivalent points, all are in units of $(2\pi/a_L)$. Please note for a GaAs/GaAs or $Al_xGa_{1-x}As/Al_xGa_{1-x}As$ [111] superlattice, the superlattice Γ point might correspond to the bulk Γ point and/or $L[(1,1,1)]$ point of the ordinary bulk Brillouin zone, and the superlattice M point might correspond to the X point and/or $L((1,-1,-1)$, $(-1,1,1)$ or $(-1,-1,1))$ points of the ordinary bulk Brillouin zone.

can be within the mini-Brillouin zone in a reduced zone scheme or within the GaAs Brillouin zone in an extended zone scheme. (We assume that GaAs and $Al_xGa_{1-x}As$ are perfectly lattice matched.)

B. Deep levels

The theory of deep levels is based on the Hjalmarson *et al.* Green's function theory,[14] which solves the secular equation for the deep level energy E

$$\det[1 - G(E)V] = 0$$

$$= \det\left(1 - P\int [\delta(E' - H)/(E - E')]\, dE'\, V\right). \tag{16}$$

Here V is the defect potential matrix,[14] which is zero except at the defect site and diagonal on that site, $(0, V_s, V_p, V_p, V_p)$, in the Vogl s^*sp^3 local basis centered on each atom. We also have $G = (E - H)^{-1}$, where H is the host tight-binding Hamiltonian operator. The spectral density operator is $\delta(E' - H)$ and P denotes the principal value integral over all energies. In the fundamental band gap of the superlattice, G is real.

1. Point-group analysis

A substitutional point defect in either bulk GaAs or bulk $Al_xGa_{1-x}As$ has tetrahedral (T_d) point-group symmetry (assuming a virtual-crystal approximation for $Al_xGa_{1-x}As$). Each such sp^3-bonded defect normally has one s-like (A_1) and one triply degenerate p-like (T_2) deep defect level near the fundamental band gap. If we imagine breaking the symmetry of bulk GaAs by making it into a GaAs/GaAs superlattice along the [111] direction, we reduce the T_d symmetry to C_{3v}. In the GaAs/$Al_xGa_{1-x}As$ superlattice the A_1 and T_2 deep levels of the bulk GaAs or $Al_xGa_{1-x}As$ produce two a_1 levels (one s-like, derived from the A_1 level and one T_2-derived p_σ-like), one doubly degenerate e level (p_π-like). Of course, for impurities far from a GaAs/$Al_xGa_{1-x}As$ interface, the s-like a_1 level will have an energy very close to the energy of a bulk A_1 level, and the p_σ-like a_1 level and the p_π-like e level will lie close to the bulk T_2 level also. But, if T_2-derived levels lie close to the valence-band maximum, then the splitting of the valence-band edge (into a doubly degenerate p_π-like maximum and a p_σ band edge at slightly lower energy) may cause the a_1 component deep level to lie lower in energy than the e levels by an energy comparable with the splitting [see Eq. (16)]. Note that this splitting exists even for defects distant from the interface and is a consequence of the different host spectral densities in the superlattice for a_1 and e states.

2. Secular equations

The secular equation, Eq. (16), is reduced by the C_{3v} point-group symmetry to the following two equations:

$$G(e;E) = V_p^{-1} \quad \text{for } e \text{ levels}, \tag{17}$$

and

$$\det \begin{vmatrix} G(s,s;E)V_s - 1 & G(s,\sigma;E)V_p \\ G(\sigma,s;E)V_s & G(\sigma,\sigma;E)V_p - 1 \end{vmatrix} = 0, \tag{18}$$

for a_1 levels, where we have

$$G(e;E) = \Sigma_{\gamma,\mathbf{k}} |(h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_3,\beta,\mathbf{k}|\gamma,\mathbf{k})|^2/2(E - E_{\gamma,\mathbf{k}}), \tag{19}$$

$$G(s,s;E) = \Sigma_{\gamma,\mathbf{k}} |(h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle + (h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle + (h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle + (h_4,\beta,\mathbf{k}|\gamma,\mathbf{k})|^2/4(E - E_{\gamma,\mathbf{k}}), \tag{20}$$

$$G(\sigma,\sigma;E) = \Sigma_{\gamma,\mathbf{k}} |3(h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_4,\beta,\mathbf{k}|\gamma,\mathbf{k})|^2/12(E - E_{\gamma,\mathbf{k}}), \tag{21}$$

and

$$G(s,\sigma;E) = \Sigma_{\gamma,\mathbf{k}} [(h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle + (h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle + (h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle + (h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle] \times [3(h_1,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_2,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_3,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle - (h_4,\beta,\mathbf{k}|\gamma,\mathbf{k}\rangle]^*/[4\sqrt{3}(E - E_{\gamma,\mathbf{k}})]. \tag{22}$$

Here $G(\sigma,s;E)$ is the Hermitian conjugate of $G(s,\sigma;E)$ and β is the site of the defect in the superlattice.

For each β the relevant host Green's functions, Eqs. (19)–(22), are evaluated, and the secular equations (17) and (18) are solved, yielding $E(e;V_p)$ and two values of $E(a_1;V_s,V_p)$. The defect potential matrix elements V_s and V_p are obtained as in Ref. 4. For $N_1 = N_2 = 10$, there are 40 possible sites β, each with four relevant deep levels: two a_1 and one doubly degenerate e; thus there are 160 levels.

For our studies of deep levels in the band gaps of superlattices, we consider only superlattices such that $(N_1 + N_2)/4$ is an integer. In such cases the sums over k in Eqs. (9)–(22) can be performed using ten special points.[15] These ten superlattice special points are sufficient to give the Green's function at any site in the superlattice with the same accuracy as Chadi's and Cohen's[16] ten bulk special points would give for bulk GaAs or AlAs, i.e., GaAs/GaAs or AlAs/AlAs superlattices. This level of accuracy is known to be adequate for bulk Green's function calculations.

III. RESULTS

A. Superlattice band structures

We first calculate the band structures of superlattices for $N_1 + N_2 = 20$ at points of high symmetry in the first Brillouin zone of the superlattice (Fig. 1). Our calculations produce $E_{\gamma,\mathbf{k}}$, the superlattice bands, as functions of N_1, the thickness of GaAs layers (Fig. 2). As the thickness of the GaAs slab, $t(\text{GaAs}) = N_1 a_1/\sqrt{3}$ (where $a_1 = 5.653$ Å[12] is the lattice constant of GaAs), or the number of GaAs layers N_1 becomes small in comparison with the thickness,

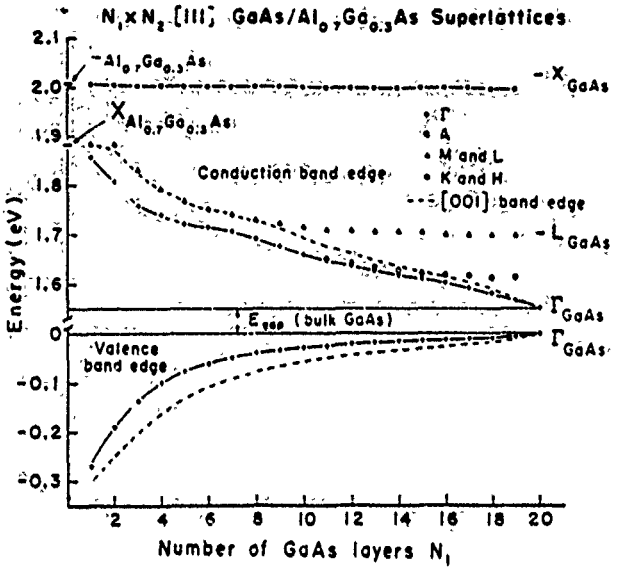


FIG. 2. Predicted energies (in eV) of the superlattice conduction-band minima and valence-band maximum with respect to the valence-band maximum *of bulk GaAs* for a [111] GaAs/$Al_xGa_{1-x}As$ superlattice vs reduced layer thicknesses N_1 and N_2 for various $N_1 \times N_2$ [111] GaAs/$Al_xGa_{1-x}As$ superlattices, with $x = 0.7$ and $N_1 + N_2$ fixed to be 20. The calculations are based on the low-temperature band structures of GaAs and $Al_{0.7}Ga_{0.3}As$, with bulk band gaps of 1.55 and 2.22 eV respectively. The conduction-band eigenvalues at Γ, L, $M(L)$, and $K(H)$ of the superlattices are shown as solid circles, open circles, solid triangles, and solid squares. The superlattice valence-band maximum is at $k = 0$. Note the broken scale on the ordinate. The conduction-band minimum and the valence-band maximum of the [001] superlattices calculated in Ref. 4 are also shown here as dashed lines for comparison. The positions of the band extrema of bulk GaAs at Γ, L, and X are shown on the right of the figure, at $N_1 = 20$.

$t(Al_xGa_{1-x}As) = N_2 a_2/\sqrt{3}$, or number of layers N_2 of $Al_xGa_{1-x}As$ slabs, the small band-gap GaAs layers become quantum wells surrounded by large band-gap $Al_xGa_{1-x}As$. As a result, the band gap of the superlattice increases from the GaAs band gap toward the $Al_xGa_{1-x}As$ band gap as N_1 decreases (for N_2 large). This behavior for [111] superlattices is similar to that found for [001] superlattices.[4] The top of the valence band is at the Brillouin-zone center, as found for [001] superlattices. We find the band gaps of $N_1 \times N_2$ [111] GaAs/$Al_xGa_{1-x}As$ superlattices to be direct for almost all the cases we calculated (i.e., both the superlattices' valence-band maxima and conduction-band minima are at Γ for different N_1, N_2, and x), which was not generally the case for [001] superlattices. (We did find one extreme case, namely a 1×11 GaAs/AlAs superlattice, in which the gap was indirect by $\simeq 0.06$ eV.) This could turn out to be a technologically important property of these superlattices, because (i) it means that [111] superlattices can be used in the smallest optoelectronic devices and (ii) for such superlattices, the electron effective mass in the plane perpendicular to the growth direction is always the electron effective mass of bulk GaAs, which is the smallest effective mass that GaAs/$Al_xGa_{1-x}As$ can have. The smaller effective mass implies larger mobility and faster device speed.

B. Quantum-well confinement

The fundamental band gaps for $N_1 \times N_2$[111] GaAs/$Al_xGa_{1-x}As$ superlattices are always smaller than for the corresponding [001] superlattices, as can be seen by a direct comparison of the superlattices with two different growth directions in Fig. 2, and, for some much larger superlattice periods, in Fig. 1 of Ref. 9. It can also be seen that the smaller the thickness of the quantum-well (GaAs) layer, the more significant is the difference between the quantum-well confinement of the direct edge for [001] and [111] superlattices. For the same number N_1 of GaAs layers, the thickness of the quantum well in a [111] superlattice is greater than in a [001] superlattice by a factor of $2/\sqrt{3} = 1.1547$, but even if this factor is included, the weaker quantum-well confinement in the [111] superlattices is still not completely explained by a simple one-band effective-mass model.[9] One of the most dramatic differences between [111] and [001] superlattices of similar GaAs layer thickness is the much larger heavy-light hole energy level splitting at the top of valence bands of the superlattices. For 10×10 GaAs/$Al_{0.7}Ga_{0.3}$ superlattices this splitting is 0.087 eV for the [111] superlattice, versus 0.035 eV for the [001] superlattice. We attribute the smaller quantum-well confinement and the larger heavy-light hole splitting in the [111] superlattice to the topological difference in the bonding between the [111] and the [001] superlattices. For example, at an interface in a [111] GaAs/AlAs superlattice, an As atom is bonded to either three Ga atoms and one Al atom or to three Al's and one Ga. In contrast, an interfacial As in a [001] superlattice is bonded to two Ga and two Al atoms. This topological difference, which has an intrinsic three-dimensional character, affects both the valence and conduction bands, altering their masses in a way that is not described quantitatively correctly by a one-band effective-mass model.

The band gap is somewhat more sensitive to changes of the GaAs layer thickness than to changes in the $Al_xGa_{1-x}As$ layer thickness (see Fig. 3), as was the case for [001] superlattices.

C. Defect levels

The substitutional defect energy levels for sp^3-bonded impurities can be evaluated using the techniques of Hjalmarson *et al.*,[14] as described above for superlattices.

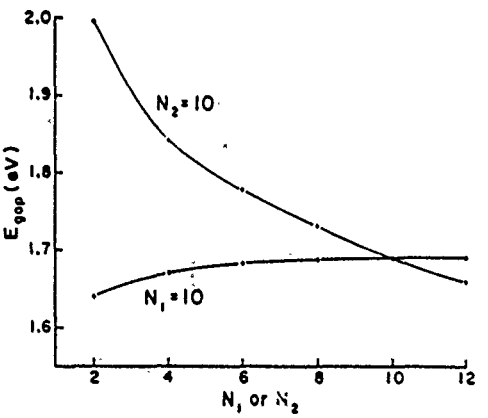


FIG. 3. Predicted fundamental energy-band gaps at $k = 0$ of a [111] $(GaAs)_{N_1}(Al_xGa_{1-x}As)_{N_2}$ superlattice as a function of reduced GaAs layer thickness N_1 and $Al_xGa_{1-x}As$ layer thickness N_2 for $x = 0.7$. Note that the variation of the gap with decreasing N_1 from, say, 12 to 2 is less than the variation associated with changing N_1 from 12 to 2.

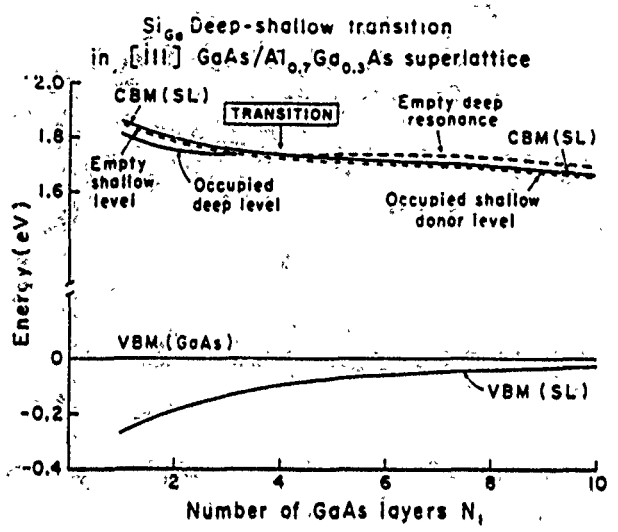


FIG. 4. Illustrating the deep-to-shallow transition as a function of increasing GaAs layer thickness N_1 in a GaAs/$Al_xGa_{1-x}As$ $N_1 \times N_2$ superlattice (SL) with $x = 0.7$ and $N_1 + N_2 = 20$ for a Si impurity on a column III site in the center of a GaAs layer of the superlattice host. The superlattice conduction-band edge (CBM) and valence-band maximum (VBM) are indicated by light solid lines. The Si deep level is denoted by a heavy line, which is solid when the level is in the gap but dashed when the level is resonant with the conduction band. The deep level in the band gap for $N_1 \leq 4$ is covered up by the conduction band as a result of changes in the host for $N_1 > 4$. The impurity's deep level lies in the gap for $N_1 \leq 4$ and is occupied by the extra Si electron; the Si, in this case, is thus a "deep impurity." For $N_1 > 4$, the deep level lies above the conduction-band edge as a resonance. The daughter electron from the Si impurity which was destined for this deep level is autoionized, spills out of the deep resonance level, and falls to the conduction-band edge (light solid line) where it is subsequently bound (at low temperature) in a shallow level associated with the long ranged Coulomb potential of the donor (indicated by the short light dashed line). It is important to realize that *both* the deep level and the shallow levels coexist and are distinct levels with qualitatively different wave functions. The issue of whether an impurity is "deep" or "shallow" is determined by whether or not a deep level associated with the impurity lies in the band gap. The computed deep-shallow transition occurs for $N_1 \simeq 4$ layers. While the qualitative physics is completely reliable, it is possible that the transition layer thickness may differ somewhat from $N_1 = 4$ in real superlattices.

1. Dependence on layer thickness

Figure 4 displays the dependence on GaAs reduced layer thickness N_1 of the deep Ga-site A_1 level of a Si impurity[17] *in the middle of a GaAs layer* in a GaAs/$Al_{0.7}Ga_{0.3}As$ superlattice. As the size N_1 of the GaAs layer shrinks, the *deep levels remain relatively constant* in energy with respect to the GaAs valence-band maximum while the conduction-band edge of the superlattice increases in energy—progressively uncovering the once resonant deep level of Si and converting this shallow donor impurity into a deep trap.[4,18] This shallow-deep transition as a function of GaAs well-size N_1, is similar for both [111] and [001] superlattices. But because the band gap of the [111] superlattice is less sensitive to quantum confinement, the transition of Si from a shallow to a deep impurity will occur at a smaller critical GaAs layer thickness N_1. (Based on the general structure of the curves of Fig. 4, it probably occurs for $N_1 \simeq 4$ and a GaAs layer thickness of order $w = 13$ Å or less.)

A similar analysis can be made of the behavior of other impurities, for example, N in GaAs. Nitrogen has a resonant state in the conduction band of GaAs. We have predicted[4,5] that N in a GaAs layer of a [001] GaAs/$Al_xGa_{1-x}As$ superlattice will have an impurity state in the band gap if the thickness of the GaAs layer is small enough, for example, ≤ 48 Å or 17 molecular layers. Here we predict that N in the GaAs layers of a [111] GaAs/$Al_xGa_{1-x}As$ superlattice will have an impurity state in the band gap if the thickness of the GaAs layers is thinner than ≤ 32 Å or ten molecular layers.

2. As *vacancy levels*

Figure 5 displays the deep energy levels in the band gap of an As vacancy in a $N_1 = N_2 = 10$ GaAs/$Al_{0.7}Ga_{0.3}As$ superlattice, as a function of β, the site of the vacancy in the superhelix or superslab. A vacancy is simulated here by letting the coefficients of the vacancy state wave function at the vacancy site vanish, because a vacancy state cannot have an expansion coefficient at nonexisting orbitals.[19] Several features of the results in Fig. 5 are worth noting: in contrast with the [001] superlattice case, the [111] superlattices are not "symmetrical" any more: the two interfaces $\beta = 0$ and $\beta = 20$ are not equivalent to each other any more: at the $\beta = 0$ interface, there are three hybrids directed inside the $Al_xGa_{1-x}As$ layer and one hybrid in the GaAs layer, but at the $\beta = 20$ interface, there are three hybrids inside the GaAs layer and one hybrid in the $Al_xGa_{1-x}As$ layer. A direct consequence of this nonsymmetry is the asymmetrical position-dependence of the T_2-derived As vacancy level. Near the interface $\beta = 0$ the *p*-like T_2 bulk As vacancy level splits into two levels: an a_1 level and a doubly degenerate e state. The a_1 level corresponds to a *p*-like level with its orbital perpendicular to the atomic layers and has almost the energy of a bulk GaAs T_2 level, while the e level corresponds to two *p*-like orbitals inside the atomic layer, and has an energy almost the same as the bulk $Al_xGa_{1-x}As$ T_2 level. However, near the interface $\beta = 20$ the a_1 level has an energy almost the same as the bulk $Al_xGa_{1-x}As$ level with its orbital oriented perpendicular to the atomic layers, and the e level has an energy almost the same as the bulk GaAs level. The splitting between the a_1 and e levels at the interface are smaller than the splitting in the corresponding [001] superlattice,[21] of order less than 0.1 eV, and becomes smaller yet when the vacancy is more than three or so atomic layers from the interface—but may not vanish even if the defect is distant from the interface, as a result of the splitting of the valence-band edge in the superlattice and the resulting changes of the host spectral density.[4] The other general behavior of the As vacancy state in [111] superlattices is very similar to the behavior in [001] superlattices.

Similar behavior to that found for the As vacancy levels is to be expected for all As site sp^3-bonded deep impurity levels in GaAs/$Al_xGa_{1-x}As$ superlattices, although the issue of whether a specific deep level lies in the fundamental band gap or not depends on the defect potential for that

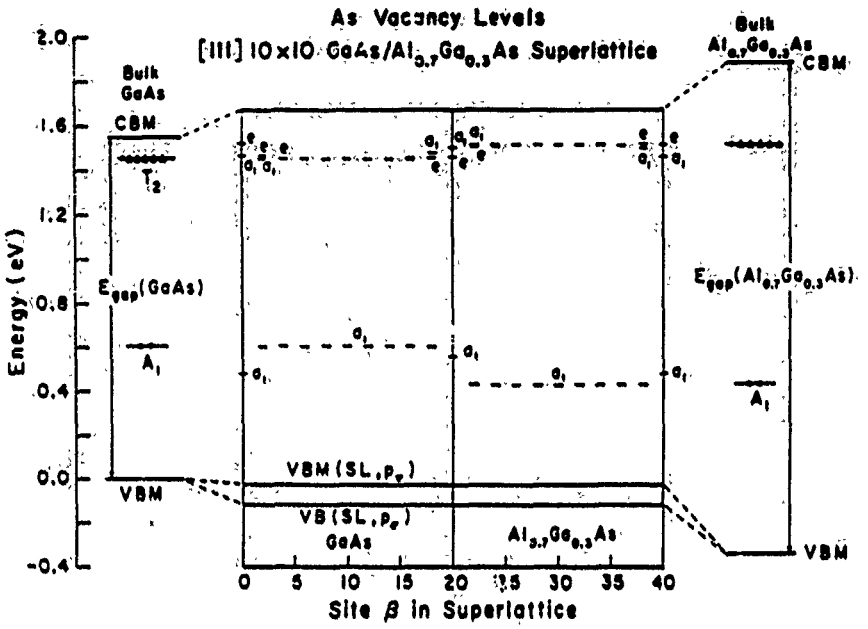


FIG. 5. Predicted energy levels of an As vacancy in a $(GaAs)_{10}(Al_{0.7}Ga_{0.3}As)_{10}$ superlattice, as a function of β, the position of the vacancy (even values of β correspond to As sites). Note the splitting of the T_2 levels at and near the interfaces ($\beta = 0$, 20, and 40), and that these vacancy levels lie at higher energy in an $Al_{0.7}Ga_{0.3}As$ layer than in a GaAs layer. The a_1 and e ordering changes at successive interfaces. The zero of energy is the valence-band maximum of bulk GaAs, and the corresponding valence band (VBM) and conduction band (CBM) edges and deep levels in bulk GaAs and bulk $Al_{0.7}Ga_{0.3}As$ are given to the left and right of the central figure, respectively. The top of the central figure is the conduction-band edge of the superlattice, and the bottom, corresponds to the split valence band in the superlattice—the valence-band maximum of the superlattice being of e symmetry (p_π) and the split-off a_1 (p_σ) band maximum lying 0.087 eV lower in energy. The A_1 level in the $Al_xGa_{1-x}As$ layer of the superlattice is lower than the corresponding level in the GaAs layer because of the band offset of 0.33 eV. The electron (hole) occupancies of the neutral deep levels in bulk GaAs and bulk $Al_{0.7}Ga_{0.3}As$ are denoted by solid circles (open triangles).

impurity and on N_1, N_2, and x. But if a T_2 level does lie in the gap, the lower the energy level is, the larger the superlattice induced splitting of it would be, due to the greater influence of the split top of the valence band.

3. Cation vacancy levels

The A_1 bulk levels for a Ga vacancy in GaAs and for a cation vacancy in $Al_xGa_{1-x}As$ all lie very deep in the host

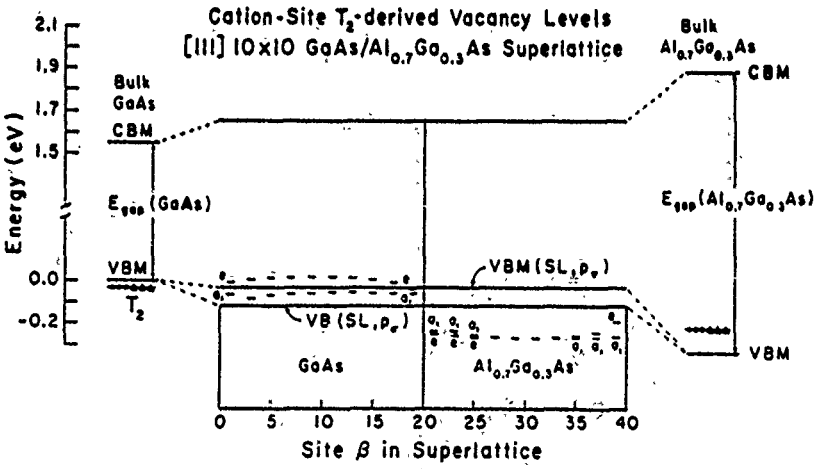


FIG. 6. Predicted T_2-derived levels of a cation vacancy in a $(GaAs)_{10}(Al_{0.7}Ga_{0.3}As)_{10}$ superlattice, as a function of site index β, the position of the vacancy. The zero of energy is the valence-band edge of bulk GaAs, and the corresponding valence band and conduction-band edges and deep levels in bulk GaAs and bulk $Al_{0.7}Ga_{0.3}As$ are given to the left and right of the central figure, respectively. The top of the central figure is the conduction-band edge of the superlattice and the two bottom lines correspond to the p_π-like valence band maximum and the split p_σ-like edge below it. Electrons occupying the levels are denoted by solid circles. Holes initially in levels *below the valence-band maximum* are denoted by open triangles and bubble up to the valence-band maximum, where the long ranged Coulomb potential can trap them in shallow acceptor levels (not shown). The cation vacancy in bulk GaAs and in an $Al_xGa_{1-x}As$ layer of the superlattice is a triple shallow acceptor, providing three such holes to the valence band. In bulk $Al_xGa_{1-x}As$ and in a GaAs layer of the superlattice, the vacancy produces a deep trap for either electrons or holes. In the superlattice, the low-energy level in a GaAs layer is of a_1 symmetry, and the high-energy level is of e symmetry, as indicated on the figure.

either in the bulk or at an interface of a GaAs/$Al_xGa_{1-x}As$ superlattice.

The T_2-derived cation vacancy levels produce deep levels near the valence-band maxima of both [001] and [111] GaAs/$Al_xGa_{1-x}As$ superlattices[4] (see Fig. 6). In bulk GaAs, the T_2 cation vacancy level is predicted[11] to lie $\simeq$0.03 eV below the valence-band maximum. In bulk $Al_{0.7}Ga_{0.3}As$, the predicted vacancy level is in the gap, 0.11 eV above the $Al_{0.7}Ga_{0.3}As$ valence-band maximum. If a 0.33-eV band offset is assumed for the GaAs/$Al_{0.7}Ga_{0.3}As$ superlattice, this vacancy level should lie 0.22 eV below the top of valence band of GaAs. In the [111] GaAs/$Al_xGa_{1-x}As$ superlattice near the interface, the T_2 cation vacancy level splits into a_1 and e sublevels. For $x \simeq 0.7$, some or all of these sublevels may lie in the gap of the superlattice. The energy of a deep level with specified symmetry is determined by a balance between the conduction-band states, which push the level down in energy, and the valence-band states, which repel it upward. Hence it can be easily understood that the cation vacancy levels, because they are very close to the top of valence band of the superlattice, are most affected by the heavy (e state)-light (a_1 state) hole splitting of the superlattice, as we have seen in the [001] superlattice case. For [111] superlattices, because the heavy-light hole splitting is even larger, its effect is even larger.

If the predictions are taken literally, then near the interface the cation vacancy produces a very interesting level structure, depending on the site of the vacancy. To begin with, in a GaAs layer of the superlattice, the bulk-GaAs T_2 valence-band resonance becomes either an e level in the gap of the superlattice or an e resonant state (depending on the position of the vacancy), and an a_1 resonant state. The e levels lie roughly $\simeq$0.07 eV above the a_1 levels, much more than in the case of [001] superlattices. This is because the T_2 levels are near the valence-band maximum and, in the superlattice, the T_2-like valence-band maximum is split into a_1 (light-hole, p_σ-like) and e (heavy-hole, p_π-like) edges, and because the a_1 valence-band edge has a lighter effective mass and a stronger quantum-well confinement effect: the band edge for a_1 (p_σ-like) states lies 0.087 eV below the edge for e states. The a_1 states lie lower because the band states deep in the valence band that repel them are at lower energy in the superlattice (this light-hole character of the p_σ-like valence-band states cause them to be driven down in energy by the quantum confinement of the superlattice). In an $Al_{0.7}Ga_{0.3}As$ layer, the cation-deep vacancy levels are now all resonant with the valence band of the superlattice (although in bulk $Al_{0.7}Ga_{0.3}As$ they lie in the gap). This is due mainly to the band offset and the fact that, roughly speaking, the deep levels do not move (much) whereas the valence-band edges do, as one goes from bulk GaAs to the superlattice to bulk $Al_{0.7}Ga_{0.3}As$.

The neutral Ga vacancy in the bulk of GaAs (assuming its deep level lies in the valence band) is a triple acceptor (Fig. 6). Consider this vacancy at the near-interfacial site $\beta = 1$ in GaAs; its levels, in order of decreasing energy are e and a_1. In all of the cases of Fig. 6, the cation vacancy is either a triple shallow acceptor or a deep trap for both an level in the gap. If it were the case that a cation vacancy near an interface had all of its sublevels resonant with the valence band of the superlattice, then the vacancy would be a triple shallow acceptor. If the vacancy were near the center of the GaAs layer, the vacancy would be a deep trap for both electrons and holes. Because the cation vacancy lies so near the valence-band maximum, relatively small amounts of lattice relaxation or charge-state splittings of the defect levels could alter the predictions for the character of the impurities.

All other Ga site sp^3 bonded substitutional impurity T_2 levels will have similar position dependencies to the ones found here for the Ga vacancy. But if the energy levels are far away from the top of the valence band, the effect of the valence-band splitting will be smaller.

IV. SUMMARY

We have found that the basic physics of electronic structure and deep impurity levels is the same for [111] and [001] superlattices, with three major differences: (i) The topological difference of the bonding in the [111] growth direction and the [001] growth direction cause significant differences of electronic structure and defect state behavior in the superlattices: [111] superlattices almost always have direct band gaps, quantum-confinement effects are weaker, and therefore shallow-deep transitions for impurities such as Si or N occur at smaller GaAs quantum-well sizes. Thus, the [111] GaAs/$Al_xGa_{1-x}As$ superlattices may prove to be better materials for electronic applications than the [001] GaAs/$Al_xGa_{1-x}As$ superlattices. (ii) The point-group symmetry of substitutional impurities in [111] superlattices is C_{3v} rather than C_{2v}, as in [001] superlattices. Therefore bulk T_2 impurity levels are split into one nondegenerate a_1 level and one doubly degenerate level e, instead of three nondegenerate levels. (iii) The larger e-a_1 splitting of the top valence band leads to larger splittings of the bulk T_2 levels, especially for those in GaAs layers and close to the top of valence band. We hope the theoretical work presented will stimulate more experiments on deep levels in superlattices.

ACKNOWLEDGMENT

We are grateful to the U. S. Air Force Office of Scientific Research (Contract No. AF-AFOSR-85-0331), and the U. S. Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support.

[1] L. Esaki and R. Tsu, IBM Res. Note RC-2418 (1969).

[2] R. Dingle, H. L. Störmer, A. C. Gossard, and W. Wiegmann, Appl. Phys. Lett. 33, 665 (1978).

[3] For a review of deep level theories, see, J. D. Dow, in *Highlights of Condensed Matter Theory* (*Proceedings of the International School of Physics "Enrico Fermi" Course 89, Varenna, 1983*), edited by F. Bassani, F. Fumi, and M. P. Tosi (North-Holland, Amsterdam, 1985), pp. 465 *et seq.* and references therein.

[4] S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B 38, 10677 (1988).

[5] J. D. Dow, S. Y. Ren, and J. Shen, in NATO Conference Proceedings, edited by C. Y. Fong (in press).

[6] For example, W. I. Wang, T. S. Kuan, E. E. Mendez, Phys. Rev. B 31, 6890 (1985); J. P. Faurie, A. Million, and J. Piaguet, Appl. Phys. Lett. 41, 713 (1982); L. A. Kolodziejski, T. Sakamoto, R. L. Gunshor, and S. Datta, Appl. Phys. Lett. 44, 799 (1984); R. N. Bicknell, R. Yanka, N. C.

Giles-Taylor, D. K. Banks, E. L. Buckland, and J. F. Schetzina, Appl. Phys. Lett. 45, 92 (1984).
[7]T. Hayakawa, K. Takahashi, M. Kondo, T. Suyama, S. Yamamoto, and T. Hijikata, Phys. Rev. Lett. 60, 349 (1988).
[8]C. Mailhiot and D. L. Smith, Phys. Rev. B 35, 1242 (1987).
[9]S. Y. Ren and J. D. Dow (to be published).
[10]A preliminary account of this work was given by J. D. Dow and S. Y. Ren, Bull. Am. Phys. Soc. 33, 630 (1988).
[11]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[12]Landolt-Börnstein, in *Semiconductors: Physics of Group IV and III-V Compounds*, Vol. 17a of *Numerical Data and Functional Relationships in Science and Technology*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer, Berlin, 1982). The room-temperature lattice constant of AlAs is actually 5.660 Å, but we take it to be equal to the GaAs lattice constant of 5.653 Å.
[13]D. J. Wolford, T. F. Kuech, J. A. Bradley, M. A. Grell, D. Ninno, and M. Jaros, J. Vac. Sci. Technol. B 4, 1043 (1986).
[14]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980); see also W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977) for the concepts that form the foundation of this work.
[15]S. Y. Ren and J. D. Dow, Phys. Rev. B 38, 1999 (1988).
[16]D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).
[17]We use the same impurity potential for Si in GaAs as in Ref. 4.
[18]M. Chandrasekhar, H. R. Chandrasekhar, A. Kangarlu, U. Venkateswaran, F. A. Chambers, and J. M. Meese, Superlatt. Microstruct. 4, 107 (1988). These authors reported the first experimental confirmation of the Si deep impurity level in a GaAs multiquantum well, as we predicted, Ref. 4.
[19]Mathematically we use the condition $G_{im}(E) = 0$ to find the vacancy energy levels, here $G_{im}(E)$ is the on-site Green's function of the host semiconductor with a specific symmetry. Under the central cell perturbation approximation, the coefficient of a deep defect state wave function at the defect site, C_{im} is determined by (Ref. 20): $(C_{im})^2 = -[G_{im}(E)]^2/[dG_{im}(E)/dE]$. Because $dG_{im}(E)/dE$ is always smaller than zero, the condition $C_{im} = 0$ naturally is equivalent to $G_{im}(E) = 0$, i.e., the widely used approximation of M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969), $V = \infty$.
[20]S. Y. Ren, Sci. Sinica 27, 443 (1984).
[21]The reason is that at an interface of a [001] superlattice, either one or the other of the T_2-derived b_1 and b_2 levels of the As vacancy will be within the GaAs layer; the other one will be within the $Al_xGa_{1-x}As$ layer. But at an interface of a [111] superlattice, neither the a_1 nor the e As T_2-derived vacancy T_2 state will be completely within the GaAs or the $GaAs/Al_xGa_{1-x}As$ layer.

d states, exchange splitting, and Mn electronic configuration in $Cd_{1-x}Mn_xTe$

A. Franciosi
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

A. Wall
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455 and IBM Corporation, Rochester, Minnesota 55901

Y. Gao and J. H. Weaver
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

M.-H. Tsai and J. D. Dow
Department of Physics, Notre Dame University, Notre Dame, Indiana 46556

R. V. Kasowski
Central Research and Development Department, E. I. du Pont de Nemours and Company, Experimental Station, Wilmington, Delaware 19898

R. Reifenberger and F. Pool
Department of Physics, Purdue University, West Lafayette, Indiana 47907
(Received 10 August 1989)

We examine the electronic structure of the prototypical semimagnetic semiconductors $Cd_{1-x}Mn_xTe$ with a combination of inverse photoemission, resonant synchrotron-radiation photoemission, and local-density pseudofunction model calculations. We show that the Mn $3d$ states lie outside the fundamental gap, the $d\uparrow$-$d\downarrow$ exchange splitting is 8.3 eV (almost twice as large as expected from earlier theoretical estimates), and that the ground-state configuration of Mn in the solid is primarily $(d\uparrow)^5(s\uparrow)(p\uparrow)$ rather than $(d\uparrow)^5s^2$.

The novel magnetotransport and magneto-optical properties of ternary semimagnetic semiconductors depend critically on the character and location of the Mn-derived d states, which are still quite controversial.[1] In this paper we present what is, to our knowledge, the first investigation of the excited electron states of a semimagnetic semiconductor by means of inverse-photoemission spectroscopy.[2] When combined with new photoemission spectroscopy results for the occupied electronic states and new self-consistent local-density pseudofunction[3] calculations for zinc blende MnTe, our results force a reevaluation of current ideas about the magnitude of electron-electron correlation and the electronic configuration of Mn in ternary semimagnetic semiconductors.

Our measurements were performed on $Cd_{1-x}Mn_xTe$ single crystals cleaved *in situ* following the methodology described in Refs. 1, 4, and 5. Inverse-photoemission measurements were performed in the bremsstrahlung isochromat spectroscopy (BIS) mode[2] at a photon energy of 1486.6 eV with a spectrometer resolution of 0.7 eV.[5] The photoemission measurements were performed in the angular integrated mode and photoelectron energy distribution curves (EDC) were obtained with overall energy resolution from 0.15 ($h\nu$=40 eV) to 0.45 eV ($h\nu$=120 eV).[6]

Two methods can be used to analyze the $3d$ contribution to the valence density of states (DOS) of semimagnetic semiconductors: they involve ternary-binary valence-band EDC difference curves[1] and resonance-antiresonance EDC difference curves obtained at Mn $3p$-$3d$ resonance-antiresonance photon energies.[7-9] In Fig. 1(a) we show an example of the first method for a $Cd_{0.55}Mn_{0.45}Te$ alloy.[10] The binding energies in Fig. 1(a) are referred to the top of the valence-band maximum E_v as derived from a linear extrapolation of the valence-band edge.[11] In Figs. 1(b) and 1(c) we show examples of resonance-antiresonance difference curves for $Cd_{0.55}Mn_{0.45}Te$ and $Cd_{0.80}Mn_{0.20}Te$, respectively.[12] The Mn $3d$ contribution in Figs. 1(a)–1(c) gives rise in all cases to a dominant emission feature 3.5 eV below E_v, a low binding-energy shoulder in the 0–2-eV range, and a broad satellite in the 6–9-eV range. The results of Figs. 1(b) and 1(c) are in agreement with those of Ref. 8, and the relatively weak concentration dependence of the Mn $3d$ features indicates that they mostly reflect the unchanging local Mn-Te coordination. An interpretation of these features has been given in terms of first principles one-electron calculations[13,14] that neglect final-state effects and exhibit a relatively poor agreement with experiment,[8,13,14] or using a semiempirical configuration interaction cluster model[9] that can address the screening of the $3d$ hole, but cannot directly incorporate Mn $3d$-Te $5p$ hybridization effects in the initial state.

We reevaluate here the merits of the one-electron picture on the basis of our new calculations of the electronic structure of antiferromagnetic zinc blende MnTe. Extended x-ray-absorption fine-structure investigations[15]

 ©1989 The American Physical Society

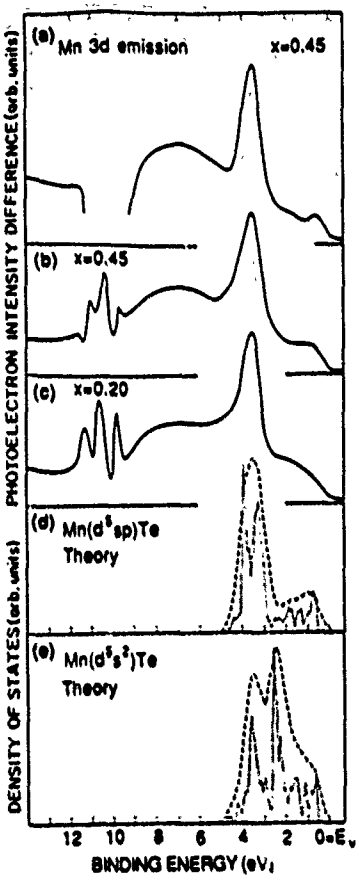


FIG. 1. (a) Ternary-binary difference curve for the valence-band emission of $Cd_{0.55}Mn_{0.45}Te$-CdTe at the Mn $3p$-$3d$ resonance ($h\nu$=50 eV). (b) Resonance-antiresonance difference curve ($h\nu$=50 and 47 eV) for $Cd_{0.55}Mn_{0.45}Te$ emphasizing the Mn $3d$ contribution to the valence band. (c) Resonance-antiresonance difference curve ($h\nu$=50 and 47 eV) for $Cd_{0.80}Mn_{0.20}Te$. (d) Theoretical l-projected density of states showing the Mn $3d$ character in antiferromagnetic zinc blende MnTe. The ground state corresponds to a Mn configuration close to d^5sp (solid line). Broadening with a Gaussian instrumental window function (dashed line) facilitates comparison with experiment. (e) Mn $3d$ character in antiferromagnetic zinc blende MnTe for the artificial d^5s^2 configuration.

and the photoemission data of Fig. 1 indicate that such a hypothetical compound should reflect the same Mn-Te local coordination and $3d$ contribution to the electronic structure found in the ternary alloys. The electronic structure calculations employed an r-space pseudofunction method in the local-density approximation, which we have modified to handle spin polarization.[3] We used a unit cell containing two Mn atoms and two Te atoms,[14] and considered both antiferromagnetic and ferromagnetic phases of Mn. The spin-dependent basis set employed naturally allows hypothetical zinc blende MnTe to adopt a ground state of arbitrary spin polarization in each atomic cell.[16] Our total-energy calculations favored the antiferromagnetic ground state with a Mn–Te bond length of 2.68 Å (as compared with the experimental value of 2.73 Å), an indication that the calculations had converged to correct physics. We used 1815 plane waves of each spin in our basis set and two special points[17] to approximate integration over the first Brillouin zone. The calculations indicate that the ground state of MnTe corresponds to highly spin-polarized Mn-derived conduction states, with Mn atomic configuration close to $(d\uparrow)^5(s\uparrow)(p\uparrow)$, i.e., the Hund's-rule ground state, rather than the customary $(d\uparrow)^5s^2$ ground state of the atom.[18]

We show in Fig. 1(d) the projected $3d$ density of states for the ground state of MnTe. The agreement in Fig. 1 between the one-electron theory and the photoemission results is unprecedented for diluted magnetic semiconductors, and only the broad satellite in the 6–9-eV range is not accounted for in the calculations.[19]

To confirm the importance of Mn $(d\uparrow)^5(s\uparrow)(p\uparrow)$ bonding and spin polarization of both valence and conduction bands, we show in Fig. 1(e) the Mn $3d$ spectral density of states from a calculation for MnTe in which the basis set and the spin-polarization were restricted such that the ground-state Mn electron configuration is close to $(d\uparrow)^5(s\uparrow)(s\downarrow)$, with some contribution from the spin-unpolarized configurations $(d\uparrow)^5(s\uparrow)(p\downarrow)$, $(d\uparrow)^5(s\downarrow)(p\uparrow)$, and $(d\uparrow)^5(p\uparrow)(p\downarrow)$. The result is similar to that of previous calculations,[13,14] but we emphasize that this result does not correspond to the ground state of MnTe and it predicts a major $3d$ feature at 2.4 eV that is not observed experimentally. We conclude that the Mn electronic configuration in the $Cd_{1-x}Mn_xTe$ alloy series is $(d\uparrow)^5(s\uparrow)(p\uparrow)$ rather than $(d\uparrow)^5s^2$.

In Fig. 2(a) we show BIS spectra for CdTe and $Cd_{0.80}Mn_{0.20}Te$ in the 0–16-eV energy range above the Fermi level E_F. The position of E_F at the surface was 0.62 ± 0.1 and 0.75 ± 0.1 eV above the valence-band maximum for $Cd_{0.80}Mn_{0.20}Te$ and CdTe, respectively. Each spectrum in Fig. 2(a) was obtained as the sum of quantitatively consistent spectra from 20 different cleaves.[20] Charging effects were observed for all samples with $x \geq 0.45$. The spectra were normalized for comparison to the CdTe DOS feature at 6.7 eV,[21] and subtracted from each other in order to obtain the difference curve in Fig. 2(b). Results for the x=0.35 alloy are qualitatively consistent with those for the x=0.20 alloys.[22,23]

The BIS spectrum for CdTe is in good quantitative agreement with the results of DOS calculations by Chelikowsky and Cohen.[24] A first structure at about 4 eV originates from states in the first two conduction bands along Δ and near X in the Brillouin zone. The structure at 6.7 eV primarily derives from states in low-symmetry directions of the Brillouin zone. A double emission feature in the 9–10-eV range originates from states in the third conduction band at L, and from a combination of states from higher bands in the Λ direction.[23,24] In $Cd_{1-x}Mn_xTe$ new MnTe-related features are observed in Fig. 2 at 4.2, 10.2, 12.2, and 13.8 eV (open squares) partially superimposed to the binary DOS features. Because of the normalization employed, minima and maxima in the difference curve of Fig. 2(b) correspond to binding energies at which the $Cd_{1-x}Mn_xTe$ DOS is small or large, respectively, as compared to the CdTe DOS. It is therefore reasonable to compare the difference curve in Fig. 2(b) with the results of our calculations for MnTe.

We show in Fig. 2(b) (solid line) the total DOS above E_F for the antiferromagnetic zinc blende ground state of MnTe. For comparison we also give in Fig. 2(c) the

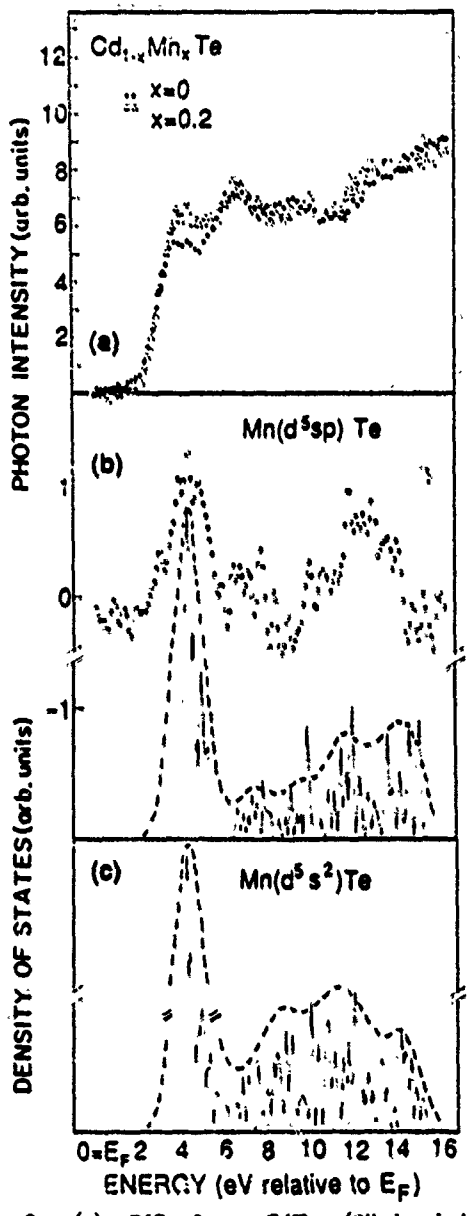


FIG. 2. (a) BIS from CdTe (filled circles) and $Cd_{0.50}Mn_{0.50}Te$ (open triangles). The conduction-band minimum E_C is 0.9 eV above the Fermi level E_F for CdTe, 1.2 eV for $Cd_{0.5}Mn_{0.5}Te$. (b) A difference curve (open squares) derived from the spectra in Fig. 2(a) is compared with the result of local density functional calculations of the total density of states above E_F (solid line) for antiferromagnetic zinc blende MnTe. The theoretical spectrum has been rigidly shifted to align the calculated features to the experimental ones. Convolution with a Gaussian broadening function (dashed line) facilitates comparison with experiment. (c) Theoretical total density of states for antiferromagnetic zinc blende MnTe where Mn is in the artificial low-spin d^5s^2 configuration.

analogous result for the reduced spin-polarization (Mn d^5s^2) case. The theoretical DOS is also shown after convolution with a 1.5-eV-wide Gaussian broadening function (dashed line) to account for experimental resolution, lifetime, and alloy scattering effects.[25] In accord with the usual practice of local-density theory, the calculated total DOS has been rigidly shifted[26] to higher binding energy to align the experimental and the theoretical main spectral features. We observe a fair agreement between theory and experiment. The difference curve exhibits a minimum at 8.5 eV and two broad maxima in the 2.0–6.0 and 10–14-eV energy ranges and it is in somewhat better agreement with the theoretical d^5sp curve [minima at 6 and 8.5 eV, maxima in the (3–6) and (10–15)-eV ranges] than with the d^5s^2 curve [minimum at 6.5 eV, maxima in the (3–6)- and (8–15)-eV ranges]. We emphasize however, that one could not conclusively rule out the d^5s^2 configuration just on the basis of the results of Fig. 2. Angular momentum projection of the DOS allows us to identify the 4.2-eV feature as primarily Mn 3*d* derived. The emission features at 10.2, 12.2, and 13.8 eV are associated primarily with Mn *p* states with some admixture of Te *p* character.

We emphasize that in Fig. 2(b) we see no evidence of the Mn-derived states in the gap predicted by earlier theories.[13,14,27] BIS studies of Mn thin films (1–5 Å) on $Cd_{1-x}Mn_xTe$ (Ref. 22) clearly show *d* emission in the gap, instead, when *unreacted* Mn is present. We can now estimate the $d\uparrow$-$d\downarrow$ exchange splitting from the position of the main Mn 3*d* features relative to the Fermi level, and obtain $3.5+0.62+4.2=8.32\pm0.4$ eV. Recent first-principles calculations for $Cd_{0.50}Mn_{0.50}Te$ alloys[14] predicted a Mn $3d\uparrow$ band $\sim$2.5 eV below E_F and a $d\downarrow$ feature within 2 eV of E_F, giving a splitting of 4.5 eV that is approximately half (54%) of our experimental value of 8.32 ± 0.4 eV and is substantially smaller than our new theoretical value of 6.8 eV. We note that for Mn atoms in AgMn intermetallics[28] the experimental splitting between majority- and minority-spin states has also been found larger (50%) than the value obtained from conventional local-density calculations (3.3 eV).

Earlier semiempirical calculations are in better agreement with the experimental exchange splitting. Ehrenreich *et al.*[29] predicted a 7.0-eV exchange splitting in their tight-binding calculations for $Cd_{1-x}Mn_xTe$. The semiempirical cluster model[9] fixed the value of the Coulomb correlation energy at about -8 eV. Such a value is in good agreement with our experimental result, and comparable with those assumed for Mott insulators such as MnO (9 eV) (Ref. 9) for which conventional local-density functional calculations also systematically underestimated the band gap and the $d\uparrow$-$d\downarrow$ splitting. The similar energy position of the 6–9-eV many-electron satellites[19] in $Cd_{1-x}Mn_xTe$ and MnO is then seen to be a reflection of the similar 3*d* correlation energy in these materials.

Our results indicate that optical transitions from the top of the *sp* valence states to the unoccupied *d* states should involve a photon energy of 4.8 ± 0.3 eV in the absence of final-state effects. Structure at 4.5 eV has been indeed observed in ellipsometry studies of $Cd_{1-x}Mn_xTe$,[30] and Kendelewicz[30] tentatively interpreted structure at 4.5 eV in reflectivity measurements as deriving from $sp\rightarrow d\downarrow$ transitions. However, in most recent studies, optical transitions at 2.2 eV have been interpreted as interband transitions from the top of the valence band to the unoccupied $d\downarrow$ states,[14,27] or as intra-atomic *d*-*d* excitations,[14] and large final-state effects have been called upon to explain a presumed discrepancy between photoemission and optical results, and photoemission and local-density functional results. We find little evidence of such a discrepancy in our

results, and suggest that the interpretation of the optical structure proposed in Refs. 14 and 27 should be reexamined. In particular, our results strongly support the interpretation of the optical spectra proposed by Kendelewicz,[30] and indicate that an alternate explanation should be found for the optical structure at lower energy. In this connection we mention that Larson *et al.*[13] recently contended that the optical structure at 2.2 eV could be alternatively explained in terms of standard (non-*d*-related) interband transition, or Mn^{2+} excitoniclike excitations that do not appear in the band-structure results.

In summary, inverse and resonant photoemission results for $Cd_{1-x}Mn_xTe$, indicate a $d\uparrow$-$d\downarrow$ exchange splitting almost twice as large as expected from earlier first-principles calculations and support a spin-polarized d^5sp-like ground state for Mn in these materials.

This work was supported in part by the Office of Naval Research and by U.S. Defense Advanced Research Projects Agency. We thank J.K. Furdyna for providing the crystals used in this study. The assistance of the staff of the University of Wisconsin Synchrotron Radiation Center, a National Science Foundation facility, is gratefully acknowledged.

[1]See, A. Franciosi, in *Diluted Magnetic (Semimagnetic) Semiconductors*, edited by R. L. Aggarwal, J. K. Furdyna, and S. von Molnar, Materials Research Symposia Proceedings, Vol. 89 (Materials Research Society, Pittsburgh, 1987), p. 175.

[2]See, for example, F. J. Himpsel and Th. Fauster, J. Vac. Sci. Technol. A 2, 15 (1984).

[3]R. V. Kasowski *et al.*, Phys. Rev. B 34, 2656 (1986).

[4]A. Wall *et al.*, J. Vac. Sci. Technol. A 5, 2051 (1987).

[5]Y. Gao *et al.*, J. Phys. E 21, 489 (1988).

[6]The experimental resolution and the position of the Fermi level were obtained from BIS and photoemission spectra of Pd and Mn thin films deposited onto the sample *in situ*.

[7]A. Franciosi, C. Caprile, and R. Reifenberger, Phys. Rev. B 31, 8061 (1985); A. Franciosi, R. Reifenberger, and J. K. Furdyna, J. Vac. Sci. Technol. A 3, 124 (1985).

[8]M. Taniguchi *et al.*, Phys. Rev. B 33, 1206 (1986); L. Ley *et al.*, *ibid.* 35, 2839 (1987).

[9]M. Taniguchi *et al.*, Solid State Commun. 62, 431 (1987).

[10]Spectra for the ternary alloy and for CdTe at resonance ($h\nu$=50 eV) have been normalized to the integrated intensity of the Cd 4*d* cores scaled by a $(1-x)$ factor derived from x-ray microprobe analysis, and subtracted from each other.

[11]The linear extrapolation method may induce a systematic error of up to 0.25 eV in the position of E_F. Such an error would not affect the measured value of the $d\uparrow$-$d\downarrow$ exchange splitting.

[12]Such spectra have been obtained from EDC's at resonance ($h\nu$=50 eV) and antiresonance ($h\nu$=47 eV) after normalization of the spectra to the integrated Cd 4*d* intensity. Such a normalization does not take into account the variation with photon energy of the Cd 4*d* and of the Te 5*s* emission, both present in the 9–11-eV binding-energy range, or small variations in the experimental resolution. Consequently, features related to such variations are observed in the 9.5–12 binding-energy range.

[13]B. E. Larson *et al.*, Phys. Rev. B 37, 4137 (1988).

[14]S.-H. Wei and A. Zunger, Phys. Rev. B 35, 2340 (1987); Phys. Rev. Lett. 56, 2391 (1986).

[15]A. Balzarotti *et al.*, Phys. Rev. B 31, 7526 (1985).

[16]U. von Barth and L. Hedin, J. Phys. C 5, 1629 (1972).

[17]D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).

[18]The calculations also explain the anomalous Curie-constant data for $Cd_{1-x}Mn_xTe$ in terms of spin polarization of the valence-band *s* and *p* states by Mn: see M.-H. Tsai *et al.*, Solid State Commun. 69, 1131 (1989).

[19]We interpret this feature as a multielectron satellite analogous to those observed in MnO and NiO. See, for example, D. E. Eastman and J. L. Freeouf, Phys. Rev. Lett. 34, 395 (1975) and G. K. Wertheim and S. Hüfner, *ibid.* 28, 1028 (1972). Only the configuration-interaction calculation for a $MnTe_4^{6-}$ cluster (Ref. 9) provides structure in that part of the spectrum, due mostly to an unscreened hole and a d^4 final state.

[20]BIS at x-ray energies provides direct information on the bulk DOS of semiconductors with little evidence of intrinsic- or damage-induced surface effects. See, for example, W. B. Jackson and J. W. Allen, Phys. Rev. B 37, 4618 (1988); and J. R. Chelikowsky, T. J. Wagener, J. H. Weaver, and A. Jin, *ibid.* B 40, 9644 (1989). Only at the longest exposure times examined (8 h) some broadening of the experimental features was occasionally observed, probably as the result of beam-induced damage.

[21]The relative amplitude of the different maxima in Fig. 2(b) depends somewhat on the normalization procedure employed.

[22]A. Wall *et al.* (unpublished).

[23]These results are in agreement with preliminary BIS results at ultraviolet energies of $Cd_{1-x}Mn_xTe$ alloys and MnCdTe(110) interface reaction products.

[24]J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B 14, 556 (1976).

[25]H. J. Levinson, F. Greuter, and E. W. Plummer, Phys. Rev. B 27, 727 (1983).

[26]The rigid shift used in Fig. 2, however, is 1.6 eV higher than what would be necessary to align the calculated *s*-*p* band gap to the value of 3.2 eV expected for MnTe.

[27]Recent linear muffin-tin orbital calculations for antiferromagnetic zinc blende MnTe by M. Podgorny, Z. Phys. B 69, 501 (1988) suggest a value of 4–4.5 eV for the $d\uparrow$-$d\downarrow$ splitting; earlier calculations for paramagnetic zinc blende MnTe by J. Masek, B. Velicky, and V. Janis, Acta Phys. Pol. A 69, 1107 (1986) yielded a value of 5.5 eV.

[28]D. van der Marel, G.A. Sawatzky, and F. U. Hillebrecht, Phys. Rev. Lett. 53, 206 (1984).

[29]H. Ehrenreich *et al.*, in *Proceedings of the Eighteenth International Conference on the Physics of Semiconductors*, edited by O. Engstrom (World Scientific, Singapore, 1987), p. 1751.

[30]P. Lautenschlager *et al.*, Phys. Rev. B 32, 3811 (1985); T. Kendelewicz, J. Phys. C 14, L407 (1981).

Electronic structure and deep impurity levels in structure-modulated zinc-blende–wurtzite semiconductor superlattices

Shang Yuan Ren
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
and Department of Physics, University of Science and Technology of China, Hefei, Anhui, China*

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
(Received 28 November 1988)

A structure-modulated superlattice is a layered material whose alternate layers are composed of materials with different crystal structures. Structure-modulated superlattices can be fabricated which have no composition modulation: for example, the zinc-blende–wurtzite CdS semiconducting superlattice. We present a theory of such zinc-blende–wurtzite structure-modulated superlattices, and find the following for a nearest-neighbor tight-binding model with ideal wurtzite c/a ratio: (i) The fundamental band gap of the superlattice equals those of both the zinc-blende and the wurtzite parent structures; (ii) in the superlattice growth direction, the dispersion relations $E_n(k)$ for the electrons and holes are the same as in the parent materials; (iii) in directions perpendicular to the superlattice growth direction, the superlattice dispersion relations $E_n(k)$ lie near those of the parent zinc-blende and wurtzite materials, and (iv) deep-level energies are almost the same as in the parent materials. The zinc-blende–wurtzite superlattice is especially interesting, because the interface between zinc-blende and wurtzite structures is ambiguous.

I. INTRODUCTION

Man-made superlattices fall into three categories: (i) composition-modulated superlattices such as zb-GaAs/zb-AlAs (where the prefix "zb-" indicates zinc-blende), (ii) structure-modulated superlattices such as zb-CdS/*w*-CdS (where "*w*-" denotes wurtzite), and (iii) combined composition- and structure-modulated superlattices such as zb-ZnTe/*w*-ZnSe. To date, most research has focused on composition-modulated superlattices. Here we present a theory of zinc-blende–wurtzite structure-modulated superlattices, which are grown along the [111] zinc-blende direction and the wurtzite c axis[1]. It is well known that numerous II-VI compound semiconductors, such as ZnS, ZnSe, CdS, and CdSe exhibit both zinc-blende and wurtzite bulk crystal structures,[2] and hence are candidates for fabrication of zinc-blende–wurtzite superlattices. Dilute magnetic semiconductor alloys, such as $Zn_{1-x}Mn_xS$ and $Zn_{1-x}Mn_xSe$, are also potential constituents of zinc-blende–wurtzite superlattices. Here we treat the case of the CdS zinc-blende–wurtzite structure.

The zinc-blende–wurtzite superlattice has a particularly interesting topological feature in that it does not contain an unambiguous interface between the zinc-blende and wurtzite structures. For example, in Fig. 1, one may select the wurtzite layer as extending either from plane E to plane I, or from plane G to plane K, or from plane H to plane L. With any of these choices, the layer thicknesses of the wurtzite layer and the zinc-blende layer are the same.

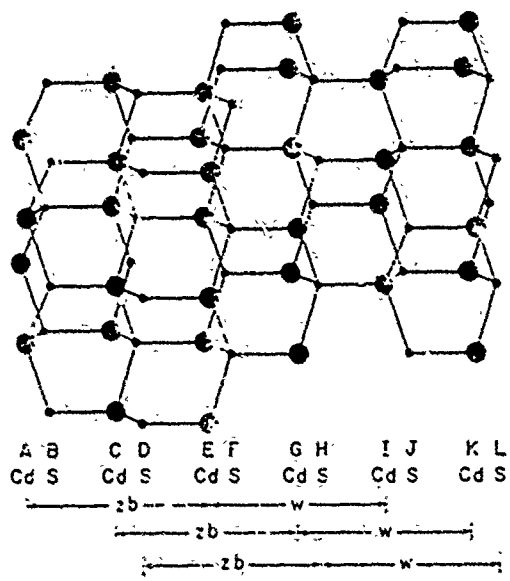


FIG. 1. Illustrating the structure of a 2×1 zinc-blende–wurtzite structure-modulated CdS superlattice with the period of two layers of zinc-blende and one layer of wurtzite structure. The large (small) circles denote Cd (S) atoms. The atomic planes are labeled A–L, with alternating planes containing only Cd or S atoms. The ambiguity of the zinc-blende–wurtzite interface is illustrated by noting various different layers which can be considered to be zinc-blende (zb) or wurtzite (w). For example, planes A–E (or C–G, or D–H) can be considered to be zinc-blende, while E–I (or G–K, or H–L) can be considered to be wurtzite.

 ©1989 The American Physical Society

II. FORMALISM

A. Host Hamiltonian

The growth direction of our CdS zinc-blende–wurtzite superlattice is defined as the direction "3" and is assumed to coincide with the zinc-blende [111] direction and the wurtzite c axis. The orthogonal directions "1" and "2" correspond to the zinc-blende $[2\bar{1}\bar{1}]$ and $[01\bar{1}]$ direction respectively. The superlattice we consider has N_{zb} two-atom-thick layers of zinc-blende structure and N_w four-atom-thick layers of wurtzite structure repeated periodically; the zinc-blende and wurtzite structures are assumed to be perfectly lattice matched. We denote this superlattice as an $N_{zb}\times N_w$ zinc-blende–wurtzite superlattice.

We first define a *superhelix* or supercell as a helical string with $2N_{zb}$ adjacently bonded atoms of zinc-blende structure with its axis aligned along the [111] direction and $4N_w$ adjacently bonded atoms of wurtzite structure with its axis aligned along the wurtzite-c direction, consisting of S, Cd, S, Cd, S, Cd, . . ., S, Cd. The center of the helix is at $\mathbf{L}$ and each of the atoms of the helix is at position $\mathbf{L}+\mathbf{v}_\beta$ (for $\beta=0,1,2,\ldots,2N_{zb}+4N_w-1$). A *superslab* of zinc-blende–wurtzite CdS consists of all such helixes with the same value of L_3 and all possible different values of L_1 and L_2; and the superlattice is a stacked array of these superslabs. If the origin of coordinates is taken to be at a S atom, the x and y axes are oriented such that a neighboring cation is at $(\frac{1}{4},\frac{1}{4},\frac{1}{4})a_L$, where a_L is the lattice constant of the zinc-blende structure.

We describe the electronic structure with a nearest-neighbor tight-binding Hamiltonian. At each site there are four sp^3 basis orbitals $|n,\mathbf{L},\mathbf{v}_\beta\rangle$, where $n=s, p_x, p_y$, or p_z and $\beta=0,1,2,\ldots,2N_{zb}+4N_w-1$. (Because the semiconductors exhibiting both wurtzite and zinc-blende structures also have large direct gaps, it is not necessary to employ an excited s^* basis orbital at each site. Such an orbital is useful for obtaining the indirect conduction-band edges found in indirect-band-gap semiconductors.[13]) In terms of these orbitals we form the tight-binding orbitals

$$|n,\beta,\mathbf{k}\rangle=N_s^{-1/2}\sum_{\mathbf{L}}\exp(i\mathbf{k}\cdot\mathbf{L}+i\mathbf{k}\cdot\mathbf{v}_\beta)\,|n,\mathbf{L},\mathbf{v}_\beta\rangle\,, \tag{1}$$

where $\mathbf{k}$ is any wave vector of the superlattice Brillouin zone. Here N_s is the number of supercells.

The minizone wave vector is a good quantum number, and so the tight-binding Hamiltonian is diagonal in $\mathbf{k}$. Evaluation of the matrix elements $\langle n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k}\rangle$ leads to a tight-binding Hamiltonian of the block tridiagonal form. For different β and β', the first three rows of block matrices are

$$\begin{vmatrix} H(0,0) & H(0,1) & 0 & \cdots & \cdots & 0 & H(0,2N_{zb}+4N_w-1) \\ H^\dagger(0,1) & H(1,1) & H(1,2) & 0 & \cdots & 0 & 0 \\ 0 & H^\dagger(1,2) & H(2,2) & H(2,3) & \cdots & 0 & 0 \\ \vdots & \vdots & & & & & \end{vmatrix}. \tag{2}$$

The last row of blocks is

$$H^\dagger(0,2N_{zb}+4N_w-1)\,0\,0\cdots 0\;\; H^\dagger(2N_{zb}+4N_w-2,2N_{zb}+4N_w-1)\;\; H(2N_{zb}+4N_w-1,2N_{zb}+4N_w-1)\,. \tag{3}$$

Here $H(\beta,\beta')$ depends on $\mathbf{k}$ and is given in terms of various 4×4 matrices for different n and n'.

The diagonal (in β) 4×4 matrix, $H(\beta,\beta)$ at site β, is

$$H(\beta,\beta)=\langle n,\beta,\mathbf{k}|H|n',\beta,\mathbf{k}\rangle=\begin{vmatrix} \epsilon_s & 0 & 0 & 0 \\ 0 & \epsilon_p & 0 & 0 \\ 0 & 0 & \epsilon_p & 0 \\ 0 & 0 & 0 & \epsilon_p \end{vmatrix}, \tag{4}$$

where the energies ϵ_s and ϵ_p in $H(\beta,\beta)$ are the atomic energy levels of the s and p states for the atom at the βth site, and may be obtained from tabulated parameters.[4] For a pure structure-modulated zinc-blende–wurtzite superlattice, we take the valence-band offset to be zero. That is to say, the cations and anions are assumed to have the same values of ϵ_s and ϵ_p independent of whether they lie in a zinc-blende layer or a wurtzite layer.

There are several distinct cases for which the off-diagonal (in β) matrix elements $\langle n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k}\rangle$ are nonzero (for $\beta\neq\beta'$).

1. Zinc-blende intrastructure matrix elements

If β and β' both refer to nearest-neighbor sites in the zinc-blende structure, we have, for example,

$$\langle n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k}\rangle=H_{c,zb;a,zb}\,, \tag{5}$$

if β refers to a cation and β' refers to an anion. $H_{c,zb;a,zb}$ is a 4×4 matrix whose rows and columns are labeled by n and n', which range over the values s, p_x, p_y, and p_z. Similarly we have matrix elements $H_{a,zb;c,zb}$. These matrix elements are

$$H_{c,zb;a,zb}=\begin{vmatrix} C_0V_1 & -C_0V_4 & -C_0V_4 & -C_0V_4 \\ C_0V_5 & C_0V_2 & C_0V_3 & C_0V_3 \\ C_0V_5 & C_0V_3 & C_0V_2 & C_0V_3 \\ C_0V_5 & C_0V_3 & C_0V_3 & C_0V_2 \end{vmatrix} . \quad (6)$$

Here we have $C_0=g_0^*$, where we have $4g_0=\exp(i\mathbf{k}\cdot\mathbf{x}_0)$, and $\mathbf{x}_0=(a_L/4)(1,1,1)$, with a_L being the lattice constant of zinc-blende CdS. Here k is the wave vector. In addition, we have

$$V_1=V(s,s)\,,$$
$$V_2=V(x,x)\,,$$
$$V_3=V(x,y)\,,$$
$$V_4=V(sa,pc)\,,$$

and

$$V_5=V(sc,pa)\,,$$

in the notation of Vogl *et al.*,[3] and

$$H_{a,zb;c,zb}=\begin{vmatrix} V_1(g_1+g_2+g_3) & V_4(g_1-g_2-g_3) & V_4(-g_1+g_2-g_3) & V_4(-g_1-g_2+g_3) \\ -V_5(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) & V_3(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) \\ -V_5(-g_1+g_2-g_3) & V_3(-g_1-g_2+g_3) & V_2(g_1+g_2+g_3) & V_3(g_1-g_2-g_3) \\ -V_5(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) & V_3(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) \end{vmatrix} . \quad (7)$$

We also have

$$4g_1=\exp(i\mathbf{k}\cdot\mathbf{x}_1)\,,$$
$$4g_2=\exp(i\mathbf{k}\cdot\mathbf{x}_2)\,,$$

and

$$4g_3=\exp(i\mathbf{k}\cdot\mathbf{x}_3)\,,$$

with $\mathbf{x}_1=(a_L/4)(1,-1,-1)$, $\mathbf{x}_2=(a_L/4)(-1,1,-1)$, and $\mathbf{x}_3=(a_L/4)(-1,-1,1)$. All of the matrix elements V are those tabulated[4] for CdS.

2. *Wurtzite intrastructure matrix elements*

The wurtzite structure has an exact hexagonal symmetry, and an approximate tetrahedral short-ranged symmetry: every atom is surrounded by a near tetrahedron of four atoms of the opposite species. The nearest-neighbor geometry is tetrahedral if the wurtzite structure has the ideal c/a ratio of $(8/3)^{1/2}=1.633$. In the ideal limit, the fundamental band gap of the wurtzite crystal structure is the same as for the zinc-blende structure.[5] Most wurtzite semiconductors are very close to the ideal limit; for example, CdS, CdSe, ZnS, and ZnTe have c/a ratios of 1.632, 1.630, 1.641, and 1.637, respectively.[2] In this paper we assume that the wurtzite is ideal, with the same bond length as its zinc-blende partner. As a result, the wurtzite matrix elements, in a nearest-neighbor tight-binding model, are related to the zinc-blende matrix elements, as follows.

If β and β' both refer to nearest-neighbor sites in the wurtzite structure, we will also have

$$(n,\beta,\mathbf{k}|H|n',\beta',\mathbf{k})=H_{c,w;a,w}\,. \quad (8)$$

Similarly we have matrix elements $H_{a,w;c,w}$. These matrix elements are

$$H_{c,w;a,w}=\begin{vmatrix} C_0V_1 & -C_0V_4 & -C_0V_4 & -C_0V_4 \\ C_0V_5 & C_0V_2 & C_0V_3 & C_0V_3 \\ C_0V_5 & C_0V_3 & C_0V_2 & C_0V_3 \\ C_0V_5 & C_0V_3 & C_0V_3 & C_0V_2 \end{vmatrix} . \quad (9)$$

Here the C_0 and V_i's are the same as defined before. This form is exactly the same as in (6). For the wurtzite structure we have two different forms of $H_{a,w;c,w}$ because there are four atoms in each unit cell (as opposed to two for zinc-blende); one is the same as in Eq. (7) and the other is

$$H_{a,w;c,w}=\begin{vmatrix} V_4(g_4+g_5+g_6) & V_4(-5g_4+g_5+g_6)/3 & V_4(g_4-5g_5+g_6)/3 & V_4(g_4+g_5-5g_6)/3 \\ -V_5(-5g_4+g_5+g_6)/3 & V_{21}g_4+V_{22}(g_5+g_6) & -V_{31}(g_4+g_5)+V_{32}g_6 & -V_{31}(g_4+g_6)+V_{32}g_5 \\ -V_5(g_4-5g_5+g_6)/3 & -V_{31}(g_4+g_5)+V_{32}g_6 & V_{21}g_5+V_{22}(g_4+g_6) & -V_{31}(g_5+g_6)+V_{32}g_4 \\ -V_5(g_4+g_5-5g_6)/3 & -V_{31}(g_4+g_6)-V_{32}g_5 & -V_{31}(g_5+g_6)+V_{32}g_4 & V_{21}g_6+V_{22}(g_4-g_5) \end{vmatrix} . \quad (10)$$

Here we have

$$V_{21}=(25V_{pp\sigma}+2V_{pp\pi})/27=V(x,x)+\tfrac{48}{27}V(x,y)\ ,$$

$$V_{22}=(V_{pp\sigma}+26V_{pp\pi})/27=V(x,x)-\tfrac{24}{27}V(x,y)\ ,$$

$$V_{31}=5V(x,y)/9\ ,$$

$$V_{32}=V(x,y)/9\ ,$$

$$4g_4=\exp(i\mathbf{k}\cdot\mathbf{x}_4)\ ,$$

$$4g_5=\exp(i\mathbf{k}\cdot\mathbf{x}_5)\ ,$$

and

$$4g_6=\exp(i\mathbf{k}\cdot\mathbf{x}_6)\ .$$

We also have $\mathbf{x}_4=(a_L/4)(-\frac{3}{3},\frac{1}{3},\frac{1}{3})$, $\mathbf{x}_5=(a_L/4)(\frac{1}{3},-\frac{3}{3},\frac{1}{3})$, $\mathbf{x}_6=(a_L/4)(\frac{1}{3},\frac{1}{3},-\frac{3}{3})$, and parameters V which are those tabulated[4] for CdS.

3. *Interstructure matrix elements*

The fact that the location of the zinc-blende–wurtzite interface is ambiguous means that one must (i) develop a formalism that treats the interfacial matrix elements the same, independent of interface choice, and (ii) select a specific interface for a particular calculation. There are two distinct forms for Hamiltonian matrix elements connecting one atom in the zinc-blende structure with another in the wurtzite structure. They are

$$H_{a,w;c,zb}=\begin{vmatrix} V_1(g_1+g_2+g_3) & V_4(g_1-g_2-g_3) & V_4(-g_1+g_2-g_3) & V_4(-g_1-g_2+g_3)\\ -V_5(g_1-g_2-g_3) & V_2(g_1+g_2+g_3) & V_3(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3)\\ -V_5(-g_1+g_2-g_3) & V_3(-g_1-g_2+g_3) & V_2(g_1+g_2+g_3) & V_3(g_1-g_2-g_3)\\ -V_5(-g_1-g_2+g_3) & V_3(-g_1+g_2-g_3) & V_3(g_1-g_2-g_3) & V_2(g_1+g_2+g_3)\end{vmatrix} \tag{11}$$

and

$$H_{c,w;a,zb}=\begin{vmatrix} V_1(C_1+C_2+C_3) & V_4(C_1-C_2-C_3) & V_4(-C_1+C_2-C_3) & V_4(-C_1-C_2+C_3)\\ -V_5(C_1-C_2-C_3) & V_2(C_1+C_2+C_3) & V_3(-C_1-C_2+C_3) & V_3(-C_1+C_2-C_3)\\ -V_5(-C_1+C_2-C_3) & V_3(-C_1-C_2+C_3) & V_2(C_1+C_2+C_3) & V_3(C_1-C_2-C_3)\\ -V_5(-C_1-C_2+C_3) & V_3(-C_1+C_2-C_3) & V_3(C_1-C_2-C_3) & V_2(C_1+C_2+C_3)\end{vmatrix}\ . \tag{12}$$

In the limit $N_{zb}=0$, this Hamiltonian reduces to the Hamiltonian of Ref. 4 for bulk wurtzite material. In the limit $N_w=0$ it becomes the sp^3 zinc-blende Hamiltonian.[3,6] Values of the parameters for this Hamiltonian can be taken from Ref. 4 or 3.

In this work we study electronic structure for superhelices as large as $N_{zb}+2N_w=20$; that is, in 40-atom-thick superslabs. The dimension of the Hamiltonian matrix at each value of k is $4(2N_{zb}+4N_w)$, because there are four orbitals per site. We diagonalize this Hamiltonian numerically for each k, finding its eigenvalues $E_{\gamma,\mathbf{k}}$ and, if necessary, the projections of the eigenvectors $|\gamma,\mathbf{k}\rangle$ on the $|n,\beta,\mathbf{k}|$ hybrid basis: $(n,\beta,\mathbf{k}|\gamma,\mathbf{k})$. Here γ is the band index (and ranges from 1 to 160 for $N_{zb}=2N_w=10$) and k lies within the superlattice mini-Brillouin zone.

B. Deep levels

The theory of deep levels is based on the Green's-function theory of Hjalmarson *et al.*,[7] which solves the secular equation for the deep-level energy E

$$\det[1-G(E)V]=0=\det\left[1-P\int\frac{\delta(E'-H)}{(E-E')}dE'\,V\right]\ . \tag{13}$$

Here V is the defect potential matrix,[7] which is zero except at the defect site and diagonal on that site (V_s,V_s,V_p,V_p) in the sp^3 local basis centered on each atom. We also have $G=(E-H)^{-1}$, where H is the host tight-binding Hamiltonian operator. The spectral density operator is $\delta(E'-H)$ and P denotes the principal-value integral over all energies. For energies E in the fundamental band gap of the superlattice, G is real.

A substitutional point defect in bulk zinc-blende CdS has tetrahedral (T_d) point-group symmetry. Each such s-and p-bonded defect normally has one s-like (A_1) and one triply degenerate p-like (T_2) deep defect level near or in the fundamental band gap. If we imagine breaking the symmetry of bulk zinc-blende CdS by making it into a zb-CdS$_2$/zb-CdS superlattice along the [111] direction, we reduce the T_d symmetry to C_{3v}. A substitutional point

defect in bulk wurtzite CdS also has C_{3v} point-group symmetry, and so in the zinc-blende–wurtzite superlattice and s- and p-bonded point defect will have C_{3v} symmetry as well and will produce two a_1 levels (one s-like, derived from the A_1 level and one T_2-derived p_σ level, and one doubly degenerate e level (p_π like).

The secular equation, Eq. (13), is reduced by the C_{3v} symmetry to the following two equations:

$$G(e;E)=V_p^{-1} \tag{14}$$

for e levels, and

$$\det\begin{vmatrix} G(s,s;E)V_s-1 & G(s,\sigma;E)V_p \\ G(\sigma,s;E)V_s & G(\sigma,\sigma;E)V_p-1 \end{vmatrix}=0 , \tag{15}$$

for a_1 levels, where we have

$$G(e;E)=\sum_{\gamma,\mathbf{k}} |(p_y,\beta,\mathbf{k}|\gamma,\mathbf{k})-(p_x,\beta,\mathbf{k}|\gamma,\mathbf{k})|^2/2(E-E_{\gamma,\mathbf{k}}) , \tag{16}$$

$$G(s,s;E)=\sum_{\gamma,\mathbf{k}} |(s,\beta,\mathbf{k}|\gamma,\mathbf{k})|^2/(E-E_{\gamma,\mathbf{k}}) , \tag{17}$$

$$G(\sigma,\sigma;E)=\sum_{\gamma,\mathbf{k}} |(p_x,\beta,\mathbf{k}|\gamma,\mathbf{k})+(p_y,\beta,\mathbf{k}|\gamma,\mathbf{k})\frac{1}{2}+(p_z,\beta,\mathbf{k}|\gamma,\mathbf{k})|^2/3(E-E_{\gamma,\mathbf{k}}) , \tag{18}$$

and

$$G(s,\sigma;E)=\sum_{\gamma,\mathbf{k}}[(s,\beta,\mathbf{k}|\gamma,\mathbf{k})]\times[(p_x,\beta,\mathbf{k}|\gamma,\mathbf{k})+(p_y,\beta,\mathbf{k}|\gamma,\mathbf{k})+(p_z,\beta,\mathbf{k}|\gamma,\mathbf{k})]^*/[\sqrt{3}(E-E_{\gamma,\mathbf{k}})] . \tag{19}$$

Here $G(\sigma,s;E)$ is the Hermitian conjugate of $G(s,\sigma;E)$ and β is the site of the defect in the superlattice.

For each site β the relevant host Green's functions, Eqs. (16)–(19), are evaluated using the special points method,[8] and the secular equations (14) and (15) are solved, yielding $E(e;V_p)$ and two values of $E(a_1;V_s)$, and $E(a_1;V_p)$. The defect potential matrix elements V_s and V_p can be considered as the same as the defect potential in the zinc-blende structure.

III. RESULTS AND DISCUSSIONS

A. Superlattice band structures

Once the Hamiltonian is defined, the superlattice band structure $E_{\gamma,\mathbf{k}}$ is determined by diagonalizing it for each wave vector k in its Brillouin zone[9] (Fig. 2), which is the same as the wurtzite Brillouin zone, except that the Γ–A length is reduced by a factor of $2/(N_{zb}+2N_w)$.

Provided the wurtzite structure has the ideal c/a ratio of $(8/3)^{1/2}$, the bulk zinc-blende, bulk wurtzite, and zinc-blende–wurtzite superlattice structures are physically indistinguishable when viewed only in the growth direction. The differences in three structures become apparent only when viewed along directions deviating from the growth axis. Therefore, for wave vectors k corresponding to the growth direction, the dispersion relations $E_n(\mathbf{k})$ and effective masses are the same as for the constituent zinc-blende and wurtzite materials. Hence zinc-blende–wurtzite superlattices made from direct-band-gap zinc-blende and (ideal c/a ratio) wurtzite materials are also direct-band-gap semiconductors with the same value of the fundamental gap at the Γ point of the Brillouin zone as the zinc-blende and the (ideal) wurtzite bulk materials: 2.60 eV for CdS. (Note that c/a for CdS is 1.632, only 0.05% different from the ideal ratio of $(8/3)^{1/2}=1.633$; hence the approximation of an ideal structure should introduce only small errors in the band structure, of order meV, the same order as found for InN.[10]) The lowest superlattice conduction band, as computed in the present model, is displayed for the 2×1 zinc-blende–wurtzite superlattice in Fig. 3; corresponding results for the highest valence band are given in Fig. 4. In the Γ–A or [111] direction, the dispersion relations $E_n(\mathbf{k})$ are the same as for the zinc-blende or the (ideal) wurtzite bulk materials, but in the Γ–M, or $[2,\bar{1},\bar{1}]$ direction and the Γ–K, or $[0,1,\bar{1}]$, direction the superlattice band structures deviate slightly from the bulk electronic structures.

A quasiparticle's resistance to acceleration by an applied electric or magnetic field is determined by its effective mass. In Fig. 5 we display the calculated conduction-band and valence-band effective masses in the Γ–A, Γ–M, and Γ–K symmetry directions, versus layer thickness of the superlattice. In our model, the conduction-band effective mass is unchanged in the superlattice, because the lowest conduction band is nondegenerate and its effective mass is isotropic for both zinc-

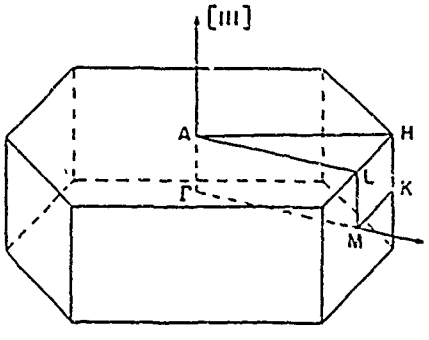


FIG. 2. The Brillouin zone for the 2×1 zinc-blende–wurtzite structure-modulated superlattice.

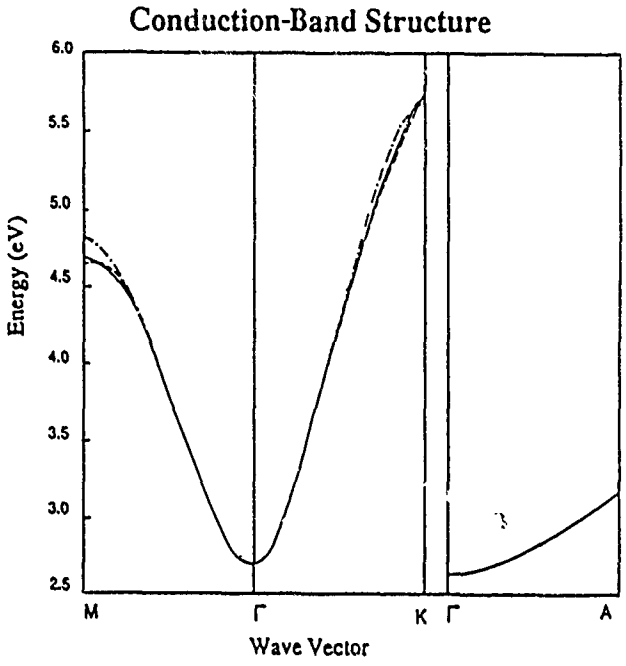


FIG. 3. The calculated lowest conduction-band structures in the three symmetry directions, $\Gamma-A$, $\Gamma-M$, and $\Gamma-K$, of the 2×1 zinc-blende–wurtzite superlattice (solid), bulk zinc-blende (dashed), and bulk wurtzite (dot-dashed) CdS. Note that the three are the same in the $\Gamma-A$ growth direction.

blende and wurtzite structures. The valence-band mass of the superlattice in the $\Gamma-A$ growth direction is independent of layer thickness and is the same as in either zinc-blende or wurtzite CdS. However, in the directions perpendicular to the growth direction, the valence-band masses change from the effective masses of wurtzite to

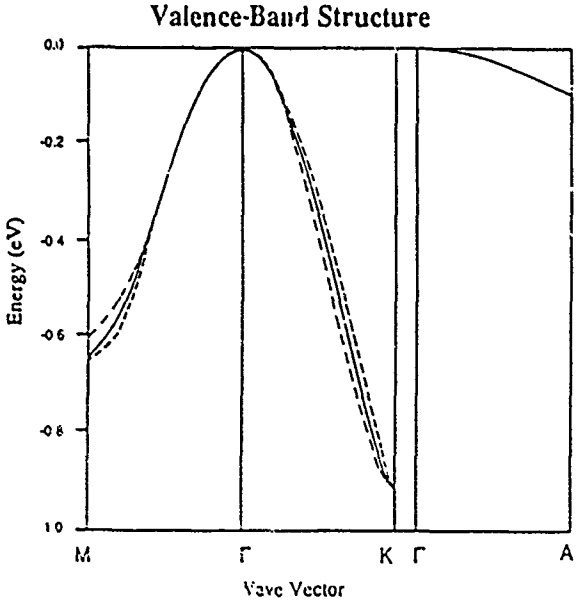


FIG. 4. The calculated highest valence-band structures in the three symmetry directions, $\Gamma-A$, $\Gamma-M$, and $\Gamma-K$, of the 2×1 zinc-blende–wurtzite superlattice (solid), bulk zinc-blende (dashed), and bulk wurtzite (dot-dashed) CdS.

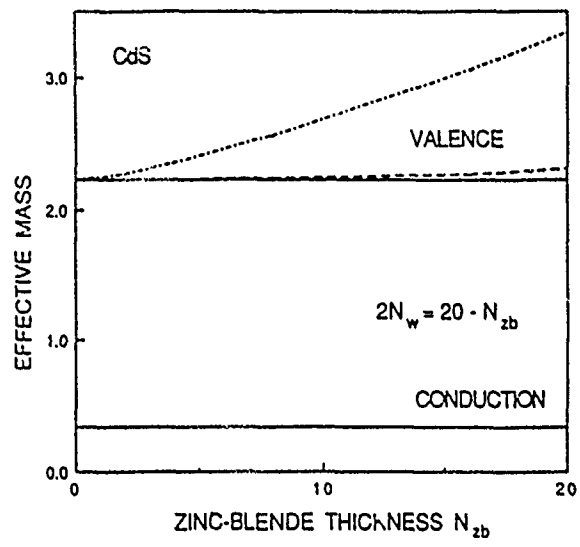


FIG. 5. The calculated reduced effective masses m^*/m_0 of the valence and conduction bands of a $N_{zb}\times N_w$ zinc-blende–wurtzite superlattice in the $\Gamma-A$ (solid), $\Gamma-M$ (dashed), and $\Gamma-K$ (dot-dashed) directions as functions of zinc-blende layer thickness N_{zb}, with $N_{zb}+2N_w=20$.

those of zinc-blende. The zinc-blende mass is anisotropic (because the valence-band maximum is triply degenerate and zinc-blende is a cubic structure) while the wurtzite mass is isotropic. The isotropy of the wurtzite valence-band mass is due to the combined effects of the lower symmetry of the wurtzite structure and the ideal c/a ratio. (For an ideal c/a ratio, the wurtzite valence-band maximum is triply degenerate, but for nonideal c/a it is nondegenerate with a doubly degenerate band slightly below it.)

B. Deep impurity levels in zinc-blende–wurtzite superlattices

Deep impurity levels are evaluated using Eq. (13) with the relevant parts of the spectral density operator $\delta(E'-H)$ expressed in terms of the Green's function (16)–(19). Since these Green's functions depend on the superlattice band structure, which differs only slightly from the bulk zinc-blende or wurtzite band structures, the deep-level energies in the CdS zinc-blende–wurtzite superlattice are virtually the same as for the same substitutional impurity in bulk zinc-blende or bulk wurtzite CdS.[11] Hence shallow-deep transitions as the layer thicknesses vary[12] are not to be expected for impurities in zinc-blende–wurtzite superlattices.

C. Long-ranged versus short-ranged order

Although the bulk zinc-blende and ideal-wurtzite structures are very different when viewed at long range, their short-ranged orders are both tetrahedral and identical (up to second-nearest neighbors). This short-ranged order leads to electronic structures of the two bulk materials that are almost identical, and to a zinc-blende–wurtzite superlattice electronic structure that is

almost the same as the bulk zinc-blende or wurtzite electronic structure. Clearly the short-ranged tetrahedral bonding, rather than any long-ranged zinc-blende, wurtzite, or superlattice order, is dominant in determining the electronic structure. While the concept of a structure-modulated superlattice is interesting, and the topology of the ambiguous zinc-blende–wurtzite interface is exciting, the practical consequences on electronic structure of growing such structures seem to be limited to introducing small anisotropies into the dispersion relations $E_n(\mathbf{k})$. In particular, such superlattices offer the possibility of introducing changes into the valence-band effective mass without significantly altering the conduction-band mass or mass isotropy.

Finally, we do expect some differences between the phonons in zinc-blende–wurtzite superlattices and the parent compounds. Although the materials will be identical when viewed along the $\Gamma-A$ (growth) direction, assuming an ideal c/a ratio, there will nevertheless be significant differences in the long-ranged Coulomb forces of zinc-blende and wurtzite structures, which should be reflected in the phonon dispersion curves.

IV. SUMMARY

The calculations presented here are, we believe, the first calculations of the electronic structures of structure-modulated semiconductor superlattices. We hope that this work will stimulate efforts to grow such interesting artificial materials.

ACKNOWLEDGMENT

We are grateful to the U.S. Air Force Office of Scientific Research (Contract No. AF-AFOSR-85-0331), Defense Advanced Research Projects Agency (Contract No. N0530-0716-09), and the U. S. Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support.

---

*Permanent address.

[1]A preliminary account of this work was given by S. Y. Ren and J. D. Dow Bull. Am. Phys. Soc. 33, 398 (1988).

[2]*Semiconductors: Physics of II-VI and I-VII Compounds, Semimagnetic Semiconductors*, Vol. 17b of *Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology, New Series*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer-Verlag, Berlin, 1982).

[3]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[4]A. Kobayashi, O. F. Sankey, S. M. Volz, and J. D. Dow, Phys. Rev. B 28, 935 (1983).

[5]J. L. Birman, Phys. Rev. 115, 1493 (1959), pointed out that an ideal-wurtzite semiconductor will have the same direct band gap as its zinc-blende partner.

[6]D. J. Chadi and M. L. Cohen, Phys. Status Solidi B 68, 405 (1975).

[7]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).

[8]S. Y. Ren and J. D. Dow, Phys. Rev. B 38, 1999 (1988).

[9]It may be noted that the zinc-blende structure can also have a hexagonal Brillouin zone; see S. Y. Ren and W. A. Harrison, Phys. Rev. B 23, 762 (1981). Therefore the zinc-blende–wurtzite superlattice has a hexagonal Brillouin zone as a natural consequence.

[10]M.-H. Tsai, D. W. Jenkins, J. D. Dow, and R. V. Kasowski, Phys. Rev. B 38, 1541 (1988).

[11]A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 28, 946 (1983). In the present work we use the values β_l from Ref. 7, rather than from this reference.

[12]S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B 38, 10677 (1988); R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Mater. Res. Soc. Symp. Proc. 77, 545 (1987).

Theory of local bond-length relaxation in $Hg_{1-x}Cd_xTe$ alloys

M.-H. Tsai, John D. Dow, and Kathie E. Newman
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

R. V. Kasowski
E. I. du Pont de Nemours and Company, Central Research and Development Department, Experimental Station, Wilmington, Delaware 19880-0356

(Received 17 November 1988; revised manuscript received 29 September 1989)

Local-density pseudofunction calculations of local Hg—Te, local Cd—Te, and average bond lengths in $Hg_{0.5}Cd_{0.5}Te$ are reported and compared with data and with other theories. Our results agree with extended x-ray-absorption fine-structure data and a phenomenological strain theory which show that the perfect compound-crystal bond lengths are virtually *conserved* in the alloy, but conflict with previous theories which predicted *large* bond relaxations *away* from the average bond length. The calculations imply that charge transfer from Cd to Hg is quite small in these alloys.

I. INTRODUCTION

Vegard's law for pseudobinary ternary alloys such as $Hg_{1-x}Cd_xTe$ states that the *average* bond length $d_{avg}(x)$, as measured by x-ray diffraction, varies linearly as a function of alloy composition x:

$$d_{avg}(x)=(1-x)d_{HgTe}(0)+xd_{CdTe}(1)\ .$$

Here the bond lengths d are related to the corresponding lattice constants a_L by a geometrical factor: $d(x)=\sqrt{3}/4a_L(x)$. In 1982 Mikkelsen and Boyce[1] showed that the *local* bond lengths (or nearest-neighbor distances) in pseudobinary alloys, as measured by extended x-ray-absorption fine-structure (EXAFS) spectroscopy, do not obey Vegard's law, but instead are nearly constant. As applied to $Hg_{1-x}Cd_xTe$, the Mikkelsen-Boyce result for the local Hg—Te and Cd—Te bond lengths $d_{HgTe}(x)$ and $d_{CdTe}(x)$ would be

$$d_{CdTe}(x)=(1-x)d_{CdTe}(0)+xd_{CdTe}(1)$$

and

$$d_{HgTe}(x)=(1-x)d_{HgTe}(0)+xd_{HgTe}(1)\ ,$$

where $d_{CdTe}(0)$ and $d_{HgTe}(1)$ have the physical interpretations of a Cd—Te impurity bond length in HgTe and a Hg—Te bond length in CdTe. In a first approximation, these bond lengths are unaffected by the alloy.

$$d_{CdTe}(x)\simeq d_{CdTe}(1)$$

and

$$d_{HgTe}(x)\simeq d_{HgTe}(0)\ .$$

Mikkelsen and Boyce, who studied III-V alloys, did find deviations from this first approximation which were linear in alloy composition x, normally with the property that the local bond lengths in the alloy relaxed *toward* the average. This idea, applied to $Hg_{1-x}Cd_xTe$, would have both the Hg—Te and Cd—Te local bond lengths expanding with increasing x, because CdTe has a larger lattice constant than HgTe. In 1985 and 1987 Sher *et al.*[2] and Hass and Vanderbilt[3] presented theories proposing that the local bond lengths relax *away* from the average in $Hg_{1-x}Cd_xTe$. Recently, Newman *et al.*[4] have developed a phenomenological strain theory for pseudobinary ternary alloys and applied it to III-V alloys. In this paper we evaluate the theory of Newman *et al.* for $Hg_{1-x}Cd_xTe$, and find that it predicts very small bond relaxations *toward* the average.

Thus there is a clear-cut disagreement among the theorists concerning both the sign and magnitude of the local Hg—Te and Cd—Te bond-length relaxations in $Hg_{1-x}Cd_xTe$ alloys. On one hand, Sher *et al.*,[2] using a tight-binding theory, and Hass and Vanderbilt,[3] using a norm-conserving pseudopotential method, have both predicted that the local Hg—Te and Cd—Te bond lengths, $d_{HgTe}(x)$ and $d_{CdTe}(x)$, (i) are significantly different in the $x=0.5$ alloy (by ~ 0.03 Å) from the perfect-crystal limits $d_{HgTe}(0)=2.796$ Å and $d_{CdTe}(1)=2.806$ Å, and (ii) relax *away* from one another as a function of alloy composition. That is, in the alloy we have, according to these theories, $d_{HgTe}(x)<d_{HgTe}(0)$ and $d_{CdTe}(x)>d_{CdTe}(1)$. On the other hand, the model of Newman *et al.*,[4] which uses a phenomenological strain theory, implies that the bond lengths are almost unchanged in the alloy and relax slightly *toward* one another. The predictions of the Newman theory are given in Fig. 1. While the conclusion of Newman *et al.* conforms with the intuitive expectation that bond lengths in the alloy should relax toward the average, the Newman theory is lacking the essential element which Sher *et al.* and Hass and Vanderbilt point to as being responsible for the counterintuitive relaxation, namely charge transfer in the alloy from Cd to Hg. Such charge transfer alters the Coulomb forces between ions and should contract bond lengths that would otherwise be expected to expand. However, both theories which propose relaxation away from the average have employed phenomenological non-self-consistent repulsive potentials to simulate the effects of d and core electrons and to produce (by construction) the observed bond lengths in the $x=0$ and 1 perfect-crystal limits. It is conceivable that

 ©1990 The American Physical Society

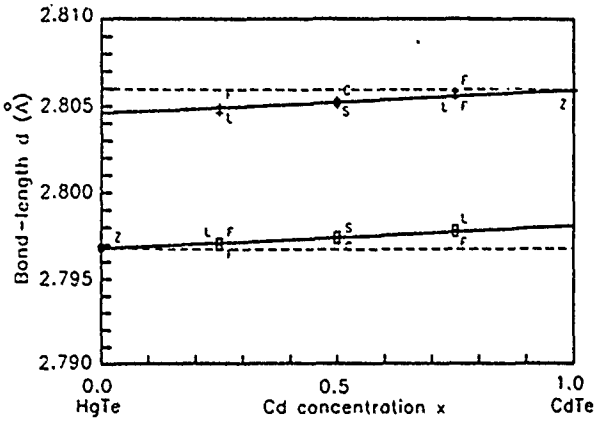


FIG. 1. Calculated Hg—Te bond lengths (squares) and Cd—Te bond lengths (pluses) in the $Hg_{1-x}Cd_xTe$ family of compounds. For $x=0$ and 1, the bond length is that to the zinc-blende compound; for $x=\frac{1}{4}$, the ordered compound is luzonite (L) with three Hg—Te bonds for every Cd—Te bond, or famatinite (F) with two long and one short Hg—Te bond(s) for every Cd—Te bond. A similar result holds for $x=\frac{3}{4}$. For $x=\frac{1}{2}$, the [001]-oriented superlattices (S) and chalcopyrite (C) compounds have equal numbers of Hg—Te and Cd—Te bonds. The solid lines are the values to be expected within the alloy, interpolated from the crystallite values, as discussed in Ref. 4. The dashed lines represent the perfect-crystal bond lengths of Hg—Te and Cd—Te, for comparison. (Cd—Te has the larger bond length.)

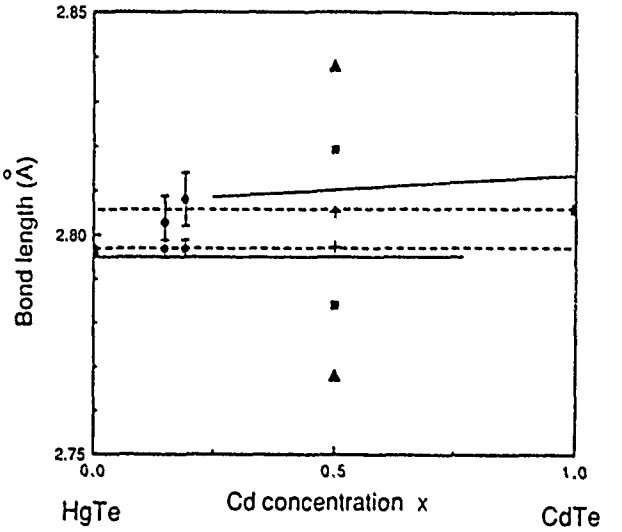


FIG. 2. Calculated local bond lengths $d_{HgTe}(x)$ and $d_{CdTe}(x)$ in $Hg_{1-x}Cd_xTe$, vs alloy composition x (solid lines), assuming straight lines defined by $d_{HgTe}(0.5)$, $d_{HgTe}(0)$, $d_{CdTe}(0.5)$, and $d_{CdTe}(1)$. The dashed lines represent the measured perfect-crystal bond lengths $d_{HgTe}(0)$ in HgTe and $d_{CdTe}(1)$ in CdTe, for reference. The EXAFS data are from Ref. 6. The phenomenological strain theory of Newman *et al.* at $x=\frac{1}{2}$ is displayed as pluses. Note that the present theory and the Newman theory are consistent with the data. The theoretical results of Sher *et al.* and of Hass and Vanderbilt are denoted by squares and triangles, respectively.

these phenomenological potentials have obscure unphysical features which inadvertently produce the large and anomalous relaxation of the bond lengths away from the average in the alloy.

On the experimental side, the recent EXAFS data of Bunker *et al.*[5,6] indicate that the Hg—Te and Cd—Te bond lengths in the $Hg_{1-x}Cd_xTe$ alloy are changed very little, if any, from the perfect-crystal values, and appear to be consistent with the predictions of Newman *et al.*, but not with those of Sher *et al.* and Hass and Vanderbilt. (See Fig. 2).

On the theoretical side, there is a clear need for a local-density-theory calculation[7] accurate enough to predict the $x=0$ and 1 perfect-crystal lattice constants and to address the question of bond-length relaxation in the alloys. Such a theory should treat all of the atom-atom interactions self-consistently on a first-principles basis.

In this paper we present local-density-theory[7] calculations of the bond lengths in the same $Hg_{0.5}Cd_{0.5}Te$ microcrystal model of the alloy as used by Hass and Vanderbilt.[3] We implement the calculations using the pseudofunction method.[8] Unlike Hass and Vanderbilt, we treat the outermost d-shell electrons of Hg and Cd as valence electrons, on the same footing as the bonding s and p electrons of Hg, Cd, and Te; thus we have no need for a phenomenological potential to simulate the effects of d electrons and can address directly the controversial issue in a completely *a priori* self-consistent model: Does charge transfer lead to *major* relaxation of the bond lengths in the alloy *away* from the average bond length?

II. EVALUATION OF THE NEWMAN THEORY FOR $Hg_{1-x}Cd_xTe$

The Newman theory models the alloy locally at selected compositions x in terms of microcrystallites. For example, famatinite and luzonite crystal structures correspond to local alloy compositions $x=\frac{1}{4}$ and $\frac{3}{4}$. The strain energies of these crystal types are expanded to second order in the bond elongations and the bond-angle distortions, with the coefficients of the expansions determined by elastic-constant data, similar to a Keating model.[9] The results of this theory for $Hg_{1-x}Cd_xTe$ are given in Fig. 1. Details of the theory can be found in Ref. 4.

III. CALCULATIONAL METHOD

All of the essential physics of this problem is linear in alloy composition x, and so the issues involved can be resolved at any single composition we choose (other than $x=0$ or 1). Three natural choices present themselves: $x=0.5$, $x\to 1$, and $x\to 0$. The latter two choices correspond to the single-impurity limit, where the issue becomes "is the bond length of a Hg impurity in CdTe *significantly smaller* than the perfect-crystal Hg—Te bond length, despite the fact that CdTe has a larger lattice constant than HgTe?" or "is the bond length of a Cd impurity in HgTe *significantly larger* than d_{CdTe} (1)?" However, it is slightly more convenient for us to determine bond lengths theoretically in the $x=0.5$ microcrystal model, which also allows a direct comparison with the work of Hass and Vanderbilt and Sher *et al.* Hence we

consider the bond lengths in $Hg_{0.5}Cd_{0.5}Te$.

Our calculations are based on local-density theory[7] and the pseudofunction method.[8] Local-density theory generally produces reliable changes of total energy, and so can be expected to predict structural properties. Its ability to predict band gaps and conduction-band structures is well known to be suspect, however.[10] The pseudofunction implementation of local-density theory has worked very well for a number of problems in semiconductor physics: It yielded a bond length in bulk Si only 0.5% different from the experimental value and a bulk modulus and its derivative with respect to pressure very close to the experimental ones.[11] It correctly described the Curie-constant data,[12] photoemission measurements,[13] and inverse photoemission spectra[13] of $Cd_{1-x}Mn_xTe$ semiconducting alloys. And it correctly predicted that a monolayer of potassium on the Si(001)-(2×1) surface[14] would form metallic chains each with Si—K bond lengths of 3.3 Å. (Subsequent experiments confirmed the predictions that the Si—K bond length is 3.14±0.1 Å and that the Si—K bond is weak,[15] as is to be expected of a metallic chain model.)

In the present study the outermost *d*-shell electrons of the cations, Hg and Cd, are treated as valence electrons, on the same footing as the bonding *s* and *p* electrons. The remaining cation electrons are treated as a spherically symmetric core. The electrons of the Te anions are separated into *s* and *p* valence electrons and a spherical core. The core charge densities of all atoms are obtained by solving the Schrödinger equation for the spherical part of the total potential; this is repeated each self-consistent iteration. The inclusion of the core charge densities in the iteration procedure is essential for obtaining accurate potentials and *d* bands, since the *d* orbitals have their sharp maxima deep in the core region. That is, a frozen-core approximation does not produce adequately accurate potentials and bond lengths. For the $Hg_{0.5}Cd_{0.5}Te$ alloy we use the same microcrystal model as that used by Hass and Vanderbilt.[3] Two geometric parameters, *a* and *u*, are adjusted to obtain the equilibrium lattice constant of the alloy, a_L, which is obtained from x-ray-diffraction data and is a linear interpolation of the lattice constants of HgTe and CdTe. We use the same microcrystal unit cell for HgTe and CdTe, but fix *u* so that Te atoms are at tetrahedral lattice points. As is usual in recent theoretical band-structure and total-energy calculations, we use two special points[3,16] to approximate the integration over the first Brillouin zone. Relativistic effects, except spin-orbit interactions, are included for all atoms.

IV. RESULTS

A. HgTe and CdTe crystals

We determined the zero-temperature, zero-pressure lattice constant of HgTe to be 6.455 Å [or $d_{HgTe}(0)=2.795$ Å], in comparison with the room-temperature experimental value of 6.459 Å (Ref. 18) [or $d_{HgTe}(0)=2.797$ Å]. Similarly, the calculated lattice constant of CdTe is 6.496 Å [$d_{CdTe}(1)=2.813$ Å], versus 6.480 Å [$d_{CdTe}(1)=2.806$ Å] measured at room temperature.[18] These lattice constants were obtained by fitting calculated total energies versus unit-cell volumes to Murnaghan's equation of state,[19]

$$E(V)=(B_0V/B_0')[(V_0/V)^{B_0'}+B_0'-1]/(B_0'-1)+\text{const}\ .$$

Here, V_0 is the equilibrium volume; B_0 is the bulk modulus and B_0' is its derivative.

Our calculated derivatives of the bulk moduli B_0 with respect to pressure, B_0', are in good agreement with available data, being 6.85 for HgTe and 6.70 for CdTe, versus 6.4±0.6 measured at room temperature[20] for CdTe. We find bulk moduli B_0 of 0.64 and 0.47 Mbar for HgTe and CdTe, compared with experimental values of 0.42 Mbar (at room temperature) (Ref. 21) and 0.49 Mbar (at 77 K) (Ref. 22) for CdTe and 0.43 Mbar (at room temperature) for HgTe. Hass and Vanderbilt[3] obtained bulk moduli of 0.047 and 0.13 Mbar for HgTe and CdTe, using an *ab initio* method which confined *d* electrons to their atomic cores. We also calculated the zone-center optical-phonon frequencies. They are 3.32 and 3.96 THz, respectively, for HgTe and CdTe, compared with experimental values of 3.64 and 4.20 THz.[21] Our calculations employed 9747 plane waves to expand the interstitial parts of the basis wave functions. For the nonspherical parts of the charge densities and potentials, 72 557 plane waves were employed.

B. $Hg_{0.5}Cd_{0.5}Te$ alloy

For the Hass-Vanderbilt microcrystal model of $Hg_{0.5}Cd_{0.5}Te$ we find an average lattice constant of 6.472 Å ($d_{avg}=2.802$ Å), essentially the same as the value extracted from measurements of HgTe and CdTe, combined with Vegard's law: $d_{avg}=0.5d_{HgTe}(x=0)-0.5d_{CdTe}(x=1)=2.801$ Å (or 6.470 Å for the lattice constant).[22] Hass and Vanderbilt[3] also did well in computing the lattice constant, having been off the experimental value by only 0.5% for $Hg_{0.5}Cd_{0.5}Te$. However, their result was predicted by an *adjusted* theory that treated the *d* electrons phenomenologically and, by construction, was exact for HgTe and CdTe. In contrast, the pseudofunction method described here is a fully *a priori* theory. We compute local bond lengths in the alloy of $d_{HgTe}(x=0.5)=2.795$ Å and $d_{CdTe}(0.5)=2.810$ Å. The difference $d_{CdTe}(x=0.5)-d_{HgTe}(x=0.5)$ is 0.015 Å, close to the experimental difference of 0.01 Å in $d_{CdTe}(1)-d_{HgTe}(0)$. The bond-length relaxation parameter ϵ is

$$\epsilon=\frac{d\text{HgTe}(x=0.5)-d_{CdTe}(x=0.5)}{d_{HgTe}(x=0)-d_{CdTe}(x=1)}$$

which is calculated to be $\epsilon=0.83$. From the EXAFS results of Pong and Bunker[6] on a number of pseudobinary ternary semiconductor alloys, ϵ typically lies in the range 0.6–0.8, but for $Hg_{1-x}Cd_xTe$ seems to be near unity. Hass and Vanderbilt obtained $\epsilon=5.8$ and Sher *et al.* found $\epsilon=2$.

IV. DISCUSSION

A. Bond-length relaxation

The present calculations reduce the uncertainty in the calculated *a priori* bond lengths in $Hg_{1-x}Cd_xTe$ alloys to about 0.01 Å from 0.3 Å. As a result, our local-density theory is consistent with the phenomenological strain theory of Newman *et al.*,[4] the lattice-constant data,[21] and the EXAFS data of Bunker *et al.*[5,6] Because our theory contains physics absent from the models of Sher *et al.* and of Hass and Vanderbilt, it implies that the theoretical uncertainties in those calculations are of comparable magnitude with the changes in bond lengths they predict. Thus the large bond-length relaxations away from the average are no longer to be expected in $Hg_{1-x}Cd_xTe$, and the apparent discrepancy between local-density theory and both the data and Newman's theory is removed.

We believe that the large uncertainties of the theories of Sher *et al.* and Hass and Vanderbilt are due to their introduction of non-self-consistent phenomenological repulsive potentials, a common feature of both theories. These potentials are adjusted to produce the observed lattice constants of HgTe and CdTe. (In the Hass-Vanderbilt theory, the zone-center optical-phonon frequencies were fitted as well.) In the model of Sher *et al.*, this phenomenological potential is rooted in Harrison's theory of the total energy[23] and accounts for the repulsive interactions between electronic charges. In the Hass-Vanderbilt theory, the repulsive potential compensates for having confined the *d* waves to the atomic cores. Such a potential may be adequate for the perfect crystals where all atoms occupy perfectly tetrahedral sites, but it should be regarded with suspicion in alloys—especially for calculations of bond lengths of four-digit accuracy. We believe that our approach, with an *a priori* self-consistent potential, is preferable for calculations requiring accurate bond lengths.

B. Charge transfer

The anomalous bond-length relaxation found by Sher *et al.* and by Hass and Vanderbilt was attributed by them to electronic charge transfer from Cd to Hg, which purportedly destabilizes the Hg—Te bond.

We have calculated the charges inside muffin-tin spheres of radius 1.4 Å for cations in HgTe, CdTe, and $Hg_{0.5}Cd_{0.5}Te$, and find them all to be the same, within 0.001 electron. The charge in a Te muffin-tin sphere of $Hg_{0.5}Cd_{0.5}Te$ is the same as the average for HgTe and CdTe, to within 0.001 electron. [The Te charge in CdTe is slightly larger than in HgTe, due to the larger (Pauling) ionicity of CdTe.] Thus, (i) the ionic charges are determined primarily by nearest-neighboring bonding, (ii) a virtual-crystal model of the $Hg_{1-x}Cd_xTe$ electronic structure is appropriate for the uppermost valence and the lowest conduction bands,[24] and (iii) we find no evidence of anomalous or large charge transfer in the $Hg_{1-x}Cd_xTe$ alloy. (Wei *et al.* reach a qualitatively similar conclusion.[25])

In summary, we believe that the extra self-consistency of our calculations, compared with previous ones, limits the charge transfer to 0.001 electron. Furthermore, the proper treatment of *d* electrons, their orbitals, and their hybridization with *s* and *p* valence electrons is an important feature of our approach. These two features combine to reduce the theoretical uncertainty for our local-density calculations and to provide a modicum of justification for theories such as that of Newman *et al.* which neglect charge transfer and consider only internal strains for computing bond lengths in alloys.

V. CONCLUSIONS

Our *ab initio* pseudofunction total-energy calculations, which treat the *d* electrons of Hg and Cd atoms self-consistently and on the same footing as valence *s* and *p* electrons, are able to obtain the Hg—Te and Cd—Te bond lengths in HgTe, CdTe, and $Hg_{0.5}Cd_{0.5}Te$ accurately to within a few tenths of a percent. Our results show that, to a good approximation, the Hg—Te and Cd—Te bond lengths are constant in $Hg_{1-x}Cd_xTe$ alloys, and that any bond-length relaxation is very small indeed.

A muffin-tin charge transfer analysis shows that charge transfer from Cd atoms to Hg atoms in $Hg_{0.5}Cd_{0.5}Te$ is insignificant, about 0.001 electron. Within the applicability of the micro-crystal model for $Hg_{1-x}Cd_xTe$ alloys, our results indicate that the properties of the nearest-neighbor bonding of the limiting crystals HgTe and CdTe are virtually preserved in these semiconducting alloys.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research (Contract Nos. N00014-84-K-0352 and N00014-85-J-0158) and the U.S. Defense Advanced Research Projects Agency (Contract No. N0530-0716-05) for their generous support. We are also very grateful to K. C. Hass for his helpful comments.

[1] J. C. Mikkelsen, Jr. and J. B. Boyce, Phys. Rev. Lett. 49, 1412 (1982).
[2] A. Sher, A. B. Chen, W. E. Spicer, and C. K. Shih, J. Vac. Sci. Technol. A 3, 105 (1985).
[3] K. C. Hass and D. Vanderbilt, J. Vac. Sci. Technol. A 5, 3019 (1987).
[4] K. E. Newman, J. Shen, and D. Teng, Superlatt. Microstruct. 6, 245 (1989).
[5] B. A. Bunker, J. Vac. Sci. Technol. A 5, 3003 (1987).
[6] W. F. Pong, R. A. Mayanovic, and B. A. Bunker, Physica B 158, 617 (1989).
[7] R. V. Kasowski, M.-H. Tsai, T. N. Rhodin, and D. D. Chambliss, Phys. Rev. B 34, 2656 (1986).
[8] L. Hedin and B. I. Lundqvist, J. Phys. C 4, 2064 (1971).
[9] P. N. Keating, Phys. Rev. 145, 637 (1966); R. M. Martin, Phys. Rev. B 6, 4546 (1972).
[10] M. S. Hybertsen and S. G. Louie, Phys. Lett. 55, 1418 (1985).
[11] M.-H. Tsai, J. D. Dow, and R. V. Kasowski, Phys. Rev. B 38,

2176 (1988).
[12]M.-H. Tsai, J. D. Dow, R. V. Kasowski, A. Wall, and A. Franciosi, Solid State Commun. 69, 1131 (1989).
[13]A. Franciosi, A. Wall, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, R. V. Kasowski, R. Reifenberger, and F. Pool, Phys. Rev. B 40, 12 009 (1989).
[14]M.-H. Tsai and R. V. Kasowski, Bull. Am. Phys. Soc. 32, 865 (1987); R. V. Kasowski and M.-H. Tsai, Phys. Rev. Lett. 60, 546 (1988).
[15]T. Kendelewicz, P. Soukiassin, R. S. List, J. C. Woicik, P. Pianetta, I. Lindau, and W. E. Spicer, Phys. Rev. B 37, 7115 (1988).
[16]D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).
[17]Since thermal-expansion and pressure corrections for the lattice constants have opposite signs, the difference between zero-pressure, zero-temperature, one-atmosphere, and room-temperature lattice constants is commonly small: for example, in bulk Si the zero-temperature, zero-pressure bond length is 2.3508 Å, compared with 2.3516 Å at room temperature and pressure.
[18]*CRC Handbook of Chemistry and Physics*, 66th ed., edited by R. C. Weast, M. H. Astle, and W. H. Beyer (Chemical Rubber Co., Boca Raton, FL, 1985/86).
[19]R. D. Murnaghan, Proc. Natl. Acad. Sci. USA 30, 244 (1944); M. T. Yin and M. L. Cohen, Phys. Rev. B 26, 5668 (1982).
[20]K. Strössner, S. Ves, W. Dieterich, W. Gebhardt, and M. Cardona, Solid State Commun. 56, 563 (1985).
[21]*Crystal and Solid State Physics*, Vol. 18 of *Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology*, edited by K. H. Hellwege (Springer, Berlin, 1984).
[22]Note that the solid lines of Fig. 2 are straight by construction.
[23]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).
[24]Y. Onodera and Y. Toyozawa, J. Phys. Soc. Jpn. 24, 341 (1986); K. C. Hass, H. Ehrenreich, and B. Velicky, Phys. Rev. B 27, 1088 (1983).
[25]S.-H. Wei and Z. Zunger, Phys. Rev. B 37, 8958 (1988); J. Vac. Sci. Technol. A 6, 2597 (1988).

Solid State Communications, Vol. 69, No. 12, pp.1131-1133, 1989.
Printed in Great Britain.

0038-1098/89 $3.00 + .00
Pergamon Press plc

EXPLANATION OF ANOMALOUS CURIE CONSTANTS OF $Cd_{1-x}Mn_xTe$ ALLOYS

M.-H. Tsai and John D. Dow
Department of Physics, University of Notre Dame
Notre Dame, Indiana 46556, U.S.A.
and
R. V. Kasowski
E. I. du Pont de Nemours and Company
Central Research and Development Department, Experimental Station
Wilmington, Delaware 19898, U.S.A.
and
A. Wall and A. Franciosi
Department of Chemical Engineering and Materials Science
University of Minnesota, Minneapolis, Minnesota 55455, U.S.A.

(Received 7 December 1988 by E. Tosatti)

The anomalous Curie-constant data for $Cd_{1-x}Mn_xTe$ alloys are explained in terms of spin-polarization of the valence band *s* and *p* states by Mn. The theory, which is based on an *a priori* spin-unrestricted pseudofunction local-density approximation, illustrates the consequences of Hund's first rule in solids.

1. Introduction

Recent magnetic susceptibility measurements of zinc-blende $Cd_{1-x}Mn_xTe$ alloys revealed Curie-law behavior, but with anomalous Curie constants that did not follow the expected dependence on alloy composition x for x>0.4 [1].

$$C = A \times S(S+1). \tag{1}$$

Here A is a constant and S is the spin of the Mn ion, which is 5/2 for the free Mn^{+2} ion in the $(3d)^5$ configuration, according to Hund's first rule [2]. For small Mn concentration x, Spalek et al. found the expected behavior, C=AxS(S+1), with a spin almost equal to the free-ion value, S = 2.47 ± 0.05, for $Cd_{1-x}Mn_xTe$. However, the Curie constants increased above their expected linear dependence on x for high Mn concentrations.

2. Effective-spin and Hund's first rule

In this paper, we shall present calculations which explain the anomalous super-linear x-dependence of the Curie constants of $Cd_{1-x}Mn_xTe$ in terms of spin-polarization of the valence bands by Mn. We shall write the Curie constant as

$$C = A \times S(x)\,[S(x)+1], \tag{2}$$

where we have the effective-spin

$$S(x) = (5/2) + \eta(x)/2. \tag{3}$$

The valence-band spin-polarization function $\eta(x)$ can be interpreted as the number of additional spins, beyond the five spins associated with the *d*-electrons, in the vicinity of a Mn nucleus of zinc-blende $Cd_{1-x}Mn_xTe$. Note that zinc-blende $Cd_{1-x}Mn_xTe$ has not been grown for x > 0.7 [1], but extrapolation to x=1 (hypothetical MnTe) of a least-squares fit of Spalek's data produces an experimental value $\eta^{exp}(x=1) = 1.2$. Moreover hexagonal, NiAs-structure MnTe exists with a measured spin-polarization function of $\eta^{exp}_{hex}(x=1) = 0.16$ [3]. (The smaller value of η may be attributed to the larger Mn-Te separation, 2.91 Å in the NiAs crystal structure, versus 2.73 Å in the zinc-blende structure.) Hence both hexagonal and hypothetical zinc-blende MnTe have "extra" electronic spins at each Mn site than can come from Mn *d*-electrons. Clearly this spin must be associated with spin-polarized *s* and *p* valence electrons. While there exist other calculations of the electronic structure of MnTe [4-7], to our knowledge none of those calculations have produced the significant spin-polarization necessary to explain the susceptibility and Curie-constant data.

In this paper, we report such a calculation, based on the local-density approximation [8] as implemented with the pseudofunction method [9], which we have modified to handle spin-polarization. The basic physical idea underlying the calculation is Hund's first rule [2]: A Mn atom in $Cd_{1-x}Mn_xTe$ will adopt a ground state that maximizes its total spin S(x). Although this rule strictly applies only to free magnetic atoms, it must have an analogue in solids. With this in mind, we

have developed our local-density theory with a spin-dependent basis set that naturally allows hypothetical zinc-blende MnTe to adopt a lowest-energy antiferromagnetic ground state of arbitrary spin-polarization in each atomic cell.

Neutral atomic Mn adopts a $(3d\uparrow)^5(4s\downarrow)(4s\uparrow)$ ground-state configuration, with its $4s$ shell being closed and spinless. The $3d$ electrons of atomic Mn all have parallel spins, according to Hund's first rule: maximizing the total spin requires a highly symmetrized spin state, and an orbital wave function that minimizes the electrons' mutual Coulomb repulsion [2]. A similar correlation effect exists in solids, although the electronic configuration is more complicated due to coupling between atoms: the dominant configuration in zinc-blende MnTe [10,11] is $(3d\uparrow)^5(4s\uparrow)(4p\uparrow)$.

In atomic Mn the spin-7/2 $(3d\uparrow)^5(4s\uparrow)(4p\uparrow)$ configuration lies at too high an energy relative to the spin-5/2 $(3d\uparrow)(4s)^2$ configuration, and so does not contribute appreciably to the ground state, even though it has a larger spin. But in condensed matter this large-spin configuration has to be considered for the ground state because the energy necessary to transfer an s valence electron to a p-state is small. (This is analogous to the case of C, which assumes the $(1s)^2(2s)^2(2p)^2$ atomic ground-state configuration, but finds a lower energy in condensed matter by forming sp^3 bonds from the $(1s)^2(2s\,2p)^3$ configuration.) In fact, the ground state of Mn in MnTe is a linear combination of configurations, including the spin-polarized configuration $(3d\uparrow)^5(4s\uparrow)(4p\uparrow)$ and the closed-shell configuration $(3d\uparrow)^5(4s)^2$, as well as others. Our calculations show that such a spin-polarized state has a lower total energy, as expected from Hund's first rule.

3. Method of calculation

We use the pseudofunction method [9] together with local spin-density theory [3] to compute the electronic structure and ground state energy of zinc-blende MnTe, $Cd_{0.5}Mn_{0.5}Te$, and $Cd_{0.75}Mn_{0.25}Te$. This method has correctly predicted the 0.5 Bohr magneton spin-saturation magnetization of Ni [12] and has been applied successfully to a number of problems concerning non-magnetic semiconductors [13,14].

Using a unit cell containing two Mn atoms and two Te atoms [4] to describe MnTe, we considered both antiferromagnetic (AF) and ferromagnetic (F) phases of Mn, by aligning the spins on nearest-neighbor Mn ions either antiparallel (AF) or parallel (F). Our total-energy calculations favored the (known) AF ground state by 0.17 eV/molecule, with a Mn-Te bond length of 2.68 Å (as compared with the experimental value of 2.73 Å [4,15]), an indication that the calculations had adequately converged to correct physics [16,17]. The resulting conduction bands, when projected onto a particular Mn site, were highly spin-polarized, a result that we did not anticipated.

We repeated the calculations for $Cd_{0.5}Mn_{0.5}Te$ and $Cd_{0.75}Mn_{0.25}Te$, using the standard per-Mn micro-crystal models: an eight-atom unit cell with the (001)-superlattice structure and a sixteen-atom unit cell with the luzonite structure, respectively [14,18].

4. Results

By summing the net spin-density of the localized s, p, and d orbitals centered on a Mn atom, we determined the valence-band spin polarization: $\eta(1) = 1.30$, $\eta(0.5) = 0.36$, and $\eta(0.25) = 0.04$. The minority-spin d-electrons actually contribute a small negative amount to $\eta(x)$.

We compare the results of our *a priori* calculations for the spin S(x) in a Mn cell with data in Fig. 1. The smooth line for S(x) is obtained by fitting to a parabola the calculated spin S(x) for three values of x: x=1, 0.5, and 0.25. The resulting parabola is $S(x)=S_0+ax+bx^2$, with S_0=2.41, a=0.34, and b=0.30. Note the excellent agreement between the *a priori* theory and the data.

Initially our studies of MnTe and $Cd_{1-x}Mn_xTe$ were motivated not by a desire to understand magnetic susceptibility and Curie-constant data but by photoemission experiments [10,11]. Our spin-unpolarized calculations produced density of states spectral features, for the d-bands in particular, that agreed with previous theories [4-7] but disagreed with the data by typically > 1eV. The introduction of spin-polarization produced (i) d-related photoemission peaks 3.5, 1.6, and 0.8 eV below the valence band maximum (in agreement with the data [10,11]), (ii) inverse photoemission spectra with the relative positions of s, p, and d features in agreement with the data [10,11], and (iii) triply-peaked Te $5s$ partial densities of states induced by spin-polarization, again

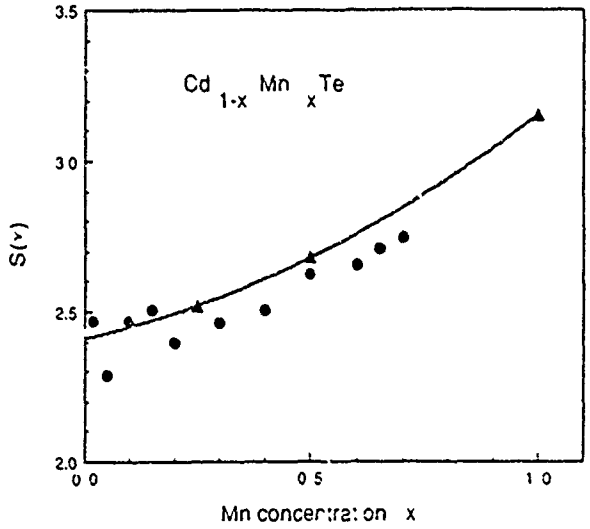


Fig. 1. Effective total spin S(x) of a Mn atom in zinc-blende $Cd_{1-x}Mn_xTe$ versus Mn concentration x. The solid line is the parabolic interpolation of the *a priori* calculations for x=1/4, x=1/2, and x=1 (solid triangles). The circles are extracted from Curie-constant data [1].

in agreement with the data [10,11]. Thus the theoretical framework developed to explain the spectroscopic data explains the *magnetic* properties of these materials as well.

One might ask why the present theory produces such dramatic spin-polarization effects when other local-density theories have not. We believe that the difference in theories ultimately is traceable to the basis sets. The pseudofunction basis allows one to obtain a Hund's-rule-like ground state because the basis contains realistic localized spin-unrestricted states. Plane-wave basis sets emphasize the interstitial regions of a crystal, and so, many plane waves are required to describe a localized *d*-function. To test our hypothesis that the difference in the results of the pseudofunction and plane-wave methods is attributable to the basis sets, we restricted our basis orbitals to be the same for spin-up and spin-down electrons. As expected, we reproduced the principal *d*-band feature at −2.4 eV below the valence band maximum found in previous plane-wave implementations of local-density theory [4,5]; this feature lies ≈1 eV lower in our unrestricted pseudofunction calculation and in the data. Hence we concluded that charge transfer involving orbitals with different spins is important in $Cd_{1-x}Mn_xTe$ and that such charge transfer, since it involves localized *d*-states, is better handled in a pseudofunction basis than in a plane-wave basis: The pseudofunction method handles the spin-polarization and the physics of Hund's first rule in a natural, transparent way.

5. Conclusion

The anomalous non-linear-in-x Curie constants of $Cd_{1-x}Mn_xTe$ alloys are attributed to spin-polarization of the valence electrons by the Mn. The amount of valence spin-polarization is correctly predicted by *a priori* spin-polarized pseudofunction local-density theory. This theory also correctly predicted that the ground state of MnTe is antiferromagnetic and removed a number of significant discrepancies between previous theories and photoemission and inverse photoemission spectra.

Acknowledgements– We are grateful to the U.S. Defense Advanced Research Projects Agency (Contract No. N0530-0710-05) and the U.S. Office of Naval Research (Contract No. N00014-84-K-0352) for their generous support of this research.

REFERENCES

[1] J. Spalek, A. Lewicki, Z. Tarnawski, J. K. Furdyna, R. R. Galazka, & Z. Obuszko, Phys. Rev. B33, 3407 (1986).

[2] See, e.g. G. Baym, in *Lectures on Quantum Mechanics*, (W. A. Benjamin, Inc., New York, 1969), p.458.

[3] *Handbook of Chemistry & Physics*, Ed. R. C. Weast, 62nd Edition, (CRC Press, Inc., Boca Raton, Florida, 1981-1982).

[4] S. H. Wei & A. Zunger, Phys. Rev. B35, 2340 (1987).

[5] B. E. Larson, K. C. Hass, H. Ehrenreich, & A. E. Carlsson, Solid State Commun. 56, 347 (1985).

[6] B. Velicky, J Masek, V. Chab, & B. A. Orlowski, Acta Phys. Polonica A69, 1059 (1986); J. Masek, B. Velicky, & V. Janis, Acta Phys. Polonica, A69, 1107 (1986); J. Phys. C 20 59 (1987); J. Masek & B. Velicky, phys. stat. sol. (b) 140, 135 (1987).

[7] M. Podgorny, Z. Phys. B69, 501 (1988).

[8] W. Kohn & L. J. Sham, Phys. Rev. 140, 1133 (1965); U. von Barth & L. Hedin, J. Phys. C 5, 1629 (1972).

[9] R. V. Kasowski, M.-H. Tsai, T. N Rhodin, & D. D. Chambliss, Phys. Rev. B34, 2656 (1986).

[10] A. Wall, A. Franciosi, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, & R. V. Kasowski, J. Vac. Sci. Technol., to be published; A. Franciosi, A. Wall, Y. Gao. J. H. Weaver, M.-H. Tsai, J. D. Dow, R. V. Kasowski, R. Reifenberger, & F. Pool, to be published.

[11] A. Franciosi, in *Diluted Magnetic (Semimagnetic) Semiconductors*, R. L. Aggarwal, J. K. Furdyna, & S. von Molnar, eds., (Materials Research Society, Pittsburgh, Pennsylvania, 1987).

[12] C. Kittel in *Introduction to Solid State Physics*, 5th Edition, (John Wiley & Sons, New York, 1976), p. 466.

[13] M.-H. Tsai, J. D. Dow, & R. V. Kasowski, Phys. Rev. B 38, 2176 (1988); M.-H. Tsai & R. V. Kasowski, Bull. Amer. Phys. Soc. 32, 865 (1987); R. V. Kasowski & M.-H. Tsai, Phys. Rev. Lett. 60, 546 (1988); M.-H. Tsai, R. V. Kasowski, & T. N. Rhodin, Surf. Sci. 179, 143 (1986); R. V. Kasowski, M.-H. Tsai, & J. D. Dow, J. Vac. Sci. Technol. B 5, 953 (1987); M.-H. Tsai, D. W. Jenkins, J. D. Dow, & R. V. Kasowski, Phys. Rev. B 38, 1541 (1988); M.-H. Tsai, R. V. Kasowski, & J. D. Dow, Solid State Commun. 64, 231 (1987).

[14] M.-H. Tsai J. D. Dow, K. E. Newman, & R. V. Kasowski, "Theory of local bond-length relaxation in $Hg_{1-x}Cd_xTe$ alloys," to be published.

[15] A. Balzarotti, M. Czyzyk, A. Kisiel, N. Motta, M. Podgorny, & M. Zimnal-Starnawska, Phys. Rev. B 30, 2295 (1984).

[16] The Mn-Te bond-length for the ferromagnetic ground state was 2.70 Å. We used 1815 plane waves of each spin in our basis set and two special points [17] to approximate integrations over the first Brillouin zone.

[17] D J. Chadi & M. L. Cohen, Phys. Rev. B8, 5747 (1973).

[18] K. C. Hass & D. Vanderbilt, J. Vac. Sci. Technol. A5, 3019 (1987).

Inverse photoemission and resonant photoemission characterization of semimagnetic semiconductors

A. Wall,[a] A. Franciosi, Y. Gao, and J. H. Weaver
Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota 55455

M.-H. Tsai and J. D. Dow
Department of Physics, Notre Dame University, Notre Dame, Indiana 46556

R. V. Kasowski
Central Research and Development Department, E. I. du Pont de Nemours and Company, Experimental Station, Wilmington, Delaware 19898

(Received 12 December 1988; accepted 9 January 1989)

The new magnetotransport and magneto-optical properties of the semimagnetic $Cd_{1-x}Mn_xTe$ semiconductor alloy series depend critically on the nature of the Mn-derived *d* states. We examine here the electronic structure of these alloys with a combination of inverse photoemission spectroscopy, core-level photoemission line-shape analysis, valence-band resonant photoemission, and local density pseudofunction theory. The spectroscopic data reflect the local Mn–Te coordination and are in remarkable agreement with our one-electron calculations. We see no evidence of Mn-derived *d* states in the gap, and observe an experimental d^5–d^4 exchange splitting of 8.4 ± 0.4 eV, i.e., almost twice as large as expected from earlier theoretical estimates. The ground-state configuration of Mn in the solid is primarily $(d\uparrow)(s\uparrow)(p\uparrow)$, and the super-exchange interaction has an important role in determining the stability of such a configuration relative to $(d\uparrow)^5s^2$.

I. INTRODUCTION

Ternary semimagnetic semiconductors[1–13] are alloys in which magnetic atoms such as Mn randomly replace some of the cations in a II–VI semiconductor lattice. The new phenomena that this class of materials bring to semiconductor science stem from the spin–spin exchange and super-exchange interactions[6,14] that involve the magnetic moments localized on the substitutional magnetic impurity, and the band and impurity states of the II–VI host.[9] Such interactions result in large *g* factors, giant magnetoresistance, large values of the Faraday rotations, and novel magnetotransport and magneto-optical properties.[9]

Critical parameters to model and understand the properties of these materials are the ground-state energy of the Mn $3d^5$ configuration, the 3*d* hybridization with anion-derived *p* states, the Coulomb correlation energy of the Mn 3*d* electrons, and the magnitude of the super-exchange interaction. Gunnarsson *et al.*[15] recently pointed out in a stimulating paper the many difficulties involved in calculating *ab initio* the renormalized Coulomb integrals in Anderson's Hamiltonian for the nonmetallic systems of interest here. Experiment has to provide the electronic information indispensable to model the properties of these materials and test the theoretical approximations.

In this paper we examine the prototypical $Cd_{1-x}Mn_xTe$ alloy system. We have described in a recent review[1] the controversy that surrounds the position and character of the *d* states in this system and emphasized how inverse photoemission may provide a crucial test of the calculations of the Coulomb interaction by sampling the density of unoccupied electron states above the Fermi level E_F. Here we present what is, to our knowledge, the first investigation of the excited electron states of a semimagnetic semiconductor by means of inverse photoemission spectroscopy. We also extend earlier photoemission spectroscopy studies of the electronic structure, and interpret the data in the light of new self-consistent local-density pseudofunction calculations for zinc-blende MnTe. Our results force a critical reevaluation of many current ideas about the electronic configuration of Mn, the magnitude of electron–electron correlation, and the validity of the one-electron approximation in ternary semimagnetic semiconductors. A short summary of selected results has been reported earlier.[16]

II. EXPERIMENTAL DETAILS

The crystals used in the present studies were grown at Purdue University through a modified Bridgman method, and characterized through x-ray diffraction and x-ray microprobe analysis as to crystallographic phase and Mn content. All results were obtained on single-phase, single crystals cleaved *in situ* ($x = 0$, 0.20, 0.35, 0.45, and 0.60) in a suitable spectrometer at operating pressures of about $\sim 5\times10^{-11}$ Torr. The data summarized here have been obtained from a number of cleavage surfaces of variable quality, ranging from highly disordered to flat mirrorlike surfaces. The data appear independent of cleave quality and provide information on the bulk electronic structure.

Inverse photoemission measurements were performed directing a collimated monochromatic electron beam at the cleaved surface and monochromatizing the emitted photons with a 0.5-m Rowland circle monochromator with a quartz grating and a microchannelplate detector.[17] The combination of the monochromator and a self-supporting $A\ell$ filter allowed only photons with energy $h\nu = 1480.6$ eV to reach

 0734-2101/89/030656-07$01.00 © 1989 American Vacuum Society

the detector. The energy of the emitted photons is determined by the radiative decay of the incoming electrons from initial-state energies E_i of the solid to final empty states at energy $E_f(h\nu = E_i - E_f)$. The final-state energies were scanned by varying the accelerating voltage of the Pierce-type electron gun while keeping constant the monochromator pass energy. The resulting spectra provide information on the total density of final states E_f weighted by an optical emission cross section that is a time-reversed photoemission cross section. This experimental technique has historically been termed bremsstrahlung isochromat spectroscopy[18] (BIS) and is finding application in the analysis of the excited electron states of metals, semiconductors, and solid surfaces.[19] The overall energy resolution of the spectrometer was ~0.7 eV, as determined from the experimental width of the Fermi level cutoff in the BIS spectra from Mn or Au thick films evaporated *in situ* on the sample.

Photoemission measurements were performed at the Synchrotron Radiation Center of the University of Wisconsin-Madison. Monochromatic synchrotron radiation obtained using a 3-m toroidal grating monochromator and the 800-MeV electron storage ring Aladdin was focused on the sample at an angle of ~45° from the sample normal. Angular integrated photoelectron energy distribution curves (EDC's) were obtained for $40 < h\nu < 120$ eV by means of a commercial hemispherical energy analyzer. The overall energy resolution (0.15–0.45 eV) and the position of the Fermi level were measured with the same method employed in the BIS experiment.

III. RESULTS AND DISCUSSION

A. Valence band

The use of synchrotron radiation photoemission allows one to choose photon energies in the region of the Mn $3p \rightarrow 3d$ transition and examine the Mn $3d$ contribution to the valence bands. Resonant photoemission at the Mn $3p$–$3d$ threshold yields a characteristic enhancement of the $3d$ cross section at resonance (50 eV) versus antiresonance (47 eV) that reflects the quantum-mechanical equivalence of different processes leading from the ground state to the same final state. We pioneered the use of this method to examine ternary semimagnetic semiconductors,[20] and more recently the same method has been applied by us and by other authors to $Cd_{1-x}Mn_xTe$.[4,8,21] One process is the direct excitation:

$$3p^6 3d^5 4(sp)^2 + h\nu \rightarrow 3p^6 3d^4 4(sp)^2 \epsilon f.$$

The second process involves a $3p$ core excitation and a super Coster–Kronig decay:

$$3p^6 3d^5 4(sp)^2 + h\nu \rightarrow 3p^5 3d^6 4(sp)^2 \rightarrow 3p^6 3d^4 (sp)^2 \epsilon f.$$

The interference of the two processes yields a characteristic Fano line shape in the overall excitation cross section.

Two methods used to analyze the $3d$ contribution to the valence density of states (DOS) involve ternary–binary valence-band EDC difference curves[4] and resonance–antiresonance EDC difference curves.[4] In Fig. 1(a) we show an application of the first method to a $Cd_{0.55}Mn_{0.45}Te$ alloy. Spectra for the ternary alloy and for CdTe at resonance ($h\nu = 50$ eV) have been normalized to the integrated intensity of the Cd $4d$ cores scaled by a $(1-x)$ factor derived from x-ray microprobe analysis, and subtracted from each other. The binding energies in Fig. 1(a) are referred to the top of the valence-band maximum E_v as derived from a linear extrapolation of the valence-band edge. The dip in the 9–11 eV region reflects the variation in the absolute intensity of the Cd $4d$ emission going from the ternary alloy to the binary. We did not perform a subtraction of the secondary background since such a process yields arbitrary variations in the relative intensity of the features in the 0–9 eV range. In Figs

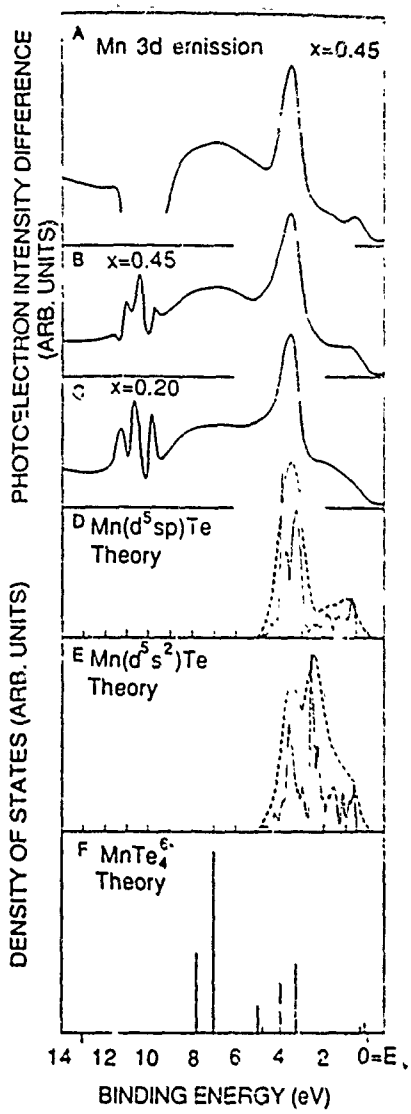


FIG. 1. (a) Ternary–binary difference curve for the valence-band emission of $Cd_{0.55}Mn_{0.45}Te$–CdTe at the Mn $3p$–$3d$ resonance ($h\nu = 50$ eV). (b) Resonance–antiresonance difference curve ($h\nu = 50$ and 47 eV) for $Cd_{0.55}Mn_{0.45}Te$ emphasizing the Mn $3d$ contribution to the valence band. (c) Resonance–antiresonance difference curve ($h\nu = 50$ and 47 eV) for $Cd_{0.80}Mn_{0.20}Te$. (d) Theoretical l-projected density of states showing the Mn $3d$ character in antiferromagnetic zinc-blende MnTe. The ground state corresponds to a Mn configuration close to $(d^5)^5(s^1)(p^1)$ (—). Broadening with a Gaussian instrumental window function (- - -) facilitates comparison with experiment. (e) Mn $3d$ character in antiferromagnetic zinc-blende MnTe for the $(d^5)^5s^2$ configuration. (f) Configuration interaction calculations of the photoemission final states for a $MnTe_4^{6-}$ cluster, from Ref. 4.

1(b) and 1(c) we show applications of resonance–antiresonance difference curve method to $Cd_{0.44}Mn_{0.44}Te$ and $Cd_{0.80}Mn_{0.20}Te$ samples, respectively. Such spectra have been obtained from EDC's at resonance ($h\nu = 50$ eV) and antiresonance ($h\nu = 47$ eV) after normalization of the spectra to the integrated Cd 4*d* intensity. Such a normalization does not take into account the variation with photon energy of the Cd 4*d* and of the Te 5*s* emission, both present in the 9–11 eV binding-energy range, or the small modulation in the relative intensity of the Cd 4*d* surface/bulk components, or the varying experimental resolution, which produce features in the difference curves in the 9.5–12 eV binding-energy range.

The different criteria for normalization notwithstanding, Figs. 1(a)–1(c) are remarkably consistent. The Mn 3*d* contribution gives rise to a dominant emission feature 3.5 eV below E_v, a smaller low-binding-energy shoulder in the 0–2 eV range, and a broad satellite in the 6–9 range. The results of Figs. 1(b) and 1(c) are in agreement with those of Refs. 4 and 21, and similar to what has been found for the selenide and sulfide series.[4,22] The lack of an important concentration dependence of the Mn 3*d* features suggests that they mostly reflect the constant Mn–Te bonding situation.[4,21]

Previous local density functional and tight-binding calculations of the 3*d* contribution to the valence states[6,7,23,24] for MnTe and $Cd_{1-x}Mn_xTe$ alloys showed only limited agreement with photoemission results. This, together with the interpretation of optical data,[5] suggested that large final-state effects might be present, and stimulated the introduction of the configuration interaction cluster model by Fujimori and co-workers.[4] This semiempirical model can address the screening of the 3*d* hole, but cannot incorporate directly Mn 3*d*–Te 5*p* hybridization effects in the initial state. We have performed new spin-unrestricted local-density calculations of the band structure of antiferromagnetic zinc blende MnTe. Recent extended x-ray absorption fine-structure (EXAFS) studies[3] indicate that such a hypothetical compound should reflect the same Mn–Te local coordination found in the ternary alloys. The lack of a strong composition dependence in the results of Fig. 1 suggests that the 3*d* contribution to the electronic structure should also be relatively similar in the ternary alloy and in the binary parent compound.

Our electronic structure calculations employed a spin-unrestricted *k*-space pseudofunction method in the local density approximation. Details on the method and a complete discussion of the results will be given in a longer forthcoming paper.[25] We found that the ground state of system is an antiferromagnetic phase of MnTe where the Mn atomic configuration is close to $(d\uparrow)^5(s\uparrow)(p\uparrow)$, as predicted by Hund's rule. Relative to the customary $(d\uparrow)^5(s)^2$ state, the spin maximization enhances the super-exchange interaction and modifies the Mn 3*d* and Te *s* binding energies. We show in Fig. 1(d) the projected 3*d* density of states for the ground state of MnTe (solid line). We also show (dashed line) the same results convoluted with a Gaussian ($\sigma = 0.30$ eV) to account for the effect of the experimental energy resolution. The agreement between theory and experiment is remarkable. If we focus on Fig. 1(a), i.e., on the spectrum obtained with the more accurate ternary/binary normalization procedure, 3*d* features are observed at 0.7, 1.6, and 3.5 eV. Correspondingly, the one-electron calculations forecast 3*d* features at 0.8, 1.6, and 3.5 eV [Fig. 1(d)]. Only the broad satellite in the 6–9 eV range is not accounted for in the calculations.

To examine the importance of the Mn electronic configuration and of the resulting additional spin polarization of valence and conduction bands, we calculated a MnTe electronic structure in which restriction of the basis set and spin polarization yielded a Mn ground-state electron configuration closer to $(d\uparrow)^5(s\uparrow)(s\downarrow)$, with some contribution from the $(d\uparrow)^5(s\uparrow)(p\downarrow)$, $(d\uparrow)^5(s\downarrow)(p\uparrow)$, and $(d\uparrow)^5(p\uparrow)(p\downarrow)$ configurations. The resulting Mn 3*d* character is shown in Fig. 1(e). The results of Fig. 1(e) are similar to those of previous calculations,[6,7] which forecast a major 3*d* feature at 2.4 eV that is not observed experimentally. The discrepancy between our results for the ground state and those of earlier calculations will be discussed in detail in a forthcoming paper, but we anticipate here that the increased total spin in our results allows a simple explanation of the anomalous Curie temperature observed in $Cd_{1-x}Mn_xTe$ alloys.[26]

The origin of the 6–9 eV satellite is, in our opinion, still controversial. Only the calculations by Fujimori[4] provide structure in this spectral region. In Fig. 1(f) we plot the result of the configuration interaction calculation for a $MnTe_4^{6-}$ cluster from Ref. 4. The multiplet lines derive from d^4 and $d^5\underline{L}$ final-state components, corresponding to an unscreened 3*d* hole or to a 3*d* hole screened by charge transfer from a Te 5*p*-derived ligand orbital $\underline{L}$ (π or σ) to a *d* orbital (e_g or t_{2g}). Comparison of Fig. 1(f) with Figs. 1(a)–1(c) shows good agreement, and, in fact, the addition of lifetime and experimental broadening[4] yields a perfect reproduction of the experimental results of Figs. 1(b) and 1(c). However, the model of Ref. 4 is semiempirical, in the sense that the energy differences between the final-state configurations and the lifetime parameters in Fig. 1(f) are not known *a priori*, but are fixed through comparison with experiment.

The cluster model does forecast a 6–9 eV satelite mostly due to the unscreened d^4 final state. It is not clear, however, why an unscreened satellite of this kind would exhibit a relative intensity increasing with Mn concentration, as observed in Fig. 1 and in a recent study of the electronic structure of metastable $Cd_{1-x}Mn_xTe$ alloys formed in a wide range of composition through reaction of Mn thin film with CdTe(110). We find that the intensity of the satellite relative to the main Mn 3*d* 3.5-eV feature increases as Cx, with $C = 0.50 \pm 0.05$ in the composition range explored ($0 < x < 0.95$).[27] At this stage the specific origin of the satellite remains to be ascertained, although one can state in general that in a band structure picture the 0.6-, 1.8-, and 3.5-eV features, should be assigned to a fully screened one-electron final state while the 6–9 eV features should be associated with a multielectron satellite analogous to the one observed at about the same energy in MnO[28] and NiO.[28]

B. Core level deconvolution

Another crucial test of first principle calculations for ternary semimagnetic semiconductors appears the influence of

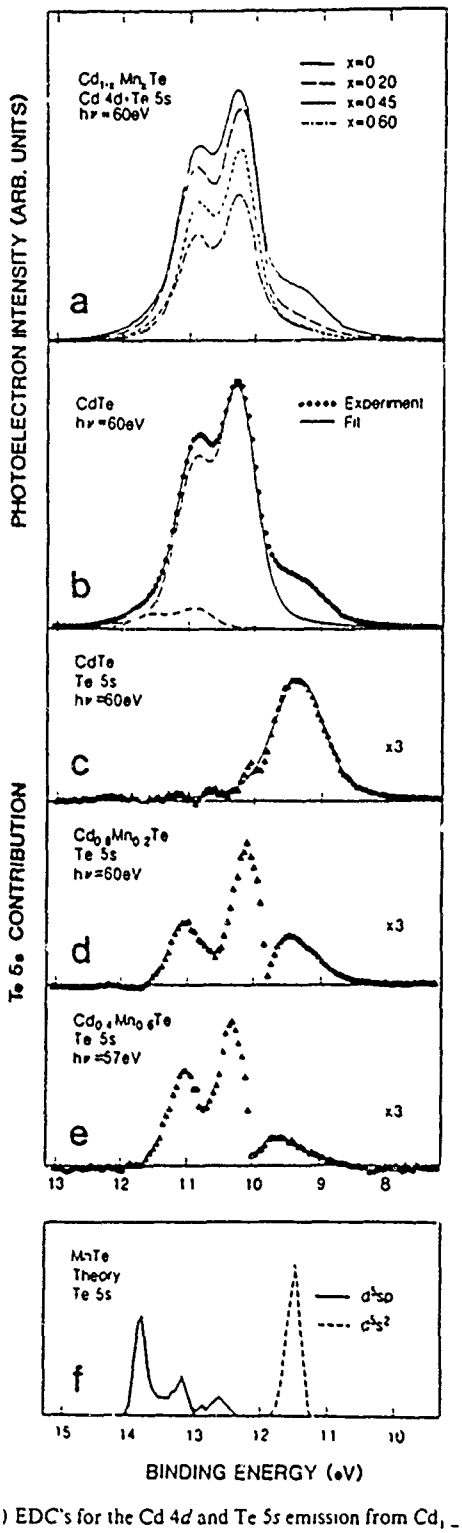


FIG 2 (a) EDC's for the Cd 4*d* and Te 5*s* emission from $Cd_{1-x}Mn_xTe$ ($x = 0$, 0.20, 0.45, and 0.60) at $h\nu = 60$ eV for all samples except $Cd_{0.40}Mn_{0.60}Te$ ($h\nu = 57$ eV). (b) Least-square fit (—) of the experimental EDC for CdTe (●) in the 7–13 eV range. Two Cd 4*d* doublet shifted 0.65 eV from each other reflect surface and bulk 4*d* components. (c) Te 5*s* contribution to the EDC for CdTe in Figs. 2(a)–2(b) after subtraction of the Cd 4*d* contribution from the overall spectrum (▲). For comparison we show the result of a least-squares fit in terms of a Lorentzian function convoluted with a Gaussian (—). (d) Te 5*s* contribution to the EDC for $Cd_{0.8}Mn_{0.20}Te$ in Fig. 2(a), after subtraction of the Cd 4*d* contribution from the overall spectrum. (e) Te 5*s* contribution to the EDC for $Cd_{0.4}Mn_{0.6}Te$ in Fig. 2(a), after subtraction of the Cd 4*d* contribution from the overall spectrum. (f) Calculated Te 5*s* line shape for the $(d\uparrow)^5(s\uparrow)(p\uparrow)$ (—) and $(d\uparrow)^5s^2$ (- - -) configuration of Mn in antiferromagnetic zinc-blende MnTe.

super-exchange on the Te 5*s* states. Comparison of our calculations with and without full spin polarization, shows major differences in the predicted Te 5*s* line shape. The results are shown in the bottom-most section of Fig. 2. For the reduced spin polarization in the $(d\uparrow)^5s^2$ case (dashed line), theory predicts a single Te 5*s* feature located some 11.5 eV below E_v, with a bandwidth of ~0.5 eV. When the *sp* polarization is correctly taken into account (solid line) theory predicts a 5*s* band some 1.5 eV wide, shifted to higher binding energy and exhibiting structure at 12.7, 13.2, and 13.6. The polarization of the Mn *sp* electrons in our results yields a larger super-exchange interaction and a corresponding increase in dispersion and binding energy of the Te 5*s* states. This effect contributes substantially to stabilize the Mn $(d\uparrow)^5(s\uparrow)(p\uparrow)$ ground-state configuration.[25] As was the case for the DOS, the results of earlier calculations[6,7,23,24] are similar to those we obtained for a $(d\uparrow)^5s^2$-like configuration for Mn in MnTe.

The Te 5*s* line-shape characteristic of the spin-polarized $(d\uparrow)^5(s\uparrow)(p\uparrow)$ ground-state configuration appears to be the one we observe, although comparison with experiment is complicated by the partial superposition of Te 5*s* and Cd 4*d* states in the photoemission spectra. In Figs. 2(a) and 2(b) we show photoemission spectra in the 9–13 eV binding energy region for $Cd_{1-x}Mn_xTe$ alloys with $x = 0$, 0.20, 0.45, and 0.60 and $h\nu = 60$ eV (57 eV for $Cd_{0.40}Mn_{0.60}Te$). The spectra have been normalized to the integrated intensity of the Cd 4*d* cores scaled by the Cd concentration $(1 - x)$ as determined from x-ray microprobe analysis. The experimental EDC's in Fig. 2(a) show a broad low-binding-energy shoulder at ~9.35 eV that reflects the Te 5*s* emission and exhibits a complex dependence on the Mn content x of the alloy. In Fig. 2(b) we show a deconvolution of the Cd 4*d* contribution in CdTe. A least-squares fit (solid line) to the experimental EDC (solid circles) was obtained in terms of two 4*d* doublets (dashed lines) shifted 0.65 relative to each other. The high-binding-energy doublet was recently observed by John *et al.*[29] and associated with a surface shift of the 4*d* line. John *et al.*[29] associated the Te 5*s* feature with a second Cd 4*d* surface component shifted in the opposite direction. Although this suggestion and the arguments provided to support it are quite ingenious, the Te 5*s* feature exhibits a cross section markedly different from that of the two Cd 4*d* doublets, a lack of escape depth sensitivity that rules out a surface-related character, and an x dependence in the ternary alloy that rules out a Cd-related origin.[30]

In the fitting procedure of Fig. 2(b) each 4*d* doublet was obtained through convolution of Lorentzian line shapes with a Gaussian function to account for the experimental resolution. Binding energy, width, intensity, branching ratio, and spin–orbit splitting of the doublets were all determined by the least-squares fitting program, and are in good agreement, within experimental uncertainty,[31] with those derived by John *et al.*[29] The two 4*d* doublets obtained were then subtracted from the overall EDC to isolate the Te 5*s* contribution, shown in Fig. 2(c) (solid triangles). The experimental Te 5*s* contribution in the binary compound is compellingly similar to the one predicted by theory for the low spin polarization, low super-exchange $(d\uparrow)^5s^2$ case, showing a *single*

emission feature centered some 9.35 eV below E_F (theory: 11.5 eV). If we now use the fitting parameters obtained for CdTe and, neglecting alloy broadening, scale down the Cd contribution by $(1-x)$, we can subtract[32] the Cd contribution from the ternary alloy EDC's of Fig. 2(a). We show the result for $Cd_{0.80}Mn_{0.20}Te$ and $Cd_{0.40}Mn_{0.60}Te$ in Figs. 2(d) and 2(e), respectively. The Te 5*s* line in the ternary alloy exhibits a shift to high binding energy relative to the binary, an increase in bandwidth and the emergence of three major emission features. For $Cd_{0.40}Mn_{0.60}Te$ structure is observed at 9.6, 10.3 and 11 eV, with a total 5*s* bandwidth of ~2.5 eV. This is in good agreement with the general trend expected from our calculations. The calculations, however, appear to overestimate the actual binding energy of the major Te 5*s* band features.

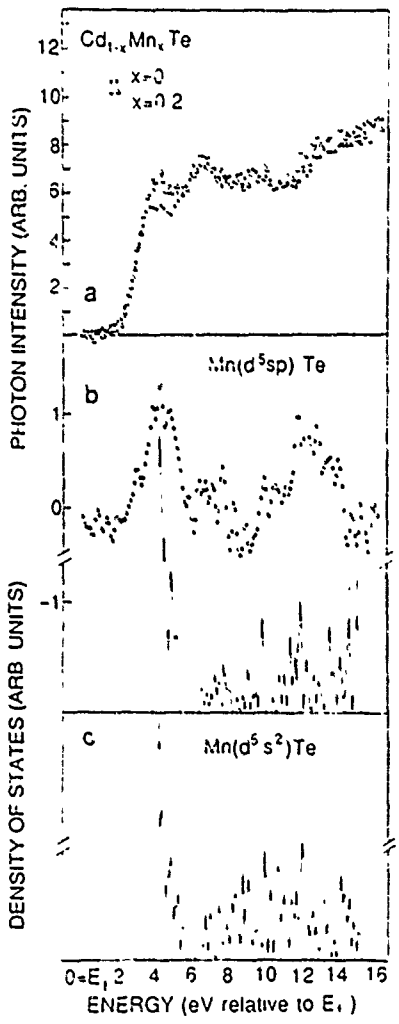


FIG 3 (a) Bremsstrahlung isochromat (BIS) spectra from CdTe (●) and $Cd_{0.80}Mn_{0.20}Te$ (△). The conduction-band minimum E_c is 0.9 eV above the Fermi level E_F for CdTe, 1.2 eV for $Cd_{0.8}Mn_{0.2}Te$. (b) A difference curve (□) derived from the BIS spectra of Fig. 3(a) is compared with the result of local density functional calculations of the total density of states above E_F (—) for antiferromagnetic zinc-blende MnTe. The theoretical spectrum has been rigidly shifted to align the calculated features (vertical bars) to the experimental ones. (c) Theoretical total density of states for antiferromagnetic zinc blende MnTe where Mn is in the artificial low-spin $(d^5)^5s^2$ configuration

C. Conduction-band states

The nature of the 3*d* contribution to the density of states above E_F and the value of the Coulomb correlation energy in ternary semimagnetic semiconductors has been quite controversial.[1] In Fig. 3(a) we show BIS spectra for CdTe and $Cd_{0.80}Mn_{0.20}Te$ in the 0–16 eV energy range above the Fermi level E_F. The position of E_F at the surface was 0.62 ± 0.1 eV and 0.75 ± 0.1 eV above the valence-band maximum for $Cd_{0.80}Mn_{0.20}Te$ and CdTe, respectively. Each spectrum in Fig. 3(a) was obtained as the sum of several quantitatively consistent spectra from different cleaves, and corresponds to a total of some 150 h of data integration with a primary electron beam current of about ~200 μA. Electron-beam induced contamination was avoided by obtaining a new cleave every 6–8 h. Although special care was taken to obtain reliable contacts on each single crystal with indium and conductive epoxy, charging effects were observed for all samples with $x \gtrsim 0.45$. We were able to obtain reliable data for samples with $x = 0$, 0.20, and 0.35.

In Fig. 3(a) the BIS spectrum for CdTe (solid circles) shows structure at 4.0, 6.7, and 9 eV above E_F, superimposed on a smoothly increasing secondary photon background deriving from the decay of inelastically scattered electrons. The spectrum for $Cd_{0.80}Mn_{0.20}Te$ (open triangle) was normalized for comparison to the 6.7-eV DOS feature of CdTe.[1,9] A difference curve was obtained from the two spectra and is shown in an expanded scale (open squares) in Fig. 3(b). Results for the $x = 0.35$ alloy are qualitatively consistent with those for the $x = 0.20$ alloys and are not presented here.[27] The conduction-band minimum E_c, expected at 0.90 eV for CdTe and at 1.2 eV for $Cd_{0.80}Mn_{0.20}Te$,[9] is seen in Fig. 3(a) as the end point of a tailing of states extending from 2.3 to ~1.0 eV.

New Mn–Te related features in the ternary alloys are observed in Fig. 3(b) at 4.2, 10.2, 12.2, and 13.8 eV (open squares). Analogous features are observed for the $x = 0.35$ alloy.[33] An interpretation of these features is provided by the results of our pseudofunction local density calculations. We show in Fig. 3(c) (solid line) the total DOS above E_F for the antiferromagnetic zinc-blende ground state of MnTe. For comparison we also give in Fig. 3(c) the analogous result for the reduced spin-polarization $Mn(d^5)^5s^2$ case. The calculated total DOS has been rigidly shifted to higher binding energy to align the main spectral features [vertical bars in Fig. 3(b)] to those observed experimentally, in accord with the usual practice of local-density theory.[34] We observe a compelling correspondence between theoretical and experimental features. *l*-projection of the theoretical DOS allows us to identify the main 4.2-eV feature as Mn 3*d*-derived, and the emission features at 10.2, 12.2, and 13.8 eV as primarily associated with Mn *p* states with some admixture of Te *p* character. We found no evidence of *d*-related emission in the gap. Upon deposition of 3 Å of unreacted Mn on the cleaved surface, we clearly detected dominant 3*d* emission in the gap and the formation of a Fermi cutoff, demonstrating that our experimental sensitivity would have been sufficient to detect Mn 3*d* states in the gap, if present.[27,33]

We can now estimate the d^5–d^4 exchange splitting from the photoemission- and BIS-determined position of the main

Mn 3*d* features and the position of the Fermi level. We obtain an exchange splitting of $3.5 - 0.7 - 4.2 = 8.4 \pm 0.4$ eV. Recent first-principle calculations for $Cd_{0.50}Mn_{0.50}Te$ alloys[7] predicted a Mn 3*d* ↑ band only some 2.5 eV below E_v, and a sharp *d*↓ feature within 2 eV from E_v, with a resulting *d*↑–*d*↓ splitting of 4.5 eV. Localized muffin-tin orbital calculations for antiferromagnetic zinc-blende MnTe by Podgorny[23] yielded a value of 4 to 4.5 eV for the *d*↑–*d*↓ splitting. Earlier calculations for paramagnetic zinc-blende MnTe by Masek *et al.*[24] yielded a slightly larger value (5.5 eV). All of the calculated values are substantially lower than our experimental value of 8.4 ± 0.4 eV or our calculated value of 6.8 eV, which is 24% too low. In the case of Mn atoms in AgMn intermetallics,[35] the experimental splitting between majority- and minority-spin states was also found larger (50%) than the value obtained from first-principle calculations (3.5 eV).

Better agreement is encountered with the results of earlier semiempirical calculations. Tight-binding semiempirical coherent potential approximation (CPA) calculations for $Cd_{1-x}Mn_xTe$ ($x = 0.3$ and 0.7) by Eherenreich *et al.*[36] fixed the position of the *d*↓ states 3.5 eV above E_v order to account for structure in the optical spectrum of the alloy relative to CdTe.[37] This, together with the photoemission-determined *d*↑ position 3.5 eV below E_v, yields a value of 7 eV for the exchange splitting, similar to the result of 6.8 eV that we obtain from our one-electron calculations after renormalization of the *s*–*p* band gap to the experimental transport value of 3.2 eV. The semiempirical cluster model by Fujimori[4] fixed the value of the Coulomb correlation energy to ~8 eV in order to fit the experimental photoemission data [see Fig. 1(f)]. Such a value is in good agreement with our BIS results, and is of the same order of magnitude as those assumed for Mott insulators such as MnO (9 eV)[6] for which local density functional calculations also systematically underestimate band gap and *d*↑–*d*↓ splitting.

A comparison of our results with optical data should identify transitions from *sp* valence-band states to the unoccupied *d* states, that our BIS study puts 4.8 ± 0.3 eV above the valence-band maximum E_v. In fact, structure at 4.5 eV has been observed in ellipsometry studies of $Cd_{1-x}Mn_xTe$,[37] and Kendlewicz[38] has recently interpreted structure at 4.5 eV in reflectivity measurements as deriving from $sp-d$↓ transitions. Our work clearly supports this interpretation of the optical spectra, and the contention of Larson *et al.*[6] that the optical structure at lower binding energy (2.2 eV) should be explained in terms of standard interband transition in zinc-blende semiconductors or Mn^{+2} excitonlike excitations that do not appear in band-structure calculations, rather than in terms of $sp-d$ transitions or d–d excitations.[6]

IV. CONCLUSIONS

Our photoemission and inverse photoemission measurements of $Cd_{1-x}Mn_xTe$ alloys, together with new local density functional calculations of the ground state of MnTe rule out the large final-state effects once thought necessary to explain the discrepancies between theory and photoemission experiments. The BIS technique provided the first direct experimental determination of the position of the *d*↓ states in a ternary semimagnetic semiconductor. The value of 4.8 ± 0.3 eV for the position of the *d*↓ band is consistent with early interpretations of the optical data which recently came under criticism on the basis of presumed discrepancies between photoemission and optical results, and photoemission and local density functional results. We find no evidence of such discrepancies in our results. From photoemission and BIS data we determined a value of 8.4 ± 0.4 eV for the *d*↑–*d*↓ exchange splitting, of the same order of magnitude as those assumed for Mott insulators, and almost twice as large as those predicted from earlier local density calculations. Finally, theory and experiments indicate that Mn in $Cd_{1-x}Mn_xTe$ adopts a spin polarized ground state with a $(d\uparrow)^5(s\uparrow)(p\uparrow)$-like configuration in the solid, a result that has important implications in our modeling of the magneto-transport and magneto-optical properties of these materials.

ACKNOWLEDGMENTS

This work was supported in part by DARPA and by ONR. We thank J.K. Furdyna for providing the crystals used in this study. The assistance of the staff of the University of Wisconsin Synchrotron Radiation Center is gratefully acknowledged. We are in debt to A. Baldereschi, C. Balzarotti, O. Gunnarsson, F. Hass, L. Ley, A. Mycielski, B. Orlowski, A. Ramdas, S.H. Wei, and A. Zunger for fruitful discussions and for providing us with results of their work prior to publication.

a) Present address: IBM Corporation, Rochester, Minnesota

[1] See A. Franciosi, in *Diluted Magnetic (Semimagnetic) Semiconductors*, edited by R. L. Aggarwal, J. K. Furdyna, and S. von Molnar (Materials Research Society, Pittsburgh, PA, 1987), p. 175, and references therein.
[2] L. Ley, M. Taniguchi, J. Ghijsen, R. L. Johnson, and A. Fujimori, Phys. Rev. B 35, 2839 (1987).
[3] A. Balzarotti, N. Motta, A. Kisiel, M. Zimnal-Starnawska, M. T. Czyzyk, and M. Podgorny, Phys. Rev. B 31, 7526 (1985); P. Letardi, N. Motta, and A. Balzarotti, J. Phys. C 20, 2853 (1987).
[4] M. Taniguchi, A. Fujimori, M. Fujisawa, T. Mori, J. Souma, and Y. Oka, Solid State Commun. 62, 431 (1987).
[5] Y. R. Lee, A. K. Ramdas, and R. L. Aggarwal, Phys. Rev. B 33, 7383 (1986).
[6] B. E. Larson, K. C. Hass, H. Ehrenreich, and A. E. Carlsson, Phys. Rev. B (in press); H. Ehrenreich, K. C. Hass, B. E. Larson, and N. F. Johnson, in *Diluted Magnetic (Semimagnetic) Semiconductors*, edited by R. L. Aggarwal, J. K. Furdyna, and S. von Molnar (Material Research Society, Pittsburgh, PA, 1987), p. 157.
[7] S.-H. Wei and A. Zunger, Phys. Rev. B 35, 2340 (1987), and Phys. Rev. Lett. 56, 2391 (1986).
[8] A. Wall, S. Chang, P. Philip, C. Caprile, A. Franciosi *et al.*, J. Vac. Sci. Technol. A 5, 2051 (1987).
[9] J. K. Furdyna, J. Appl. Phys. 53, 7637 (1982).
[10] N. B. Brandt and V. V. Moshchalkov, Adv. Phys. 33, 193 (1984).
[11] R. R. Galazka and J. Kossut, *Narrow Gap Semiconductors Physics and Applications*, in Lecture Notes in Physics Series (Springer, Berlin, 1980), Vol. 133.
[12] L. A. Kolodziejski, R. L. Gunshor, N. Otsuka, X. C. Zhang, S. K. Chang, and A. V. Nurmikko, Appl. Phys. Lett. 47, 882 (1985).
[13] C. Mailhiot and D. L. Smith, Phys. Rev. B 33, 2566 (1986), G. Y. Wu, D. L. Smith, C. Mailhiot, and T. C. McGill, Appl. Phys. Lett. (in press).
[14] P. W. Anderson, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), Vol. 14, p. 99.
[15] O. Gunnarsson, O. K. Andersen, O. Jepsen, and J. Zaanen (unpublished).

[16]A. Franciosi, A. Wall, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, R. V. Kasowski, R. Reifenberger, and F. Pool (unpublished).
[17]Y. Gao, M. Grioni, B. Smandek, J. H. Weaver, and T. Tyrie, J. Phys. E 21, 489 (1988).
[18]J. K. Land and Y. Baer, Rev. Sci. Instrum. 50, 221 (1979).
[19]F. J. Himpsel and T. Fauster, J. Vac. Sci. Technol. A 2, 815 (1984); P. D. Johnson and N. V. Smith, Phys. Rev. B 27, 2527 (1983).
[20]A. Franciosi, C. Caprile, and R. Reifenberger, Phys. Rev. B 31, 8061 (1985); A. Franciosi, R. Reifenberger, and J. K. Furdyna, J. Vac. Sci. Technol. A 3, 124 (1985).
[21]M. Taniguchi, L. Ley, R. L. Johnson, J. Ghijsen, and M. Cardona, Phys. Rev. B 33, 1206 (1986).
[22]A. Franciosi, S. Chang, C. Caprile, R. Reifenberger, and U. Debska, J. Vac. Sci. Technol. A 3, 926 (1985); A. Wall, A. Franciosi, D. W. Niles, C. Quaresima, M. Capozi *et al.*, Phys. Rev. B (in press).
[23]M. Podgorny, Z. Phys. B 69, 501 (1988).
[24]J. Masek, B. Velicky, and V. Janis, Acta Phys. Pol. A 69, 1107 (1986).
[25]A more complete analysis will be presented in M.-H. Tsai and J. D. Dow (unpublished); see also R. V. Kasowski, M.-H. Tsai, T. N. Rhodin, and D. D. Chambliss, Phys. Rev. B 34, 2656 (1986).
[26]M.-H. Tsai, J. D. Dow, R. V. Kasowski, A. Wall, and A. Franciosi, Solid State Commun. (in press).
[27]A. Wall, A. Raisanen, G. Haugstad, and A. Franciosi (unpublished).
[28]D. E. Eastman and J. L. Freeouf, Phys. Rev. Lett. 34, 395 (1975); G. K. Wertheim and S. Hüfner, *ibid.* 28, 1028 (1972).
[29]P. John, T. Miller, T. C. Hsieh, A. P. Shapiro, A. L. Wachs, and T.-C. Chiang, Phys. Rev. B 34, 6704 (1986).
[30]A. Wall, A. Franciosi, Y. Gao, J. H. Weaver, and J. R. Chelikowsky (unpublished).
[31]A best fit in the 8–13 eV range was obtained for the following values of the fitting parameters (we give in parenthesis for comparison the values derived by John *et al.*[29]): Lorentzian full width at half-maximum (FWHM) = 0.25 eV (0.21 eV), Gaussian FWHM = 0.473 eV (0.44 eV), spin-orbit splitting 0.66 eV (0.69 eV); branching ratio = 0.666 (0.658); surface/bulk ratio = 0.13 (0.63, for different surface orientation after ion bombardment); surface shift = 0.66 eV (0.57 eV).
[32]A scale factor to normalize EDC's from one cleave relative to another cleave can be obtained from the Mn 3*d* emission intensity scaled by the concentration *x*, from the Te 4*d* core emission, or assuming that the total integrated Te 5*s* emission remains unchanged in the alloy series. All of these methods yield results qualitatively consistent for the Te 5*s* line shape in Figs. 2(c), 2(d), and 2(e).
[33]These results are in agreement with those of BIS investigation at ultraviolet energies of $Cd_{1-x}Mn_xTe$ alloys and Mn–CdTe(110) interface reaction. See A. Wall, A. Franciosi, and B. Reihl (unpublished).
[34]However, the rigid shift used in Fig. 3 is some 1.6 eV higher than what would be necessary to align the calculated *s*–*p* band gap to the value of 3.2 eV expected for the gap of MnTe (Ref. 7).
[35]D. van der Marel, G. A. Sawatzky, and F. U. Hillebrecht, Phys. Rev. Lett. 53, 206 (1984).
[36]H. Ehrenreich, K. C. Hass, N. F. Johnson, B. E. Larson, and R. J. Lampert, in *Proceedings of the 18th International Conference on the Physics of Semiconductors*, edited by O. Engstrom (World Scientific, Singapore, 1987), p. 1751.
[37]P. Lautenschlager, S. Logothetidis, L. Vina, and M. Cardona, Phys. Rev. B 32, 3811 (1985).
[38]T. Kendlewicz, J. Phys. C 14, L407 (1981).

Strain-assisted *p*-type doping of II-VI semiconductors

Shang Yuan Ren
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 and Department of Physics, University of Science and Technology of China, Hefei, Anhui, People's Republic of China

John D. Dow and Stefan Klemm
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 30 January 1989; accepted for publication 24 April 1989)

By incorporating a II-VI semiconductor into a strained-layer superlattice, it should be possible to overcome the effects of deep hole traps near the valence-band edge and hence to dope the semiconductor *p* type in many cases. This idea is illustrated for CdTe/ZnTe superlattices.

I. INTRODUCTION

For many years, efforts to fabricate efficient light emitters, lasers, and other optical devices from large-band-gap (green to ultraviolet) II-VI semiconductors have been frustrated by the resistance of these materials to doping, especially *p*-type doping. There appear to be four main explanations that have been given for this problem: (i) Self-compensation: common dopants, such as column-I impurities, are believed to simultaneously produce anion vacancies when they occupy column-II sites, and the vacancies compensate the dopants[1], (ii) large acceptor binding energies: because of the large effective masses (typically $m^* \simeq 1.35 m_0$ for CdS) and small dielectric constants ($\kappa \simeq 8.9$), the acceptor binding energies, $E_B = 13.6\,\mathrm{eV}\,(m^*/m_0\kappa^2)$ are often 60-200 meV vs 10–60 meV for the more common III-V and group-IV semiconductor materials, and hence the acceptors are less easily thermally ionized[1,2]; (iii) low incorporation probabilities: for reasons that are not presently understood, *p*-type dopants may not incorporate efficiently—recent experiments[3] suggest that this incorporation can be dramatically increased by photoassisted doping; and (iv) deep-level formation: impurities, even the dopants themselves in some cases, form deep hole traps and impurity energy levels in the gap that capture any free holes.

In this paper we focus on the issue of deep-level formation and prescribe a method for overcoming the effects of deep hole traps that are moderately close to the valence-band maximum.

Theories of column-V impurities occupying column-VI sites in II-VI semiconductors indicated that, for the wurtzite crystal structure, these expected shallow dopants often produce deep traps instead,[4] but in the zinc-blende structure they produce shallow acceptors. The observation of a nitrogen level $\simeq 100$ meV above the valence-band maximum of ZnSe lent support to the theory.[5] Hence small changes in the local environment of a substitutional impurity are perhaps sufficient to change the character of a potential *p* dopant from a shallow acceptor which provides hole carriers and enhances *p*-type conductivity to a deep impurity which traps holes and tends to make the material semi-insulating rather than *p* type. This sensitivity of impurity character to environment suggests that a range of perturbations of the host semiconductor may convert a deep hole trap into a *p*-type dopant. In particular, this paper discusses the possibility that impurities that might otherwise produce deep hole traps $\lesssim 0.25$ eV above the valence-band maximum of their host semiconductor might be converted into *p*-type dopants by subjecting the host semiconductor to a strain obtained by incorporating the host into a lattice-mismatched superlattice. For example, the CdTe/ZnTe superlattice has a lattice mismatch of 6.3% (Refs. 6 and 7) and the strain in the superlattice can cause deep hole traps in the unstrained smaller-gap material, CdTe, to become shallow acceptors in the superlattice.

In this regard we note that acceptor levels in CdTe have been reported at 58, 59, and 60 meV above the valence-band maximum for Li, Na, and P, as well as at 147 and 108 meV for Cu and Ag.[8] We believe that the Li, Na, and P levels are genuine substitutional shallow acceptor levels, because their energies are almost equal, the small differences being attributed to central-cell corrections. We identify the Cu and Ag levels as deep levels (with qualitatively different wave functions from shallow levels[9]) that lie relatively close to the valence band and act as hole traps. Because the strength of the ordinary central-cell potentials for valence electrons of Cu and Ag are intermediate between those of Li and Na on the one hand and P on the other,[10] we believe that the Cu and Ag energies are not shallow effective-mass theory energies. Therefore, a goal of the present theory, as applied to CdTe, is to predict the layer thicknesses of CdTe/ZnTe superlattices in which the Li, Na, and P levels remain shallow (in a CdTe layer) but Cu and Ag become shallow acceptors rather than deep traps. We shall see that this should occur, due to internal strain, in small-period $N_{\mathrm{CdTe}} \times N_{\mathrm{ZnTe}}$ superlattices for $N_{\mathrm{CdTe}}/N_{\mathrm{ZnTe}} < 1$ in the case of [001] superlattices.

II. QUALITATIVE PHYSICS

The application of hydrostatic pressure to a semiconductor shifts the band edges with respect to nearby deep levels.[11] Likewise a uniaxial stress splits the valence-band edge and shifts the conduction-band edge, relative to the energies of deep levels.[12] Furthermore, deep levels in superlattices, crudely speaking, have almost the same absolute energies as in the bulk—but the band edges of the superlattice are significantly perturbed from the bulk value.[13–16] Thus, in a crude approximation, we may regard the deep energy levels associated with impurities as almost constant in energy, while hydrostatic pressure, uniaxial stress, or quantum confinement in a superlattice cause the semiconductor's band edges to move considerably. This viewpoint is

 0021-8979/89/172065-04$02.40 © 1989 American Institute of Physics

illustrated in Fig. 1, where we show the fundamental band edges of CdTe with a deep trap level, and illustrate how the band edges move when perturbed (assuming the absolute deep-level energy does not change). An important point is that when the valence-band maximum moves up enough in energy to cover the deep trap level, the impurity changes from a hole trap to a shallow acceptor, because the hole is autoionized, and bubbles up to the valence-band maximum where it is trapped by the long-ranged Coulomb potential of the impurity. (See Fig. 1.) Thus by "covering up" the deep trap levels in the gap with the valence band, it is possible to convert the traps into shallow acceptors and achieve *p*-type doping. This "covering up" is termed a deep-to-shallow transition.

Hydrostatic pressure causes the valence-band maximum to move up slightly in energy while the conduction-band minimum moves up even more.[11] The motion of the valence-band maximum relative to the deep level is typically of magnitude ≈ 1 meV/kbar,[11] and so hydrostatic pressures can only "cover up" deep levels $\gtrsim 0.1$ eV above the valence-band maximum, even if the pressures p are very large (almost impractically large): $p \gtrsim 100$ kbar.

Uniaxial stress can be more effective in inducing a deep-to-shallow transition that converts a deep trap to a shallow acceptor, but externally applied uniaxial stresses much greater than 10 kbar ordinarily can fracture a semiconductor. Under uniaxial stress, even a 10-kbar stress can cover up levels within ≈ 0.14 eV of the valence-band maximum.

Internal strains in a lattice-mismatched superlattice such as CdTe/ZnTe can move the valence-band edge tenths of an eV. We shall show that small-period superlattices can be grown which have such large internal strains that deep trap levels within about 0.25 eV of the valence-band maximum can be covered up, making them shallow acceptors. This should be an important mechanism for producing *p*-type doping, because in CdTe there are many deep hole traps with energies around 0.1–0.2 eV above the valence-band maximum, including traps associated with Cu and Ag.[8]

For [001] superlattices, every internal stress σ can be represented as a combination of hydrostatic and uniaxial stresses:

$$\sigma_{\text{hydro}} = (\sigma_{xx} + \sigma_{yy} + \sigma_{zz})/3 \tag{1}$$

and

$$\sigma_{\text{uni}} = \sigma_{zz} - \sigma_{xx}. \tag{2}$$

In a typical 1×1 CdTe/ZnTe superlattice the CdTe internal strain corresponds to hydrostatic pressures of ≈ -10.7 kbar and a uniaxial stress of ≈ 16.0 kbar. Hence the heteroepitaxy of the superlattice allows one to reach local pressures and stresses (without damaging the semiconductor) that would be unattainable when applied globally to bulk semiconductors. These enormous strains can move the valence-band maximum up in energy by ≈ 0.25 eV, covering up any deep traps near the valence-band maximum of unstrained CdTe.

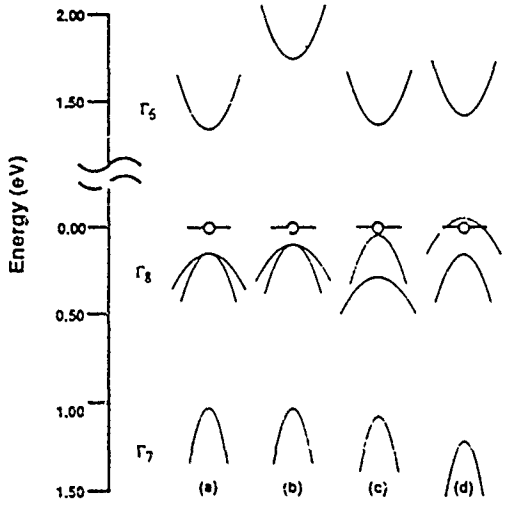


FIG. 1. Schematic energy-band structure (energy in eV vs wave vector) of CdTe, illustrating the qualitative changes in the valence-band-level structure with respect to the deep-level energy. (a) The bulk semiconductor with a deep level within 0.2 eV of the valence-band edge. The Γ_8 and Γ_7 bands were *p*-like bands that split due to the large spin-orbit interaction in CdTe. (b) The application of 50 kbar hydrostatic pressure decreases the energy between the impurity level and the valence band slightly. (c) 10 kbar of uniaxial stress splits the Γ_8 bands and can "cover up" the deep hole trap and make it an acceptor. (d) A 2×4 superlattice is even better than large uniaxial stress in that the deep level is better covered by the internal strain-induced band splitting.

To simplify the discussion, we shall make the assumption that the relevant deep levels have absolute energies that are unchanged by internal strains. With this approximation, we need only predict the shift and splitting of the valence-band maximum caused by strain, and determine if deep hole traps 0.1–0.2 eV above the valence-band maximum of unstrained CdTe will be covered up by the strain-induced shift of the valence-band maximum.

III. FORMALISM

In zinc-blende semiconductors the top of the valence band is threefold degenerate and *p* like in a nonrelativistic band theory. Inclusion of the spin-orbit interaction splits this degeneracy into a Γ_8 doubly degenerate $J = 3/2$ valence-band maximum and a nondegenerate $J = 1/2$ Γ_7 band below it. Strain further splits the bands, leading to a Hamiltonian matrix in the $|J, M_J>$ pseudo-angular momentum basis for the valence-band maxima at $k = 0$.

$$\begin{array}{c} \begin{matrix} |3/2,\ 3/2> & |3/2,\ 1/2> & |1/2,\ 1/2> \end{matrix} \\ \begin{pmatrix} -\delta E_H - (\delta E_S)/2 & 0 & 0 \\ 0 & -\delta E_H + (\delta E_S/2) & (1/\sqrt{2})\delta E_S \\ 0 & (1/\sqrt{2})\delta E_S & -\Delta_0 - \delta E_H \end{pmatrix} \end{array}. \tag{3}$$

Here we have used the notation of Ref. 17: Δ_0 is the spin-orbit interaction parameter, δE_H and $\delta E_{H'}$ are the energy shifts of the relevant bands induced by the isotropic strain or dilation:

$$\Delta = (u_{xx} + u_{yy} + u_{zz})/3, \tag{4}$$

and δE_S and $\delta E_{S'}$ are the shifts due to the anisotropic part of the strain, and the zero of energy is the unstrained valence-band maximum. Because the energy shifts (with respect to a nearby deep level) induced by the isotropic parts of the strain[11] are about an order of magnitude smaller than the anisotropic-strain shifts, the upward valence-band shift induced by internal strain is very nearly

$$\Delta E_{vb} \approx -\delta E_S/2, \tag{5}$$

and is due to the anisotropic strain.

IV. [001] SUPERLATTICES

For [001] superlattices, the valence-band shift is related to the strain tensor components $u_{zz} - u_{xx}$ by

$$\Delta E_{vb} \approx -\delta E_S/2 = (b_1 + 2b_2)(u_{zz} - u_{xx}), \tag{6}$$

where b_1 and b_2 are deformation potential constants.[17] Since we have been unable to find tabulated values of b_1 and b_2 for CdTe, we estimate $(b_1 + 2b_2) = -1.76$ eV, a value typical of semiconductors (for example, in GaAs the measured values range from -1.66 to -2.0 eV, in Ge a typical value is -2.1 eV). The strains u_{zz} and u_{xx} in CdTe can be related to the bulk lattice constants a_{CdTe} and a_{ZnTe} and to the layer thicknesses of the superlattice: N_{CdTe} and N_{ZnTe}, as discussed in the Appendix. Figure 2 shows the theoretical predictions.

V. [111] SUPERLATTICES

In [111] superlattices, the valence-band shift is most easily related to the shear ϵ in a CdTe layer:

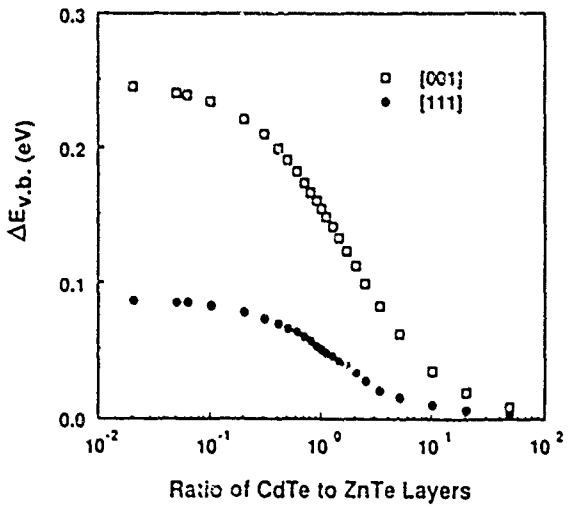


FIG. 2. Energy shift ΔE_{vb} in eV vs N_{CdTe}/N_{ZnTe}, the ratio of CdTe to ZnTe layer thicknesses. Deformation potential theory predicts that decreasing the layer thickness ratio N_{CdTe}/N_{ZnTe} for CdTe/ZnTe superlattices increases strain, thereby shifting the valence-band maximum up relative to the deep level by up to 0.25 eV for [001] superlattices (open squares). This effect is relatively weak in [111] superlattices (solid circles).

$$\epsilon = u_{xx} - u_{yy} = u_{zx}, \tag{7}$$

and is

$$\Delta E_{vb} \approx -\delta E_S/2 = -(d_1 + 2d_2)\epsilon/2. \tag{8}$$

(See the Appendix.) Taking the deformation potential constant for CdTe to be $(d_1 + 2d_2) \approx -4.59$ eV, the value for GaAs,[17] we find the results of Fig. 2.

VI. SUMMARY

Clearly internal strains can raise the valence-band edge of CdTe (and other II-VI semiconductors) a few tenths of an eV when the semiconductor is sandwiched in a strained-layer superlattice or heterostructure. This should be enough to cover up the deep levels at energies ≈ 0.1–0.2 eV above the unstrained valence-band maximum and convert these deep hole traps to p-type shallow dopants. (Of course, the precise extent of this effect should be evaluated by computing the deep trap levels in a strained-layer superlattice,[13–16] rather than assuming that the deep levels are completely unaltered by the strain, as we have here.)

By sandwiching a II-VI semiconductor between materials which have a different lattice constant, as in a strained-layer superlattice, it will be possible to move the semiconductor's valence-band maximum up in energy, so that deep hole traps lying within 0.1–0.2 eV of the unstrained valence-band edge will be autoionized and will become shallow acceptors, doping the II-VI semiconductor p type. We believe that this mechanism of conditioning II-VI semiconductors for p type doping is most likely to be successful in common-anion heterostructures and superlattices, such as CdS/ZnS, CdSe/ZnSe, or CdTe/ZnTe. The reason for choosing common-anion heterostructures is to minimize the absolute value of the valence-band offset. If the CdTe valence-band maximum lies *above* the ZnTe valence band, then quantum confinement will tend to drive the valence-band maximum down in energy (and inhibit covering of the deep levels); but if the CdTe valence-band maximum lies *below* the ZnTe maximum, holes will tend to be swept out of the CdTe layer. We hope that experimenters will attempt to grow such structures, and determine if they are indeed more easily doped p type, as we predict.

ACKNOWLEDGMENTS

We are grateful to the Air Force Office of Scientific Research, the Office of Naval Research, and the Defense Advanced Research Projects Agency for their generous support (Contract Nos. AF-AFOSR-85-0331, N00014-84-K-0352, and N0530-0716-05). We thank R.-D. Hong for stimulating conversations and R. N. Bhargava and J. Schetzina for posing problems which caused us to think along these lines.

APPENDIX

The relevant parts of the strain tensor u for evaluating ΔE_{vb} can be calculated using continuum elasticity theory.[18]

For a [001] superlattice, the relevant strain tensor elements are u_{xx} and u_{zz} [see Eq. (6)]. The boundary condi-

tions are that the in-plane lattice constants $a_{\parallel}$ of the adjacent layers match one another:

$$a_{\parallel \mathrm{CdTe}} = a_{\parallel \mathrm{ZnTe}},$$

and are related to the bulk lattice constant through the strain tensor u_{xx}:

$$a_{\parallel \mathrm{CdTe}} = a_{\mathrm{CdTe}}(1 + u_{xx,\mathrm{CdTe}}).$$

The constraints are

$$u_{xx} = u_{yy},$$

$$u_{xy} = u_{yz} = u_{zx} = 0,$$

and

$$a_{\mathrm{CdTe}}(1 + u_{xx;\mathrm{CdTe}}) = a_{\mathrm{ZnTe}}(1 + u_{xx,\mathrm{ZnTe}}).$$

We assume that the layer thicknesses are sufficiently small that we may take the strain tensors u_{CdTe} and u_{ZnTe} to be constants. The total elastic energy of the superlattice is

$$W = N_{\mathrm{CdTe}} W_{\mathrm{CdTe}} + N_{\mathrm{ZnTe}} W_{\mathrm{ZnTe}},$$

where

$$W_{\mathrm{CdTe}} = (1/2)c_{11;\mathrm{CdTe}}(2u^2_{xx,\mathrm{CdTe}} + u^2_{zz;\mathrm{CdTe}}) + c_{12,\mathrm{CdTe}}(u_{xx;\mathrm{CdTe}}u_{xx;\mathrm{CdTe}} + 2u_{xx,\mathrm{CdTe}}u_{zz;\mathrm{CdTe}})$$

is the increase of elastic energy in a CdTe layer (with a similar expression for ZnTe layer). The elastic constants c_{11}, c_{12}, and c_{44} have been tabulated[7] for both CdTe and ZnTe. The solutions for u_{zz} and u_{xx} in a CdTe layer are obtained by minimizing the superlattice's total energy for fixed N_{CdTe} and N_{ZnTe}:

$$u_{xx\,\mathrm{CdTe}} = D_{12}(a_{\mathrm{CdTe}} - a_{\mathrm{ZnTe}})/D,$$

and

$$u_{zz;\mathrm{CdTe}} = -2c_{12,\mathrm{CdTe}}u_{xx,\mathrm{CdTe}}/c_{11\,\mathrm{CdTe}}$$

where we have

$$D_{11} = N_{\mathrm{CdTe}}(2c_{11,\mathrm{CdTe}} + 2c_{12,\mathrm{CdTe}} - 4c^2_{12\,\mathrm{CdTe}}/c_{11,\mathrm{CdTe}})/a_{\mathrm{CdTe}},$$

$$D_{12} = N_{\mathrm{ZnTe}}(2c_{11,\mathrm{ZnTe}} + 2c_{12,\mathrm{ZnTe}} - 4c^2_{12,\mathrm{ZnTe}}/c_{11\,\mathrm{ZnTe}})/a_{\mathrm{ZnTe}},$$

and

$$D = -(a_{\mathrm{ZnTe}}D_{11} + a_{\mathrm{CdTe}}D_{12}).$$

For [111] superlattices, the relevant strain is the anisotropic shear strain ϵ. [See Eq. (8)]. The boundary conditions are that the in-plane lattice constants match one another:

$$a_{\mathrm{CdTe}} = a_{\mathrm{ZnTe}},$$

where we have

$$a_{\mathrm{CdTe}} = a_{\mathrm{CdTe}}(1 + \Delta_{\mathrm{CdTe}} - \epsilon_{\mathrm{CdTe}}),$$

with a similar expression for ZnTe. The constraints are

$$\Delta = u_{xx} = u_{yy} = u_{zz},$$

$$\epsilon = u_{xy} = u_{yz} = u_{zx},$$

and

$$a_{\mathrm{CdTe}}(1 + \Delta_{\mathrm{CdTe}} - \epsilon_{\mathrm{CdTe}}) = a_{\mathrm{ZnTe}}(1 + \Delta_{\mathrm{ZnTe}} - \epsilon_{\mathrm{ZnTe}}).$$

The increase in elastic energy of the superlattice is

$$W = N_{\mathrm{CdTe}}W_{\mathrm{CdTe}} + N_{\mathrm{ZnTe}}W_{\mathrm{ZnTe}},$$

where in each CdTe layer we have

$$W_{\mathrm{CdTe}} = (3/2)(c_{11,\mathrm{CdTe}} + 2c_{12,\mathrm{CdTe}})\Delta^2_{\mathrm{CdTe}} + 6c_{44,\mathrm{CdTe}}\epsilon^2_{\mathrm{CdTe}}.$$

By minimizing the total elastic energy of the superlattice for fixed N_{CdTe} and N_{ZnTe}, we find

$$\epsilon_{\mathrm{CdTe}} = -D_{12}(a_{\mathrm{CdTe}} - a_{\mathrm{ZnTe}})/D,$$

where

$$D_{11} = 12N_{\mathrm{CdTe}}c_{44;\mathrm{CdTe}}/a_{\mathrm{CdTe}},$$

$$D_{12} = 12N_{\mathrm{ZnTe}}c_{44,\mathrm{ZnTe}}/a_{\mathrm{ZnTe}},$$

$$D_{21} = (c_{11,\mathrm{CdTe}} + 2c_{12,\mathrm{CdTe}} + 4c_{44,\mathrm{CdTe}})a_{\mathrm{CdTe}}/(c_{11,\mathrm{CdTe}} + 2c_{12,\mathrm{CdTe}}),$$

$$D_{22} = -(c_{11,\mathrm{ZnTe}} + 2c_{12,\mathrm{ZnTe}} + 4c_{44,\mathrm{ZnTe}})a_{\mathrm{ZnTe}}/(c_{11,\mathrm{ZnTe}} + 2c_{12,\mathrm{ZnTe}}),$$

and

$$D = D_{11}D_{22} - D_{12}D_{21}.$$

[1] See, for example, C. H. Henry, K. Nassau, and J. W. Shiever, Phys. Rev. B 4, 2453 (1971), and references therein.

[2] D. W. Jenkins and J. D. Dow, Phys. Rev. B 39, 3317 (1989).

[3] R. N. Bicknell, N. C. Giles, and J. F. Schetzina, Appl. Phys. Lett. 49, 1735 (1986).

[4] A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 28, 946 (1983).

[5] Z. L. Wu, J. L. Merz, C. J. Werkhoven, B. J. Patrick, and R. N. Bhargava, Appl. Phys. Lett. 40, 345 (1982).

[6] The lattice constants of CdTe and ZnTe are 6.481 and 6.100 Å, respectively. See Ref. 7.

[7] Landolt-Bornstein, *Numerical Data and Functional Relationships in Science and Technology, New Series, Vol. 17b, Semiconductors: Physics of Group II-VI and I-VII Compounds, Semimagnetic Semiconductors*, edited by O. Madelung, M. Schulz, and H. Weiss (Springer, Berlin, 1982).

[8] R. N. Bhargava, J. Cryst. Growth 59, 15 (1982), and references therein.

[9] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980); for a discussion of the ideas on which this work was based, see W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[10] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[11] S. Y. Ren, J. D. Dow and D. J. Wolford, Phys. Rev. B 25, 7661 (1982). Recently R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow [Phys. Rev. B 38, 12549 (1988)] performed similar calculations for other semiconductors and found similar results.

[12] S. Y. Ren, D. Q. Mao, and M. F. Li, Chin. Phys. Lett. 3, 313 (1986). D. W. Jenkins, S. Y. Ren, and J. D. Dow, Phys. Rev. B 39, 7881 (1989).

[13] S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B 38, 9677 (1988).

[14] J. D. Dow, S. Y. Ren, and J. Shen, *Deep impurity levels in semiconductors, semiconductor alloys and superlattices*, in *NATO Advanced Sciences Institutes Series B, Vol. 183, Properties of Impurity States in Superlattice Semiconductors*, edited by C. Y. Fong, I. P. Batra, and S. Ciraci (Plenum, New York, 1988).

[15] R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Materials Res. Soc. Symp. Proc. 77, 545 (1987).

[16] S. Y. Ren and J. D. Dow, J. Appl. Phys. 65, 1987 (1989).

[17] M. Chandrasekhar and F. H. Pollak, Phys. Rev. B 15, 2127 (1977).

[18] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 3rd ed. (Pergamon, Oxford, 1986), p. 35.

Proposed explanation of the *p*-type doping proclivity of ZnTe

John D. Dow
Department of Physics and Astronomy, Arizona State University, Tempe, Arizona 85287-1504 and Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556*

Run-Di Hong, Stefan Klemm, Shang Yuan Ren, and M.-H. Tsai
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

Otto F. Sankey
Department of Physics, Arizona State University, Tempe, Arizona 85287-1504

R. V. Kasowski
Central Research and Development Department, E. I. du Pont de Nemours and Company, Experimental Station, Wilmington, Delaware 19898

(Received 7 March 1990; revised manuscript received 23 July 1990)

An explanation is proposed for the fact that ZnTe is unique among the II-VI compound semiconductors in that it can be doped *p* type rather easily: a *p*-like deep-level resonance lies within the valence band of ZnTe and emerges into the fundamental band gap with increasing Se content *x* in $ZnTe_{1-x}Se_x$ random alloys. This level generates free holes when it lies below the Fermi energy in the valence band, making its parent defect a shallow acceptor. When the level moves into the gap, the impurity becomes a deep hole trap. The native and foreign antisite defects Zn_{Te} and Li_{Te} are suggested as possible parent defects of the relevant deep level; they are predicted to be shallow acceptors in ZnTe, while the corresponding defects are deep traps in other II-VI compound semiconductors. Tests of this proposal are suggested and the substitutional *s*- and *p*-bonded deep levels of ZnSe and ZnTe are predicted, extending the theory of Hjalmarson *et al.* [Phys. Rev. Lett. 44, 810 (1980)]. The possibility of doping ZnSe *p* type with (antisite) Be is also proposed and discussed.

I. INTRODUCTION

Most wide-band-gap II-VI compound semiconductors can be easily doped *n* type, but resist *p*-type doping. The most notable exception is ZnTe which is easily doped *p* type, but not *n* type. The different and singular doping behavior of ZnTe is not presently understood.

The purpose of this paper is to offer an explanation of this difference in terms of a defect that undergoes a shallow-deep transition as a function of alloy composition *x* in $ZnTe_{1-x}Se_x$, so that it is an acceptor in ZnTe but a hole trap in ZnSe. We present calculations that suggest that such a transition occurs for the native and foreign antisite defects Zn_{Te} and Li_{Te} in ZnTe. And we suggest specific tests of the explanation, which has a firm, but admittedly not unshakable, theoretical foundation. Indeed, the reader should accept this work for its intended purpose: to sketch a simple model which can be tested experimentally and which appears to offer an explanation of the *p*-type doping proclivity of ZnTe in terms of the natural occurrence of certain defects. The theoretical tools are presently not available to predict with 100% confidence if these defects form in sufficient concentration with precisely the electronic structures we find. Nevertheless, the picture we propose is simple enough to lend itself readily to experimental tests, while presenting an attractive alternative explanation of the doping properties of ZnTe to mechanisms which rely on self-compensation or the relative solubilities of impurities in different hosts.

We propose that *p*-type conductivity is most easily achieved in ZnTe of the II-VI semiconductors because defects that produce *p*-like deep traps lying slightly above the valence-band maximum in most II-VI semiconductors instead have these levels lying below the band edge in ZnTe. Each such level for a neutral defect is naturally occupied by both electrons and holes when it lies within the fundamental band gap, and so can trap holes. When the level descends below the valence-band maximum, its holes are autoionized and become carriers at room temperature. (At zero temperature they are trapped in shallow acceptor levels.) Hence these defects generate holes in ZnTe, but trap them in other semiconductors. More specifically, our proposal is that the native antisite defect Zn on a Te site, Zn_{Te} (and even the antisite dopant Li_{Te}) is a shallow acceptor in ZnTe and dopes the material *p* type, compensating any *n*-type impurities that might otherwise themselves compensate acceptors. In other II-VI semiconductors, neither the cation-on-anion-site defect nor antisite Li yields a shallow acceptor, but instead produces a deep *p*-like level in the fundamental band gap. The antisite defects Zn_{Se} (and Li_{Se}) neither dope the material *p* type nor produce thermally ionizable holes; instead they trap carriers, holes in particular.

We arrived at the conclusion that the defect responsible for the resistance of most II-VI semiconductors to *p*-type doping might be the cation-on-anion-site defect after

 ©1991 The American Physical Society

having examined predictions for deep levels of all the *s*- and *p*-bonded substitutional impurities in all of the major II-VI semiconductors. During this examination we asked the questions (i) "Can impurities explain the different doping proclivities of ZnTe and the other II-VI semiconductors" and (ii) "Does any impurity that is likely to occur in significant concentrations have a deep-level structure that is different in ZnTe from in the other II-VI semiconductors?"

The classical mechanism that has been proposed to explain the resistance of II-VI semiconductors to *p*-type doping is self-compensation: The introduction of an acceptor such as Li or Na onto a cation site is thought to simultaneously produce an anion vacancy which is a double donor and compensates the acceptor, leaving the material semi-insulating or doping the material *n* type rather than *p* type.[1-6] If the self-compensation mechanism is indeed operative in most II-VI semiconductors and cannot be blocked for at least some *p*-type dopants, then II-VI semiconductors are unlikely to become important electronic materials, regardless of the purity of the materials. The self-compensation is thought to be an intrinsic property of the doping, regardless of the quality of the undoped semiconductor: *p*-type doping produces compensating anion vacancies. Therefore, the fact that ZnTe is naturally *p* type is an important clue to understanding and overcoming self-compensation (if it occurs)—and perhaps the key to developing electronic-grade II-VI semiconductors.

The explanation often given for the singular behavior of ZnTe is that its band gap is smaller than that of other II-VI semiconductors,[1-3] and so the energy generated by the self-compensation, which is approximately the energy of the gap, is too small to produce a vacancy in ZnTe, but is adequate for vacancy formation in other II-VI semiconductors. At first, this argument appears attractive, because it correlates with the fact that self-compensation is most commonly found in large-band-gap semiconductors. Upon closer examination, however, the argument is difficult to reconcile with the fact that the band gaps of ZnSe and ZnTe are not very different or with the fact that stoichiometric CdTe, which has a significantly smaller band gap than ZnTe, resists *p*-type doping (although less so than other II-VI compounds). Therefore we conclude that an alternative explanation of the singular *p*-type character of ZnTe is needed.

Self-compensation by vacancies is not universally accepted as the mechanism by which II-VI semiconductors resist *p*-type doping. Some other explanations point to the different solubilities of impurities on different sites and interstitial self-compensation: For example, Na can occupy a cation site as an acceptor in ZnSe, but becomes a donor at interstitial sites,[7] and so interstitial Na can compensate substitutional Na. Other mechanisms have also been suggested to account for the resistance of II-VI semiconductors to *p*-type doping: The acceptor binding energies of some II-VI semiconductors can be rather large, >0.1 eV, making the acceptors difficult to ionize thermally and frustrating *p*-type conductivity.[8] The incorporation of the standard *p*-type dopants into a II-VI host can be severely limited by the chemistry of the host.[9] Unanticipated deep-level formation may occur, in the sense that impurities such as those from column V of the Periodic Table, when occupying a column-VI site, may produce deep levels in the fundamental band gap of some wurtzite-structure II-VI compound semiconductors rather than the expected shallow acceptors.[10] Indeed, it is conceivable that the II-VI semiconductor doping problem is a consequence of rather complicated defect and impurity dynamics. Currently there are many efforts in progress to overcome the *p*-doping problem by improving the quality of II-VI materials, for example, by employing molecular-beam epitaxial techniques of crystal growth; these efforts implicitly assume that impurities and defects, not self-compensation, cause the *p*-type doping problem.[11-14]

Clearly there are a variety of detailed mechanisms for explaining the resistance of most wide-band-gap II-VI semiconductors to the *p*-type doping. However, a satisfactory explanation of why ZnTe is almost singularly compatible with *p*-type doping is lacking—and could provide a vital clue for doping the other II-VI semiconductors *p*-type.

The purpose of this paper is to offer a simple but general explanation of the ZnTe doping singularity. This explanation does not depend on any detailed model of dopant compensation or theory of the resistance of II-VI semiconductors to *p*-type doping. It merely asserts (on the basis of a theory of deep levels) that ZnTe is unique in having a native antisite defect, Zn_{Te}, that is a strong *p*-type dopant, capable of generating holes and compensating donors of any origin. ZnTe is also unique in that its antisite dopant Li_{Te} is also a shallow acceptor, not a deep trap. The electronic structures of these antisite defects provide a simple, general, and experimentally verifiable explanation of why ZnTe is easily doped *p* type, and allow us to avoid the more difficult issue of specifying the detailed mechanisms by which most II-VI semiconductors occur *n* type. Hence, instead of asking "Why are ZnSe and most other II-VI semiconductors *n* type" we address the question "what unique feature of ZnSe and ZnTe causes them to *change* their doping proclivities from *n* type for ZnSe to *p* type for ZnTe"? We propose that the relevant feature is the shallow acceptor character of the Zn_{Te} antisite defect.

At first glance, most of the mechanisms that might frustrate *p*-type doping appear to be smooth functions of the chemical compositions of the II-VI semiconductors, and so it is difficult to understand why ZnTe can be doped *p* type and most other II-VI semiconductors cannot. One might still expect that $ZnTe_{1-x}Se_x$ (Ref. 15) or $(ZnTe)_{1-x}(CdSe)_x$ alloys (Ref. 16) would smoothly change from *p*-type to semi-insulating to *n*-type conductivity as x increases. However the concept of abrupt deep-shallow transitions[17-23] of impurity character as a function of host chemical composition raises the possibility that a defect intimately involved in obstructing the *p*-type doping in most II-VI semiconductors might have quite a different character in ZnTe. We predict that the native antisite defect Zn_{Te} should be a shallow acceptor in ZnTe, but that its analog Zn_{Se} should be a deep trap in

ZnSe. At the extreme, even Li (a common dopant) has an antisite defect with the same character, changing from shallow in ZnTe to deep in ZnSe. The deep-shallow transition of the antisite defects' characters occur *abruptly* as functions of alloy composition x, and can naturally explain why ZnTe is p type while other II-VI semiconductors are not.

II. MODEL

Our model uses the well-established theory of deep levels,[18,19,24,25] which has successfully explained a broad range of data, including the physics of the N trap in $GaAs_{1-x}P_x$ (which produces an energetically "shallow" level within <0.1 eV of the conduction-band edge in GaP, a deep trap more than 0.1 eV below the edge for $x \simeq 0.5$, and a resonance in the conduction band of GaAs),[17,19,18] rapid III-V-compound laser degradation,[26] bulk and surface core excitons[27–31] in Si and III-V compound semiconductors, intrinsic surface states,[32,33] Schottky-barrier heights,[25–38] and the deep-shallow behavior of defects such as the *DX* center in $Al_xGa_{1-x}As$ semiconductors[39] and $GaAs/Al_xGa_{1-x}As$ superlattices.[20–23,40]

In this model, deep energy levels are obtained by solving the secular equation

$$\det[1-(E-H_0)^{-1}V]=0 ,$$

where H_0 is the host Hamiltonian operator, V is the defect potential operator, and the level's energy E is assumed to have an infinitesimal positive imaginary part. Hjalmarson *et al.*[19] have presented predictions of A_1 deep levels based on this theory. In a Löwdin basis[41] of sp^3s^* orbitals centered at each site[42] the defect potential is diagonal and related to atomic energies.[19] The host Hamiltonian is treated in an empirical tight-binding model, the parameters of which have been fit to observed band gaps in several semiconductors and follow chemical trends from one semiconductor to another.[42] Charge-transfer effects, such as those treated by Jansen and Sankey,[7] are not included explicitly in this model. Nevertheless the model produced excellent agreement with measured surface states of ZnSe and ZnTe,[33] and so is expected to describe well any localized perturbation associated with a defect. Details of the model have been published.[24,25,32] We consider only substitutional defects because common interstitials such as Li and Zn are normally donors[7] and because most interstitials are notoriously sensitive to charge transfer and the local environment,[7,43–45] a property that we regard as unlikely to be associated with the nearly universal doping properties of II-VI semiconductors—although we recommend the work of Chadi and Chang[46] for a somewhat different viewpoint.

III. RESULTS

As a test of the model's ability to describe defects in II-VI semiconductors, we first consider Li doping of zinc-blende-structure ZnSe and find Li_{Zn} to be a shallow acceptor, in agreement with the data.[47] Furthermore, experiments[47] reveal a deep acceptor related to Li 0.2 to 0.3 eV above the valence-band maximum. Since interstitial Li is a donor,[7] this acceptor can likely be associated with the theoretically predicted 0.4 eV antisite Li_{Se} deep level, and indicates that the theory lies only slightly above the data in energy.[48,49] Thus the Li doping data for ZnSe are well described by the present theory, lending support to the other theoretical predictions.

Since the spirit of the present work is to understand global chemical trends, we limit our discussion to a mean-field one-electron theory of neutral defects, and omit the Coulomb effects normally associated with small to moderate splittings of order 0.2 eV between different charge states of a defect.[7,43–35] This simplification makes all of the energy levels associated with different charge states of a defect degenerate and allows a simple discussion of global trends without concern for many-electron effects. Furthermore, the effects of these rather small effects can be included *a posteriori* with very little difficulty.[43–35]

Each of the s- and p-bonded impurities has eight "deep" spin orbitals with energies near the fundamental band gap, two A_1 or s-like and six T_2 or p-like spin-orbitals. Normally the energies of the T_2 spin orbitals lie above the A_1 spin orbitals. In fact, spin-orbit splitting, which is included in the theory, causes the T_2 levels to split slightly into $p_{3/2}$-like Γ_8 levels and $p_{1/2}$-like Γ_7 levels. However, for simplicity we shall refer to both levels as T_2 in the text, while plotting the spin-orbit split levels in the figures. (In the T_d double-group notation, A_1 becomes Γ_6).

The predictions of the theory for substitutional s- and p-bonded defects in ZnSe and ZnTe are summarized in Figs. 1–4. These figures indicate the doping character of each element of the Periodic Table.

A notable feature of the predictions is that the T_2 defect energy levels for Zn on either site in any host are virtually identical to those of the corresponding Li defect because these atoms have essentially the same atomic orbital energies. This means, in particular, that Zn_{Se} and Li_{Se} have about the same energy levels and the same doping character, except that Zn provides one more electron than Li.

Figures 1–4 should be examined as follows: In general the deep levels move up (down) as one moves to the left (right) across a row of the Periodic Table. The host atom has A_1 and T_2 levels occupied in the valence band and A_1 and T_2 levels empty in the conduction band. In the case of the Se-site Si in ZnSe, all the defect's levels lie above the corresponding Se levels, but the difference between Si and Se is not enough to cause any of the occupied deep levels to move up into the band gap. As a result, Si_{Se} is a double acceptor (denoted 2A) because it has two fewer electrons than Se, and these missing-electrons or holes "bubble up" from the deep Si levels in the valence band to the valence-band edge, where they are trapped in shallow acceptor levels by the Coulomb potential of the ionized Si atom. Similar reasoning applied to the empty level indicates that Br on the Se site of ZnSe is a single donor. Because Br is more electronegative than Se, its energy levels lie lower than those of Se, but the

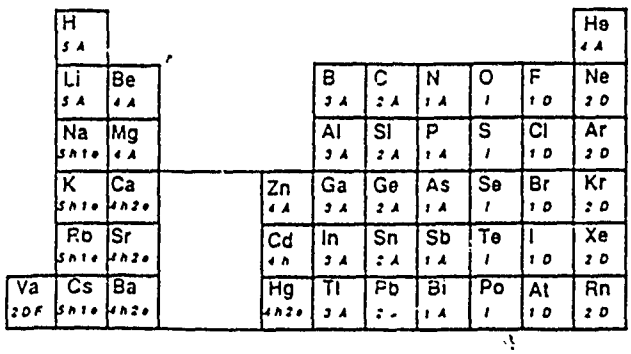


FIG. 1. Predicted doping character of Zn-site substitutional s- and p-bonded atoms in ZnSe. The character is denoted D for donor, A for acceptor, and I for isoelectronic defect (with no deep level in the gap). If the impurity produces one or more deep trap levels in the gap, this is indicated for each atom by the number of holes and electrons trapped in those levels for a single neutral defect, e.g., $5h1e$ denotes 5 holes and 1 electron—in a p-like deep level. See the text for further explanation.

Br-Se difference is not enough to pull one of the empty Se deep levels in the conduction band down into the gap. As a result, the extra electron of the neutral Br defect (relative to the Se atom the Br replaced) falls out of the deep level in the conduction band, the Br is autoionized, and the extra Br electron falls to the conduction-band edge, where it becomes a shallow donor electron orbiting the ionized Br impurity (denoted $1D$). As one moves to the left in the Periodic Table, the deep levels that were within the valence band for the host atom move up, sometimes into the band gap, becoming deep traps—as for Cs_{Se}, which (when neutral) has a p-like T_2 level in the gap occupied by five holes and one electron, and an A_1 state *below* the valence-band maximum filled by two electrons, and therefore is denoted $5h1e$. In some cases the defect is so electropositive that the relevant occupied deep levels pass through the gap into the conduction band, as for the anion vacancy (which is infinitely electropositive[50] and is denoted as the element "Va" in the figures). The A_1 and T_2 levels of the vacancy lie in the conduction band, with the six electrons of the T_2 level removed along with the Se atom during vacancy creation. Thus the Se vacancy is a double donor ($2D$) and has a false valence (F) with respect to Se, allowing it to donate $+2$ electrons rather than -6, because eight spin orbitals ($2A_1+6T_2$) have crossed the gap into the conduction band as the Se host atom has been "transmuted" into a vacancy. Similarly a Zn-site Cl atom has deep levels derived from the empty conduction-band levels of Zn, the A_1 level is pulled down

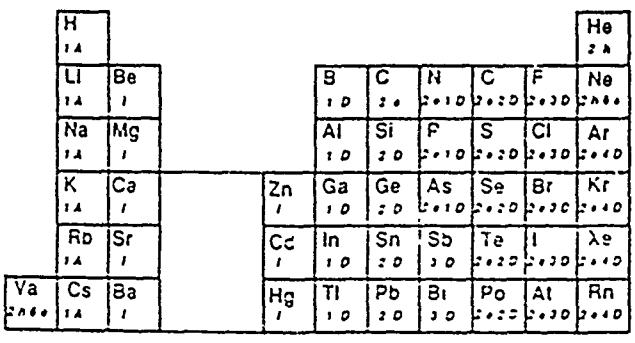


FIG. 2. Predicted doping character for Se-site substitutional atoms in ZnSe, as in Fig. 1

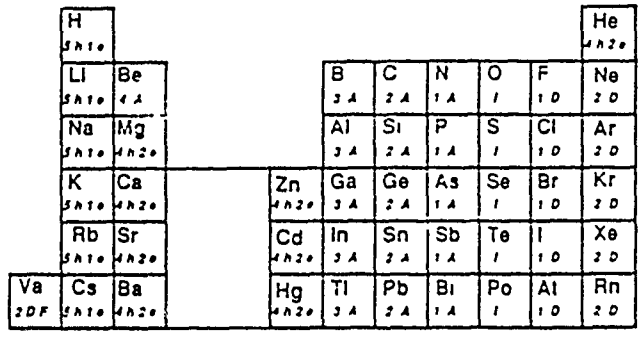


FIG. 3. Predicted doping character for Zn-site substitutional atoms in ZnTe, as in Fig. 1.

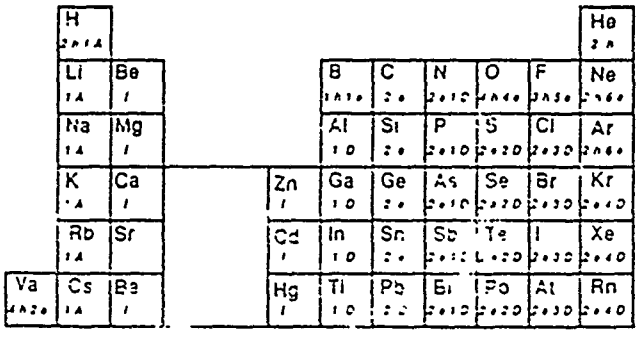


FIG. 4. Predicted doping character for Te-site substitutional atoms in ZnTe, as in Fig. 1. Note that the Zn antisite defect produces a shallow [quadruple (Ref. 51)] acceptor.

into the gap because Cl is more electronegative than Zn, while the T_2 levels remain in the conduction band. The five extra electrons of neutral Cl distribute themselves with two in the A_1 deep level and three in the conduction band; hence Cl produces a deep doubly occupied A_1 level (2 *e*) and a triple donor (3 *D*). Some defects, such as oxygen on a Se-site in ZnSe neither change the number of electrons nor introduce new deep levels into the gap, and so are termed isoelectronic (*I*).

By generating Periodic Tables such as Figs. 1–4 for a wide range of II-VI semiconductors, we have been able to determine how the predicted characters of defects are different in various semiconductors. Here we discuss only ZnSe, which has similar defect deep-level behavior to other II-VI semiconductors, and ZnTe, which exhibits unusual *p*-type doping behavior.

The theory predicts that neutral Zn_{Te} in ZnTe is a quadruple[51] acceptor (4 *A*), whereas native cation-on-anion-site defects in other II-VI semiconductors produce T_2-symmetric deep traps in the band gap: e.g., (4*h*2*e*) for Zn_{Se} in ZnSe. (See Figs. 1 and 3). Charge-transfer and Coulomb splitting effects omitted from the model may actually prevent the formation of the quadrupole acceptor state, but some level of acceptor character is nevertheless to be expected. (Sec. IV treats this issue in more detail.) This native antisite defect yields four holes to the valence band and is a powerful *p* dopant *in ZnTe only*. Moreover, as a function of alloy composition x in $ZnTe_{1-x}Se_x$ or $(ZnTe)_{1-x}(CdSe)_x$, the relevant deep level should move from below the valence-band maximum into the gap—and the character of the Zn-on-anion-site defect should change rather *abruptly* from a *p*-type acceptor to a deep trap which no longer dopes the material *p*-type. (See Figs. 5 and 6). Hence the material should change, rather abruptly, from being *p* type to *n* type. The abrupt character of this transition should be an experimental sig-

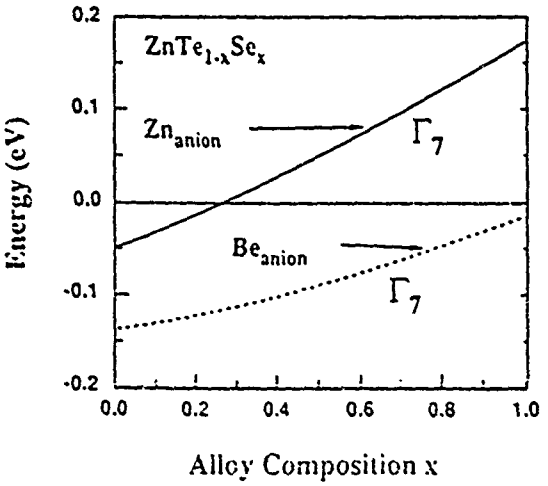


FIG. 5. Predicted dependence on alloy composition x of the energy levels (in eV) for the anion-site defects Zn_{anion} (solid line), Li_{anion} (the same solid line), and Be_{anion} (dashed) in $ZnTe_{1-x}Se_x$. The $p_{1/2}$-like level derived from the T_2 deep level is labeled Γ_7.

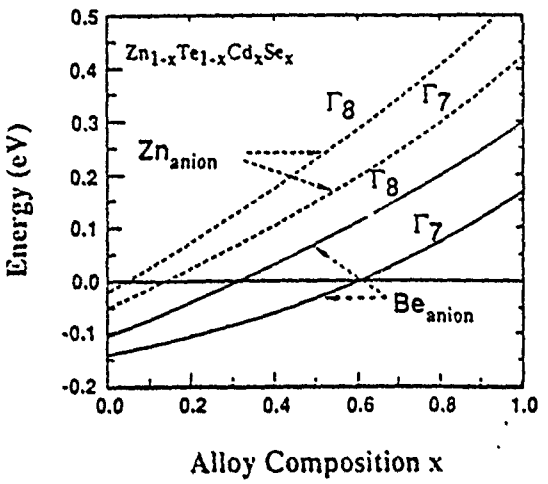


FIG. 6. Predicted dependence on alloy composition x of the energy levels (in eV) for the anion-site defects Zn_{anion} (solid line) and Be_{anion} (dashed) in $(ZnTe)_{1-x}(CdSe)_x$. The $p_{3/2}$-like ($p_{1/2}$-like) levels derived from the T_2 deep levels are labeled Γ_8 (Γ_7). For $x=0$, and for a neutral Zn_{Te} defect, the Γ_7 level is occupied by two electrons and the Γ_8 level by four holes.

nature of the antisite-defect doping mechanism, which perhaps can be detected by a sudden change of doping character with alloy composition x in $ZnTe_{1-x}Se_x$ or $(ZnTe)_{1-x}(CdSe)_x$, by optical spectroscopy (allowing for the fact that both the valence-band edge and the deep level have T_2 symmetry), or by pressure or strain measurements which cause the deep level to move into or out of the fundamental band gap.

Zn is not the only antisite defect that is predicted to produce a strong *p* dopant. Li_{Te} (although unlikely to occur in large concentrations) is likewise a shallow acceptor yielding nominally five holes[51] to the valence band. The energy of the foreign antisite Li deep level is essentially the same as that of native antisite Zn—and so substitutional antisite Li should also undergo a similar deep-shallow transition as a function of alloy composition. (Note that this paper does not consider the questions of whether the Li would be mechanically stable on such a site or whether it would naturally occupy such a site with reasonable probability.) Thus, for example, if we assume an oversimplified situation for the Li doping of ZnTe in which there are no vacancies or other defects, and if we denote the concentrations of interstitial-, Zn-site, Te-site, and Se-site Li by $[Li_I]$, $[Li_{Zn}]$, $[Li_{Te}]$, and $[Li_{Se}]$, then the conditions for *p*-type character in ZnTe and ZnSe are

$$5[Li_{Te}]+[Li_{Zn}]-[Li_I]>0\,,$$

and

$$-[Li_{Se}]+[Li_{Zn}]-[Li_I]>0\,.$$

Clearly the antisite Li could make a difference in the doping character of the host, *provided its concentration is large enough*. Now, in an equilibrium situation, the Li_{Te}

concentration is likely to be exponentially smaller than $[Li_{Zn}]$, and so Li by itself should not change the doping character of a II-VI semiconductor; but this example illustrates in a simple case how the concentrations of the different point defects determine the doping character.

Another interesting prediction of the theory is that Be on an anion site is a (quadruple[51]) acceptor in both ZnSe and ZnTe, while remaining isoelectronic when occupying a Zn site. (See Fig. 5). This means that heavy Be doping may also make ZnSe p type.[52]

On the anion site, only Zn, Mg, Cd, Li, and H and He are predicted to have different doping characters in ZnSe and ZnTe. On the cation site Ar, F, O, Bi, Sb, Sn, Ge, Si, B, H and the vacancy have different doping characters. See Figs. 1–4. Predicted energy levels are given in Figs. 7–22. (Impurities with no deep levels near the fundamental band gap behave according to the conventional rules and are not displayed.)

It is noteworthy that the anion vacancy is a double donor for both ZnSe and ZnTe—as anticipated by the extensive self-compensation literature. This is a case of false valence,[53] because the removal of a column-VI atom should add six holes to the valence band, not two electrons to the conduction band. The removal of, say, Te in ZnTe is achieved theoretically by driving its atomic energies to infinity as well as adding six holes. In the case of ZnTe and other II-VI semiconductors, this causes an A_1 and a T_2 level to move from the valence band into the conduction band.

Within the context of the conventional self-compensation picture, each acceptor in p-type doped ZnSe is compensated by a double-donor anion vacancy generated as a result of acceptor doping. If the same energy that produced an anion vacancy instead produced an anion-site antisite Zn, then in ZnSe every antisite defect would compensate roughly two acceptors (or one vacancy), but in ZnTe this would not happen. Moreover, in ZnTe roughly every two vacancies would be compensated by one antisite defect, blocking the conventional self-compensation process. Antisite defects are common in a variety of semiconductors, as argued by Van Vechten and others.[54] The formation energy should be less for the an-

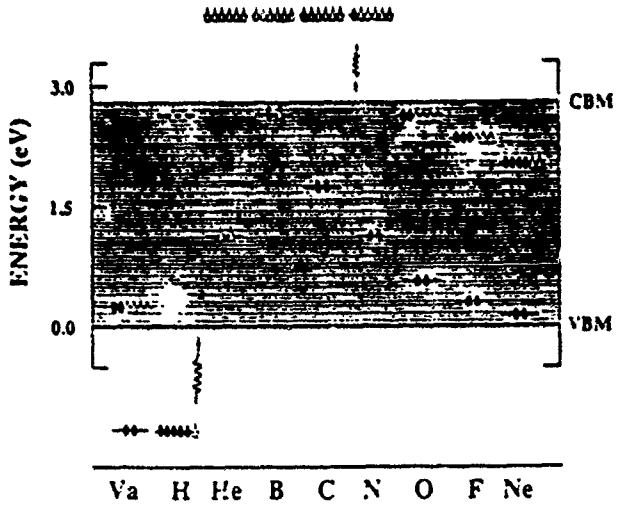


FIG. 7. Predicted deep energy levels in the gap (striped) of ZnSe due to s- and p-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa. Their occupancies, when neutral, are indicated by open triangles for holes and by closed circles for electrons

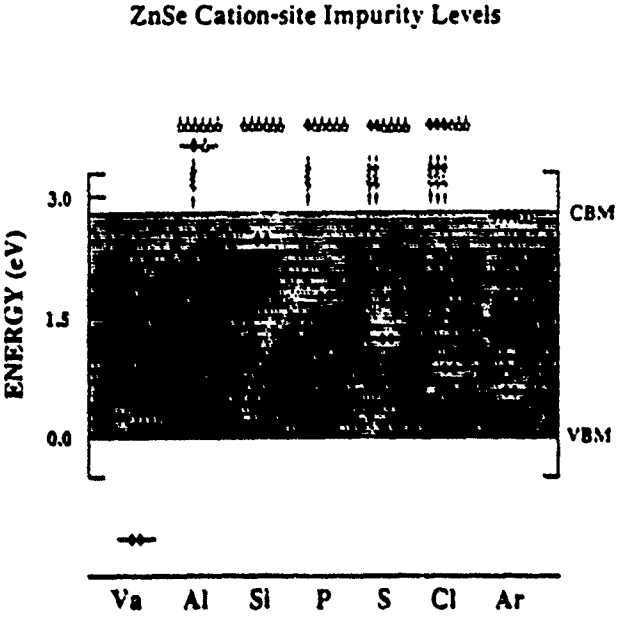


FIG. 8. Predicted deep energy levels in the gap of ZnSe due to s- and p-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

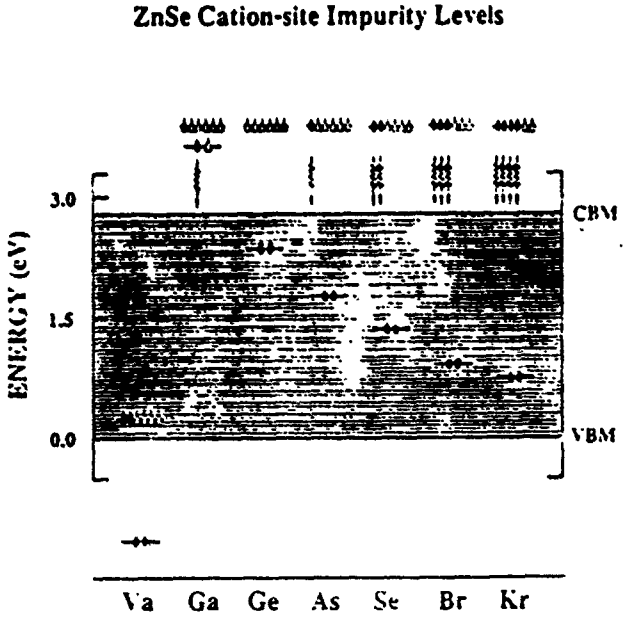


FIG. 9. Predicted deep energy levels in the gap of ZnSe due to s- and p-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

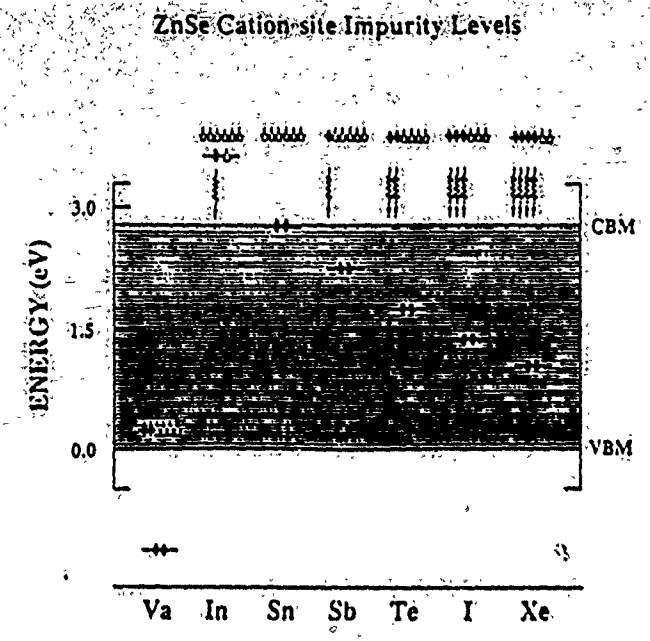


FIG. 10. Predicted deep energy levels in the gap of ZnSe due to *s*- and *p*-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

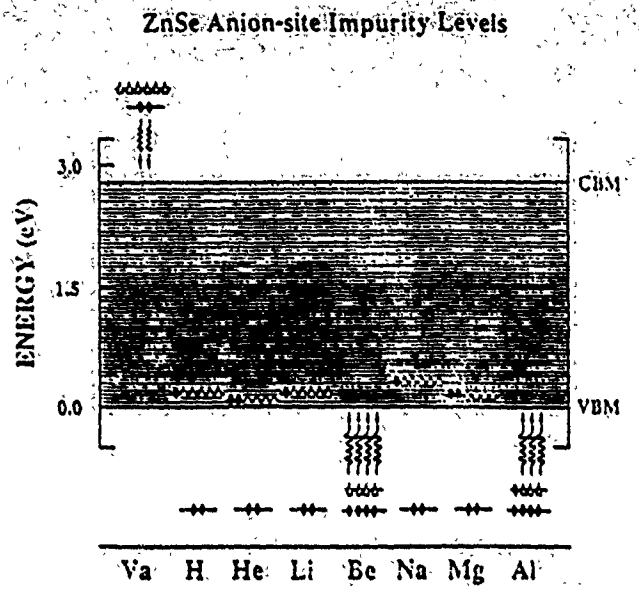


FIG. 12. Predicted deep energy levels in the gap of ZnSe due to *s*- and *p*-bonded substitutional impurities on the Se site. The impurities are indicated on the abscissa.

tisite defect than for the vacancy, and the number of antisite defects to be expected at equilibrium should exceed the number of vacancies. Stated more simply, it normally costs less energy to rearrange a bond than to break it. Hence in most crystal-growth processes, an adequate number of Zn_{Te} defects should be created, and the self-compensation would be expected to be impotent in ZnTe but not inZnSe. For example (Fig. 23), in ZnSe with one Na acceptor, one Se vacancy, and one Zn_{Se} antisite defect, the shallow Na acceptor level is filled by one of the vacancy's two electrons, with the other electron occupying the T_2 Zn deep level in the gap (leaving it with three electrons and three holes). There is no *p*-type conductivity. In contrast, the same situation in ZnTe leaves the two electrons from the vacancy occupying two of the five empty shallow orbitals of both Na and Zn_{Te}, leaving three thermally ionizable holes and *p*-type conductivity. Therefore, we propose that *p*-type conductivity is rather easily achieved only in ZnTe of the common II-VI semiconductors because only in ZnTe is the antisite defect Zn_{Te} a shallow acceptor rather than a deep trap.

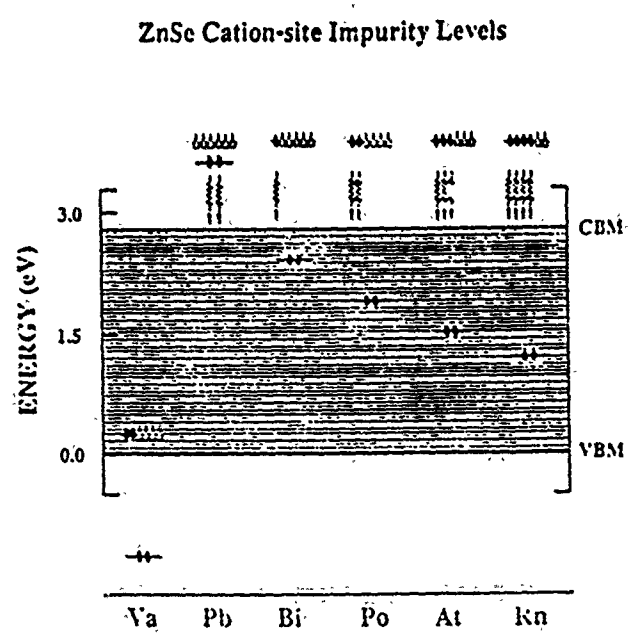


FIG. 11. Predicted deep energy levels in the gap of ZnSe due to *s*- and *p*-doped substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

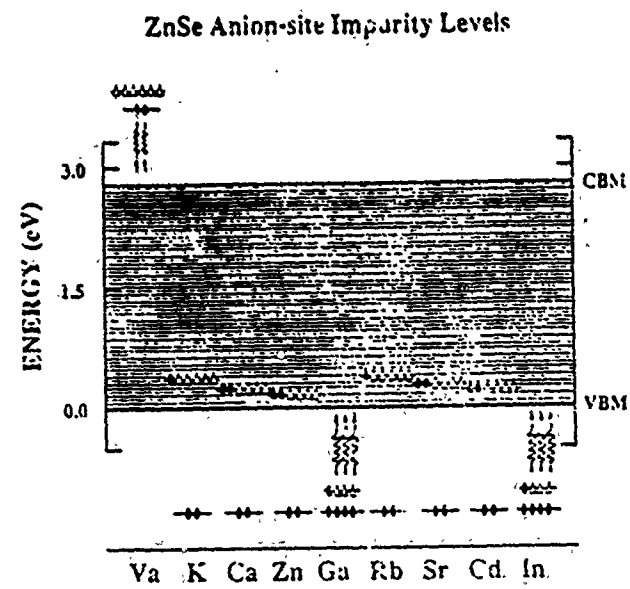


FIG. 13. Predicted deep energy levels in the gap of ZnSe due to *s*- and *p*-bonded substitutional impurities on the Se site. The impurities are indicated on the abscissa. Al (not shown) has a similar electronic structure to Ga.

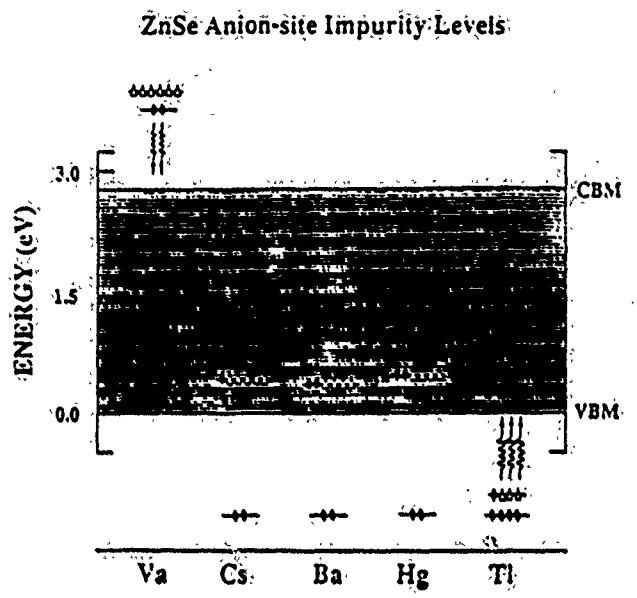


FIG. 14. Predicted deep energy levels in the gap of ZnSe due to *s*- and *p*-bonded substitutional impurities on the Se site. The impurities are indicated on the abscissa.

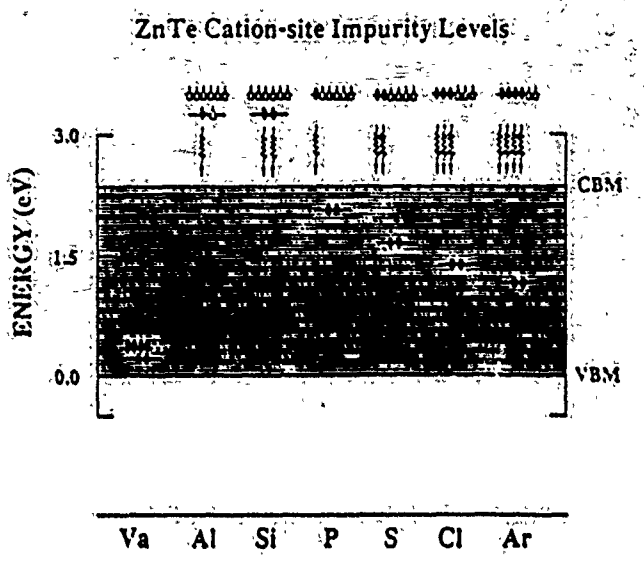


FIG. 16. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

IV. POTENTIAL PROBLEMS WITH THE MODEL

Two concerns that we have about the theory are (i) an explanation is required of why II-VI semiconductors other than ZnTe have an *n-type-doping* proclivity and (ii) the omission of Coulomb effects from the theory could cause our level predictions to be in error more than we currently believe.

A. *n*-type doping

If the antisite defects alone occur in the same concentrations in ZnSe and ZnTe, then one must explain why ZnSe apparently is more easily doped *n* type. In our model, the *n*-type dopants must first compensate all the antisite defects, after which they produce donors—and the physics is the same for ZnSe and ZnTe. However, we suspect that there are fewer Zn_{Se} defects in ZnSe than comparable antisite defects in ZnTe, basically because Zn is more similar to Te than to Se, both in its atomic size and its electronic energy levels. Doping the semiconductor *n* type with, say, Ga could exacerbate such differences. Therefore we believe that the *n*-type doping differences of ZnSe and ZnTe can be plausibly explained.

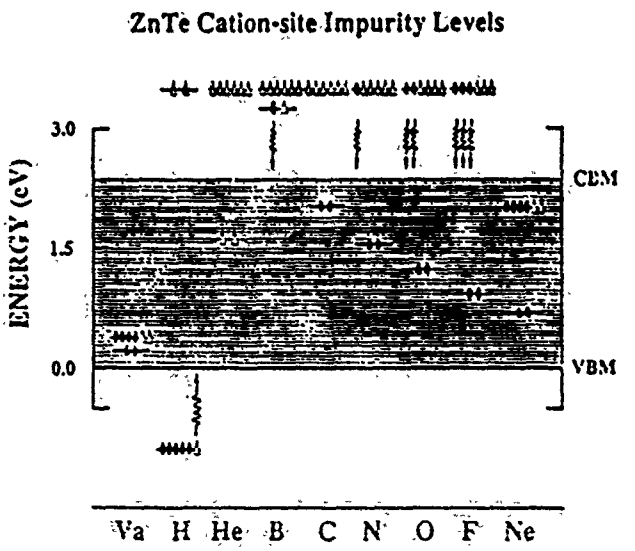


FIG. 15. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

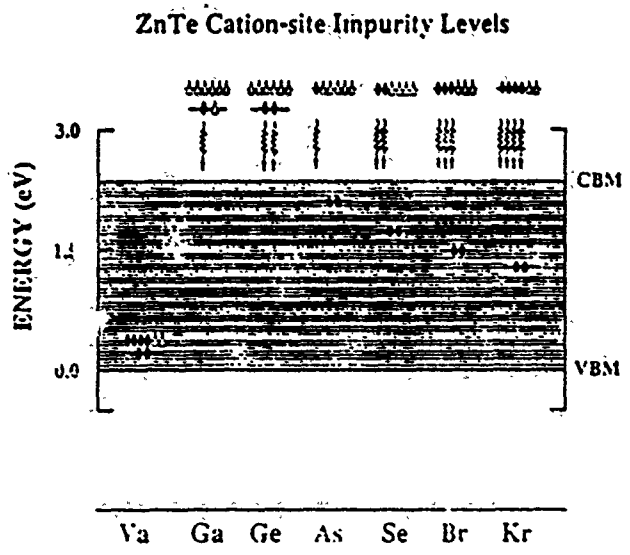


FIG. 17. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

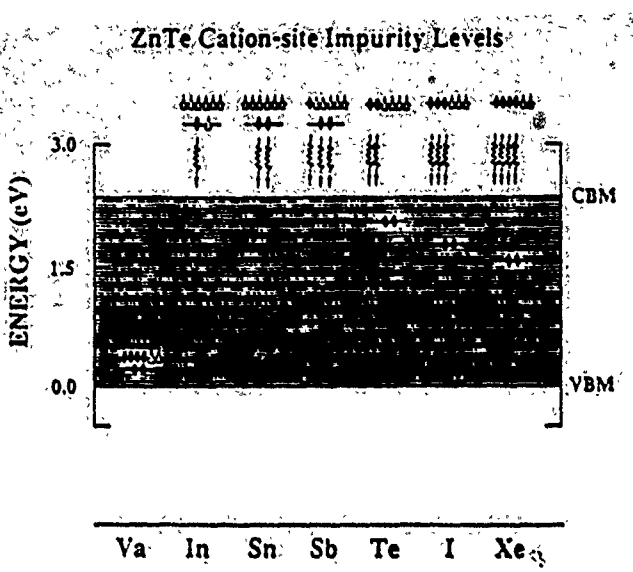


FIG. 18. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

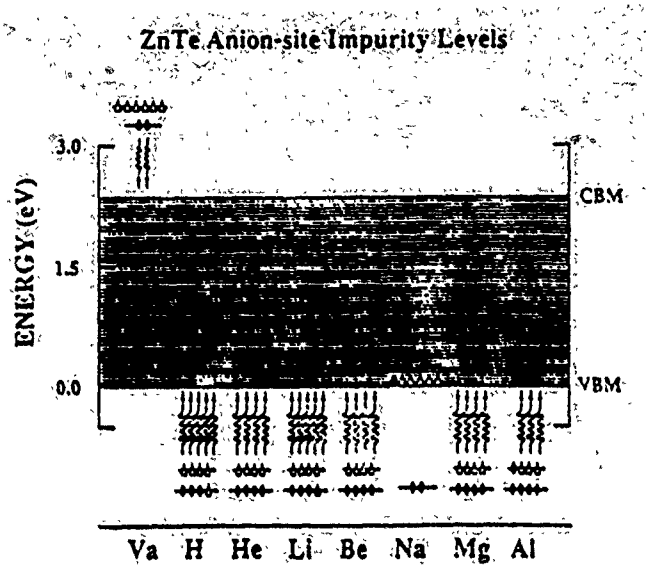


FIG. 20. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Te site. The impurities are indicated on the abscissa.

B. Coulomb effects

It is very difficult to precisely determine the Coulomb effects on deep levels in II-VI semiconductors. Therefore we have actually executed three rather different theoretical calculations which generally agree that the position of the Zn_{Te} deep level with respect to the ZnTe valence-band edge is about 0.2–0.5 eV lower than the energy of the Zn_{Se} level with respect to the ZnSe band edge (in any charge state). These theories do not agree, however, on the precise location of the neutral Zn_{Te} deep level, which the present theory places in the ZnTe valence band, while the other two theories place the level near midgap (but with a theoretical uncertainty due to finite supercell size of order 1 eV).

The first calculations, the ones we have described here, are based on an empirical tight-binding model of electronic structure, and incorporate Coulomb effects only implicitly in the parameters of the model. These calculations are, strictly speaking, for *neutral defects* in the ZnSe or ZnTe host, and do not directly consider the negative charging of Zn_{Te} as the holes in the Zn deep level bubble up to the valence-band maximum. This charging will cause the deep level to move up in energy about 0.2–0.3 eV for each additional electron trapped in the deep level[7] (an effect omitted from the model)—hence the Zn_{Te} level may move into the gap after the first or second electron is trapped on the Zn (i.e., after one or two holes "bubble up" to the valence-band maximum)—making Zn_{Te} a single or double acceptor rather than a full fourfold accep-

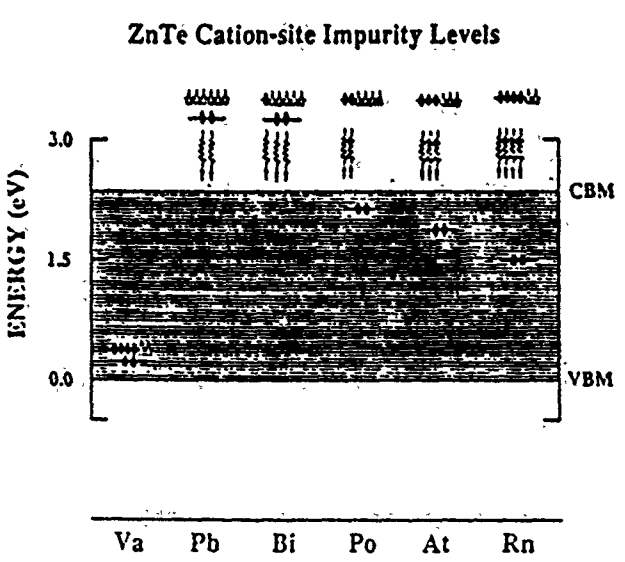


FIG. 19. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Zn site. The impurities are indicated on the abscissa.

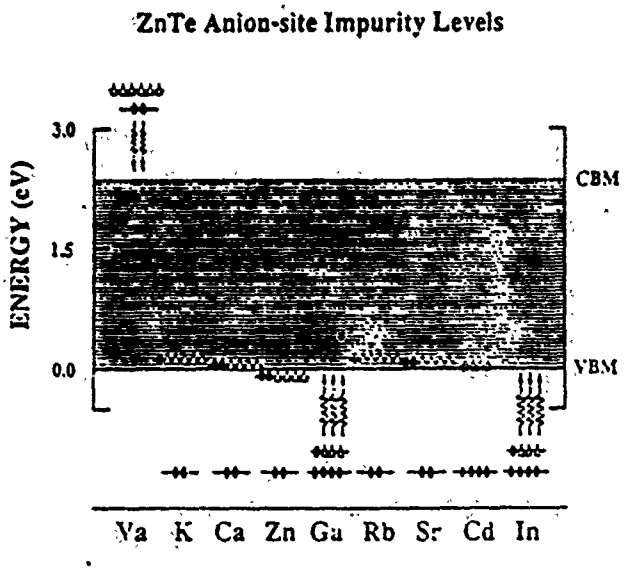


FIG. 21. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Te site. The impurities are indicated on the abscissa. Al (not shown) has a similar electronic structure to Ga.

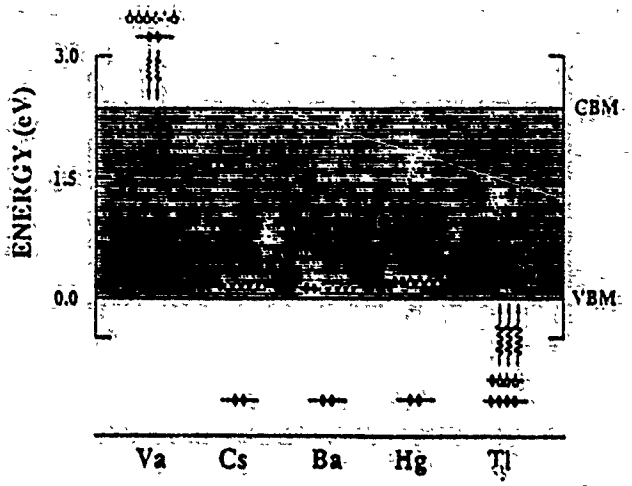


FIG. 22. Predicted deep energy levels in the gap of ZnTe due to *s*- and *p*-bonded substitutional impurities on the Te site. The impurities are indicated on the abscissa.

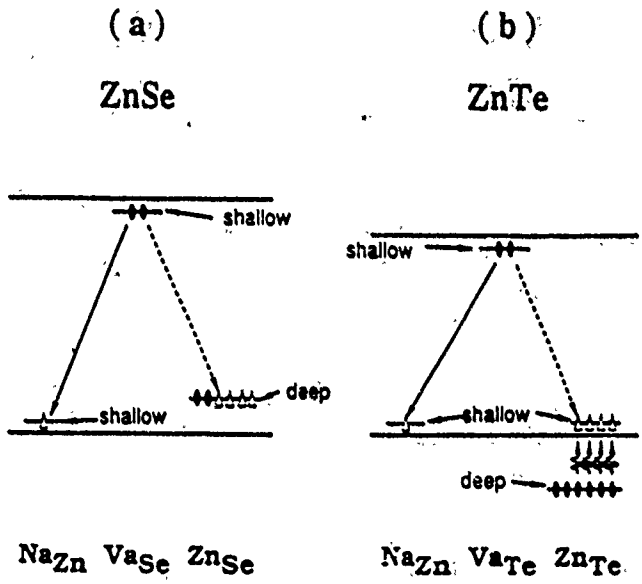


FIG. 23. Illustrating (a) self-compensation in the presence of a Zn_{Se} defect in ZnSe and (b) "anti" self-compensation" by Zn_{Te} in ZnTe. The self-compensation mechanism assumes that doping with an acceptor such as Na_{Zn} leads to production of a compensating double-donor anion vacancy. For simplicity we assume that there are only three defects: Na_{Zn}, the anion vacancy Va_{anion}, and the Zn-antisites Zn_{Se} and Zn_{Te}. In both cases, an electron (solid circle) from the anion vacancy's shallow donor level compensates (solid lines) the hole (open triangles) of the Na_{Zn} defect's shallow acceptor level. The remaining vacancy electron is trapped (dashed lines) (a) by the Zn_{Se} deep level in ZnSe or (b) by the shallow acceptor level in ZnTe, because the holes of the Zn_{Te} deep T_2 level have bubbled up (wiggly lines) to the shallow acceptor level. The remaining three holes in the Zn_{Te} shallow level are thermally ionizable in ZnTe, render the self-compensation by the vacancy impotent, and dope the material *p* type, whereas the holes in the deep level of ZnSe are not thermally ionizable, so that the Zn_{Se} level traps both electrons and holes, reducing conductivity.

tor. Thus the qualitative physics governing doping proclivity will be the same, although the number of free holes per defect will be smaller, so long as the deep level of the neutral Zn_{Te} defect is correctly predicted to lie below the valence-band maximum.

The second[7] and third[55] theories are local-density theories for a defect in a supercell of finite size, and produces levels for the neutral defect Zn_{Te} somewhat higher than the levels of the present theory. However, due to the finite size of the supercells, the predicted levels are actually bands of order 1 eV wide. Furthermore local-density theory is known to produce incorrect band gaps—and this can lead to errors in deep-level energies.[56]

The second theory is the tight-binding-like version of local-density theory discussed by Jansen and Sankey,[7] while the third is a pseudofunction[55] implementation of local-density theory. We executed the pseudofunction calculations for supercells of $N_s=16$ and $N_s=32$ atoms, and found that (relative to the valence-band maximum) the deep level moved down in energy with increasing supercell size, as the level's width decreased. Perhaps the local-density-theory levels of Zn_{Te} would also lie in the valence band if the supercell size were increased until convergence were achieved.

We conclude on the basis of these theories that we cannot definitively place the neutral Zn_{Te} deep level below the Zn_{Te} valence-band maximum—nor can we place it in the gap definitively. Therefore the issue of the location of this level will have to be determined experimentally.

In this regard we note that Li doping produces a mysterious deep level in ZnSe just above the valence-band maximum,[47] and no corresponding level in the gap of ZnTe—as the present (first) theory predicts. While this fact lends credence to the present theory, it does not confirm the theory unambiguously (because the nature of the Li-related defect is unknown).

We conclude that there is good circumstantial evidence for the Zn_{Te} defect yielding a deep level in the valence band of ZnTe, the kind of deep level needed to explain the *p*-type doping proclivity of ZnTe. The case is not airtight, appears to be unresolvable by theory, and calls for careful experimental investigation. In particular, careful studies of the dependence of doping proclivity on stoichiometry and on alloy composition in $ZnTe_{1-x}Se_x$ might confirm or exclude the present model. (Note that a rather sudden change in doping proclivity as a function of alloy content x is implied by the present model.)

V. SUMMARY

The viewpoint presented here is admittedly somewhat oversimplified in the interest of illustrating the essential physics of the problem. Nevertheless we hope that efforts to test this simple picture will be attempted because, if the picture is correct as we believe, then schemes for overwhelming the *p*-type doping problem in some other II-VI semiconductors than ZnTe will be feasible.[8] Two obvious schemes are (i) to grow ZnSe under conditions unfavorable to Zn_{Se} formation, and (ii) to engineer the electronic structure of a perturbed version of ZnSe so that the antisite defect's deep level is resonant with the

valence band.

Finally our conclusions are based only on the occurrence of defects with deep-level electronic structures similar to those predicted for the Zn (and Li) antisite defects. The model for the *p*-type doping will still be valid if another defect has similar electronic structures in ZnSe and ZnTe. Furthermore the model does not rely in any way on a detailed picture of the "self-compensation" process or a knowledge of the detailed forces on defects.

ACKNOWLEDGMENTS

We are grateful to the U.S. Office of Naval Research, the U.S. Air Force Office of Scientific Research, and the U.S. Defense Advanced Research Projects Agency for their generous support (Contract No. N00014-84-K-03552, No. AF-AFOSR-85-0331, and No. N0530-0716-05).

---

*Present and permanent address.

[1]F. A. Kroger and H. A. Vink, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1956), Vol. III, p. 310.

[2]R. F. Brebick, J. Phys. Chem. Solids 4, 190 (1958); 18, 116 (1961).

[3]G. Mandel, Phys. Rev. 134, A1073 (1964).

[4]P. J. Dean, Czech. J. Phys. B 30, 272 (1980).

[5]R. N. Bhargava, J. Cryst. Growth 59, 15 (1982), and references therein.

[6]See, for example, C. H. Henry, K. Nassau, and J. W. Shiever, Phys. Rev. B 4, 2453 (1971), and references therein.

[7]R. W. Jansen and O. F. Sankey, Phys. Rev. B 39, 3192 (1989); Solid State Commun. 64, 197 (1987); J. Vac. Sci. Technol. B 6, 1240 (1988). For Zn_{Te} and Zn_{Se} these authors find bands whose centers are at 1.2 and 1.5 eV above the respective valence-band maxima. R. W. Jansen, D. S. Wolde-Kidane, and O. F. Sankey, J. Appl. Phys. 64, 2415 (1988).

[8]S. Y. Ren, J. D. Dow, and S. Klemm, J. Appl. Phys. 66, 2065 (1989); J. Shen, J. D. Dow, and S. Y. Ren, J. Appl. Phys. 67, 3761 (1990).

[9]See, for example, M. Aven and J. Z. Devine, J. Lumin. 7, 195 (1973); Y. S. Park and B. K. Shin, *Topics in Applied Physics* (Springer, Berlin, 1977), p. 33, and references therein.

[10]A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 28, 946 (1983).

[11]T. Yasuda, I. Mitsuishi, and H. Kukimoto, Appl. Phys. Lett. 52, 57 (1988), and references therein.

[12]J. Nishijawa, K. Itoh, Y. Okuno, and F. Sakurai, J. Appl. Phys. 57, 2210 (1985).

[13]T. C. McGill (private communication).

[14]F. Ren and G. D. Watkins, Phys. Rev. Lett. 56, 2310 (1986); G. D. Watkins, *ibid.* 33, 223 (1974).

[15]A. Koukitu, J. Cryst. Growth 84, 425 (1987).

[16]S. A. Azimov, T. M. Razykov, V. I. Vyaly, and B. Kh. Kadyrov, Phys. Status Solidi A 96, 281 (1986).

[17]D. J. Wolford, B. G. Streetman, W. Y. Hsu, J. D. Dow, R. J. Nelson, and N. Holonyak, Jr., Phys. Rev. Lett. 36, 1400 (1976).

[18]W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[19]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980). For a discussion of the ideas on which Ref. 19 was based, see Ref. 18.

[20]S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B 38, 10677 (1988).

[21]J. D. Dow, S. Y. Ren, and J. Shen, in *Deep Impurity Levels in Semiconductors, Semiconductor Alloys, and Superlattices*, Vol. 183 of *NATO Advanced Research Workshop on Properties of Impurity States in Superlattice Semiconductors*, edited by C. Y. Fong, I. P. Batra, and S. Ciraci (Plenum, New York, 1988), p. 175.

[22]R. D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, in *Interfaces, Superlattices, and Thin Films*, Vol. 77 of *The Materials Research Society Symposium Proceedings*, edited by J. D. Dow and I. K. Schuller (MRS, Pittsburgh, PA, 1987), p. 545.

[23]S. Y. Ren and J. D. Dow, J. Appl. Phys. 65, 1987 (1989).

[24]J. D. Dow, *Localized Perturbations in Semiconductors*, In *Highlights of Condensed-Matter Theory*, Proceedings of the International School of Physics "Enrico Fermi," Course 89, Varenna, 1983, edited by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna, Italy, 1985), pp. 465–494.

[25]J. D. Dow, in *Microscopic Identification of Electronic Defects in Semiconductors*, Vol. 46 of the *Materials Research Society Symposiums Proceedings*, edited by N. M. Johnson, S. G. Bishop, and G. D. Watkins (MRS, Pittsburgh, 1985), p. 71.

[26]J. D. Dow and R. E. Allen, Appl. Phys. Lett. 41, 672 (1982).

[27]H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Rev. B 24, 6010 (1981).

[28]R. E. Allen and J. D. Dow, Phys. Rev. B 24, 911 (1981).

[29]H. P. Hjalmarson, H. Büttner, and J. D. Dow, Phys. Lett. 85A, 293 (1981).

[30]K. E. Newman and J. D. Dow, Solid State Commun. 50, 587 (1984). See B. A. Bunker, S. L. Hulbert, J. P. Stott, and F. C. Brown, Phys. Rev. Lett. 53, 2157 (1984), for experimental confirmation of this theory.

[31]M. A. Bowen, R. E. Allen, and J. D. Dow, Phys. Rev. B 30, 4617 (1984).

[32]J. D. Dow, R. E. Allen, and O. F. Sankey, in *Intrinsic and Extrinsic Surface Electronic States of Semiconductors. Chemistry and Physics of Solid Surfaces, V*, Vol. 35 of *Springer Series in Chemical Physics* edited by R. Vanselow and R. Howe (Springer-Verlag, Berlin, 1984), p. 483 and references therein.

[33]R. P. Beres, R. E. Allen, and J. D. Dow, Phys. Rev. B 26, 769 (1982).

[34]J. D. Dow and R. E. Allen, J. Vac. Sci. Technol. 20, 659 (1982).

[35]O. F. Sankey, R. E. Allen, and J. D. Dow, J. Vac. Sci. Technol. B 2, 491 (1984).

[36]O. F. Sankey, R. E. Allen, and J. D. Dow, Solid State Commun. 49, 1 (1984).

[37]O. F. Sankey, R. E. Allen, and J. D. Dow, J. Ultramicroscopy 14, 127 (1984).

[38]J. D. Dow, O. F. Sankey, and R. E. Allen, Materials Science Forum 4, 39 (1985).

[39]H. P. Hjalmarson and T. J. Drummond, Appl. Phys. Lett. 48,

657 (1986).
[40]R. E. Allen and J. D. Dow, Phys. Rev. B 25, 1423 (1982).
[41]P.-O. Löwdin, J. Chem. Phys. 18, 365 (1950).
[42]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[43]S. Lee, J. D. Dow, and O. F. Sankey, Phys. Rev. B 31, 3910 (1985).
[44]G. Kim, J. D. Dow, and S. Lee, Phys. Rev. B 40, 7888 (1989).
[45]G. Kim, J. D. Dow, and S. Lee, Arabian J. Sci. Engineer. 14, 513 (1989).
[46]D. J. Chadi and K. J. Chang, Phys. Rev. B 39, 10063 (1989).
[47]M. E. Agel'menev, A. N. Georgobiana, Z. P. Ilyukhina, A. D. Levit, L. S. Lepnev, and M. I. Sluch, Sb. Kratk. Seobshch. Fiz. AN SSSR Fiz. Inst. P. N. Lebedeva (USSR), 6, 18 (1987) [Sov. Phys.—Lebedev Inst. Rep. 6, 25 (1987)].
[48]It has been shown in Ref. 49 that, to an adequate approximation, the deep-level spectrum of a cluster of substitutional defects is the sum of the constituents' spectra, a fact that allows us to consider only isolated substitutional impurities and interstitials.
[49]O. F. Sankey, H. P. Hjalmarson, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. Lett. 45, 1560 (1980).
[50]M. Lannoo and P. Lenglart, J. Phys. Chem. Solids 30, 2409 (1969).
[51]The charge-state splitting effect omitted in the present theory could become significant when the defect becomes highly charged, and could alter the defect's mean-field energy levels so that the acceptor Zn_{Te} would bind fewer than four holes in shallow levels.
[52]Be, being light, will tend to diffuse rapidly, and should occupy some anion sites, much as Li appears to.
[53]C. S. Lent, M. A. Bowen, R. S. Allgaier, J. D. Dow, O. F. Sankey, and E. S. Ho, Solid State Commun. 61, 83 (1987).
[54]J. A. Van Vechten, J. Electrochem. Soc. 122, 419 (1975); 122, 423 (1975). See also J. P. Buisson, R. E. Allen, and J. D. Dow, Solid State Commun. 43, 833 (1982), and references therein.
[55]R. V. Kasowski, M.-H. Tsai, T. N. Rhodin, and D. D. Chambliss, Phys. Rev. B 34, 2656 (1986).
[56]R. E. Allen and J. D. Dow, J. Vac. Sci. Technol. 19, 383 (1981).

Scanning tunneling microscope images of native defects on the ZnSe(110) surface

Wei-Min Hu
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 and Department of Physics, University of Science and Technology of China,[a] *Hefei, China*

John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 8 February 1989; accepted 21 April 1989)

Scanning tunneling microscope images of native antisite defects at the relaxed (110) surface of ZnSe are predicted. The images of a particular sample depend on the sign of the voltage bias and the voltage sweep of the sample relative to the microscope tip, and whether that sweep causes a deep level to actively participate in the tunneling. Under certain conditions the images give the appearance of two defects at incorrect sites.

I. INTRODUCTION

In this paper we report the first theoretical scanning tunneling microscope (STM) images of surface antisite defects on the ZnSe(110) surface. This work was stimulated by measurements imaging II–VI semiconductor surfaces by scanning tunneling microscopy,[1] following earlier experiments on GaAs.[2,3] We follow the theory of Tersoff and Hamann[4] who have argued that an ideal STM obtains an image

$$I(\mathbf{r}) = \int_{\alpha}^{\beta} dE \, \Sigma_{i,b,\mathbf{R}} \times (i,b,\mathbf{R}|\delta(E-H)|i,b,\mathbf{R})|\psi_i(\mathbf{r}-\mathbf{R}-\mathbf{v}_b)|^2.$$

Here we have $\delta(E-H) = (-1/\pi)\ \mathrm{Im}\ G(E+i0)$ where $i0$ is a positive imaginary infinitesimal and G is the Green's function for the relaxed (110) surface. The index i ranges over the basis orbitals,[5] with the values s, p_x, p_y, p_z, and s^* (an excited-state orbital[6]); the position of an atom is denoted by its unit cell $\mathbf{R}$, and its site $\mathbf{v}_b$ in the cell, where b denotes either anion or cation. The $\psi_i(\mathbf{r}-\mathbf{R}-\mathbf{v}_b)$ are Löwdin orbitals centered on $\mathbf{R}+\mathbf{v}_b$. Electrons tunnel from the tip to the ZnSe conduction band when the sample is biased positive; and we integrate from $\alpha = E_c$ (the conduction-band edge) to $\beta = E_c + E_0$, where E_0 is typically 1 eV, and add in the contribution from any empty deep levels in the gap that lie within the range of the voltage sweep. A negatively biased ZnSe sample relative to the tip senses electrons below the valence-band maximum E_v, and we integrate from $\alpha = E_v - E_0$ to $\beta = E_v$, and add in the contribution of any occupied deep levels in the gap within the range of the voltage sweep. Here we compute the STM image obtained by measuring the current I for a tip at a fixed distance z above the x–y plane, which corresponds to the plane of the unrelaxed (110) surface. In the present work, we use the typical value $z = 5\ a_{\mathrm{Bohr}}$, where $a_{\mathrm{Bohr}} = 0.53$ Å, and sum $\mathbf{R}$ over 25 unit cells on the surface. (Subsurface contributions are negligible.)

Our calculations of STM images employed established[7] empirical tight-binding Hamiltonians H_0 and H[6,8,9] to represent the relaxed perfect and defective ZnSe(110) surfaces, respectively. The relaxation consisted of a rigid rotation of the anions up out of the surface plane through an angle $\omega = 25.6°$.[10–14] Details of the calculational method, which is based on the ideas of Vogl *et al.*,[6] Hjalmarson *et al.*,[8] and Allen *et al.*[9,15] are available elsewhere.[16]

II. RESULTS

A. Perfect surface

The images calculated for the perfect relaxed (110) surface of ZnSe are given in Figs. 1 and 2 and, as expected, show bumps at the Se and Zn sites, for negative and positive bias (of the sample with respect to the tip), respectively.

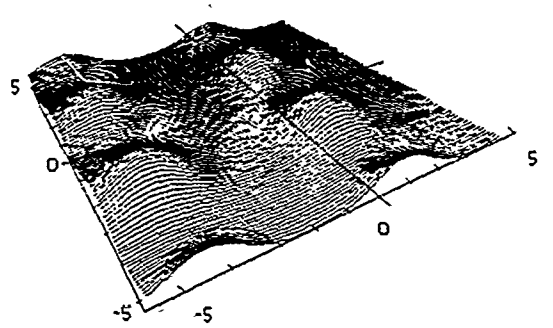


FIG. 1. Predicted STM image $I(x,y,z)$ for a relaxed perfect ZnSe(110) surface with a tip–sample distance $z = 5\ a_{\mathrm{Bohr}}$, under negative bias. Note that the Se atoms are prominent. The units in the x- and y-directions are Å.

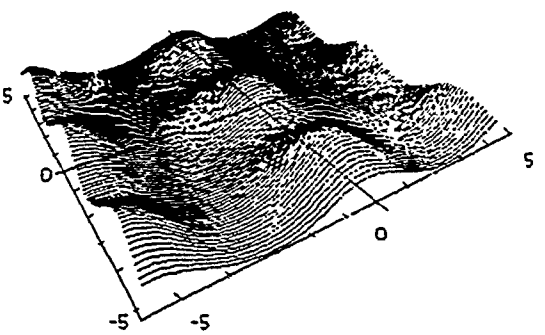


FIG. 2. Predicted STM image $I(x,y,z)$ for a relaxed perfect ZnSe(110) surface, under positive bias. Note that the Zn atoms are prominent.

 0734-211X/89/040907-03S01.00 © 1989 American Vacuum Society

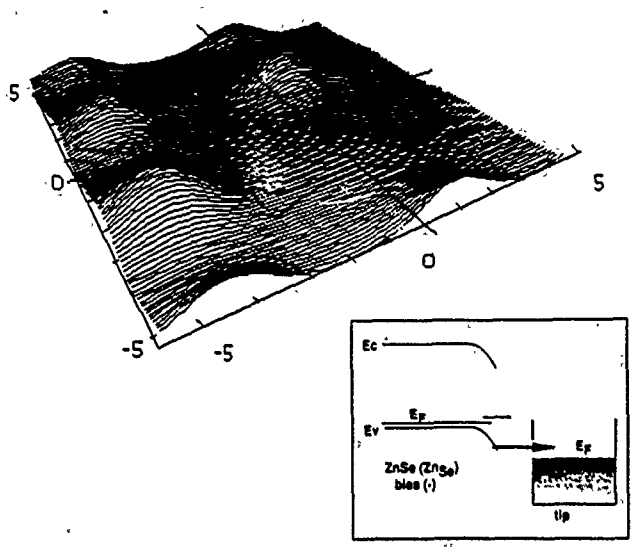


FIG. 3. Predicted STM image $I(x,y,z)$ for a relaxed ZnSe(110) surface with a Zn_{Se} defect at the cross, under negative bias such that the deep level is not in the tunneling region (see the inset). Note the depression.

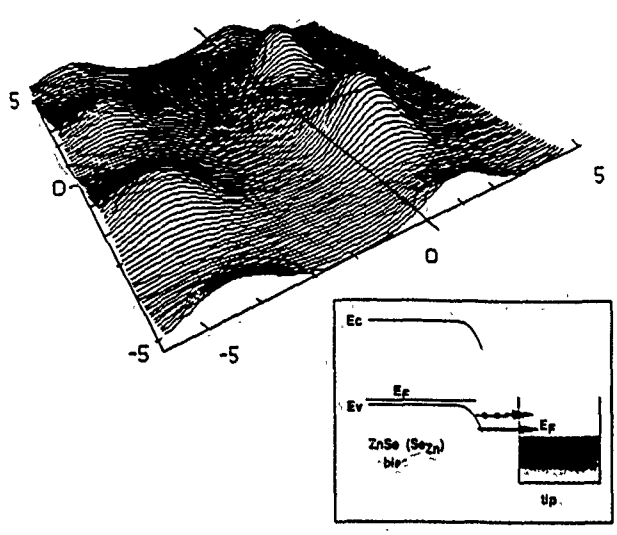


FIG. 5. Predicted STM image $I(x,y,z)$ for a relaxed ZnSe(110) surface with a Se_{Zn} defect at the origin, under negative bias such that the deep level lies in the tunneling region (see the inset). Note the two peaks at adjacent Se sites.

B. Defective surface

The images of defective surfaces depend on the bias conditions, band bending, and the Fermi energy. Here we discuss some representative images.

1. Zn_{Se}

When the Zn on Se-site antisite defect, Zn_{Se}, is present at the surface, it produces an empty deep level, capable of containing two electrons, $\simeq 1.18$ eV above the valence-band maximum at the surface. Under negative bias, if the deep level is unoccupied, electrons from the ZnSe valence band produce a tunneling current and Zn_{Se} is visible as a slight depression at its site (Fig. 3). Under positive bias such that the deep level is unoccupied but resonant with the tip's Fermi sea, two small peaks emerge at the sites adjacent to the surface defect site (Fig. 4). The qualitative features of the image are sensitive to the surface deep level structure and depend on the sign of the bias Voltage.

2. Se_{Zn}

The Se_{Zn} antisite defect behaves somewhat differently when imaged. The neutral surface defect has a doubly occupied deep level $\simeq 0.42$ eV above the valence-band edge at the surface, which, when above the tip's Fermi energy, produces a two-peaked image—with the peaks *at neighbors to the defect.* (See Fig. 5.) This happens because the neighboring Se atoms have rotated up out of the surface and are closer to the tip and because the defect's electronic state has an antibonding character, placing significant wave function amplitude

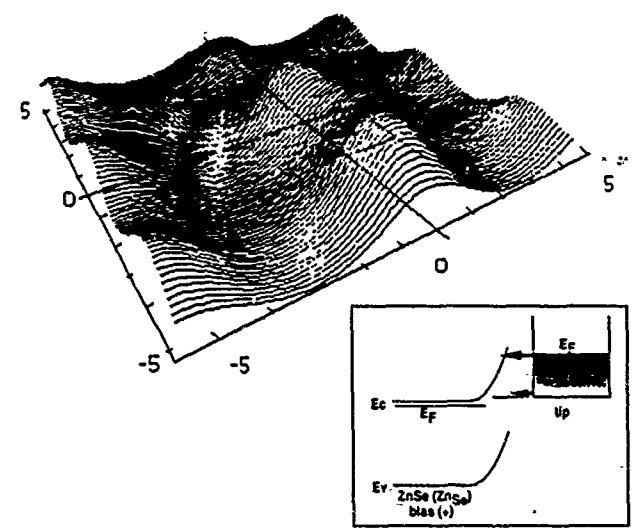


FIG. 4. Predicted STM image $I(x,y,z)$ for a relaxed ZnSe(110) surface with a Zn_{Se} defect at the cross, under positive bias such that tunneling into the deep level occurs (see the inset). Note the two peaks at adjacent Zn sites.

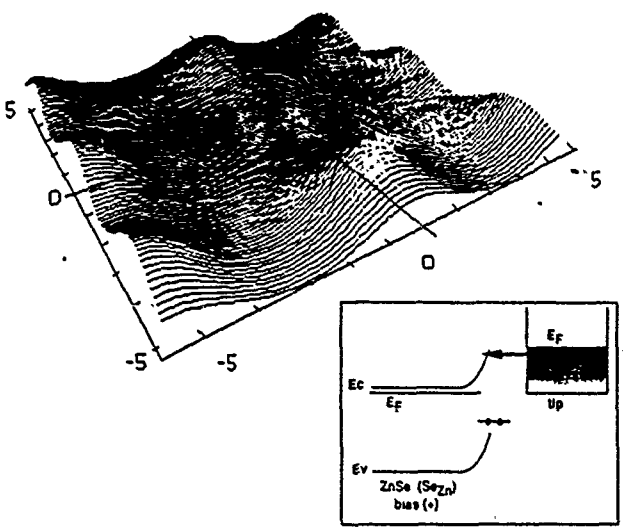


FIG. 6. Predicted STM image $I(x,y,z)$ for a relaxed ZnSe(110) surface with a Se_{Zn} defect at the origin $x = y = 0$, under positive bias such that the deep level is not in the range of tunneling energies (see the inset). Note the slight depression at the defect site.

on the neighboring atoms. Finally, a positive bias as in Fig. 6 leads to a slight depression at the defect site (Fig. 6).

III. SUMMARY

We find that the ideal STM images of antisite defects are not always bumps, but sometimes are depressions or even two-peaked structures—depending on (i) whether or not the voltage sweep causes the defect's deep levels to participate actively in the tunneling and (ii) the proximity of the defect to the tip. These facts complicate the analyses of STM data, and indicate that careful experiments combined with careful analyses are necessary. These complications are in fact advantages, however, because they will help in the identification of surface defects.

ACKNOWLEDGMENTS

We thank S. Y. Ren for stimulating discussions and we are grateful to the Defense Advanced Research Projects Agency and the Office of Naval Research (Contract Nos. N0530-0716-05 and N00014-84-K-0352) for their generous support.

[a] Permanent address.

[1] W. E. Packard, Y. Liang, N. Dai, J. D. Dow, R. Nicolaides, R. C. Jaklevic, and W. J. Kaiser, J. Microsc., Proc. Roy. Microscopical Soc. (London) 23, Suppl. (in press).
[2] R. M. Feenstra and A. P. Fein, Phys. Rev. B 32, 1394 (1985).
[3] R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. 58, 1192 (1987).
[4] J. Tersoff and D. R. Hamann, Phys. Rev. B 31, 805 (1985).
[5] P. S. Bagus, T. L. Gilbert, and C. C. J. Roothaan, J. Chem. Phys. 56, 5195 (1972).
[6] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
[7] S. Y. Ren, W. M. Hu, O. F. Sankey, and J. D. Dow, Phys. Rev. B 26, 951 (1982).
[8] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).
[9] R. E. Allen, Phys. Rev. B 20, 1454 (1979); R. E. Allen and M. Menon, Phys. Rev. B 33, 5611 (1986).
[10] D. J. Chadi, Phys. Rev. B 19, 2074 (1979).
[11] S. Y. Tong, A. R. Lubinsky, B. J. Mrstik, and M. A. Van Hove, Phys. Rev. B 17, 3303 (1978).
[12] D. J. Chadi, Phys. Rev. B 18, 1800 (1978).
[13] S. Y. Tong, W. N. Mei, and S. Xu, J. Vac. Sci. Technol. B 2, 393 (1984).
[14] M. W. Puga, G. Xu, and S. Y. Tong, Surf. Sci. 164, L789 (1985).
[15] S. Y. Ren and R. E. Allen (to be published).
[16] W.-M. Hu and J. D. Dow (to be published).

Doping $Zn_{1-x}Mn_xSe$ *n* type

Run-Di Hong and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556

(Received 10 March 1989; accepted for publication 7 April 1989)

The standard *n*-type dopant Ga is predicted to change its character from a shallow donor to a deep trap with increasing alloy composition x in $Zn_{1-x}Mn_xSe$. As a result, Ga-doped $Zn_{1-x}Mn_xSe$ alloys should be *n* type for very small x, but not for Mn-rich material.

ZnSe is a wide band-gap II-VI semiconductor that is rather easily doped *n* type,[1] but has been extremely difficult to dope *p* type until recently.[2,3] The standard *n* dopant is Ga, which occupies a cation site. In this letter we consider the Ga doping of $Zn_{1-x}Mn_xSe$ alloys and predict that there exists a critical value of the Mn composition x, x_c, such that for $x > x_c$ Ga will *not* be a donor in $Zn_{1-x}Mn_xSe$ and, hence, will not dope $Zn_{1-x}Mn_xSe$ *n* type.

Our predictions are based on the concept of a shallow–deep transition in which an s-like A_1 symmetric "deep" level of Ga lies above the conduction-band edge of ZnSe, donating its extra electron to the conduction band (see Fig. 1). With increasing alloy composition x, the conduction-band edge moves up in energy with respect to the Ga A_1 deep level and passes through it, leaving the level as a trap in the gap capable of capturing a second electron of opposite spin (and removing a carrier from the conduction band).

The theoretical predictions for the energy of the Ga deep level as a function of the Mn concentration x are given in Fig. 2, along with the conduction-band edge's dependence on x. For $x < x_c$, Ga is a donor and an *n*-type dopant, but for $x > x_c$ neutral Ga is a deep trap. The theory is basically the Hjalmarson *et al.* theory of substitutional deep impurity levels,[4,5] modified[6–8] to account for the *d* states of Mn (which do not play a significant role, because the spin-up *d* states lie well below the valence-band maximum and the spin-down states are well above the conduction-band edge[9]). The theoretical uncertainty in the predicted deep level energy for the Hjalmarson model is typically a few tenths of an electron volt, and so the predicted value of $x_c = 0.09$ may be uncertain by about 0.05. Nevertheless, the Hjalmarson model has a long history of successfully predicting chemical trends, and the qualitative dependence of the Ga level on alloy composition (Fig. 2) should certainly be observed. The major implication of the prediction is that growing Mn-rich, Ga-doped, *n*-type $Zn_{1-x}Mn_xSe$ should be very difficult, even though Ga-doped ZnSe itself is normally *n* type. Other column III dopants should exhibit this same behavior, although x_c will be slightly larger by about 0.03–0.04 for the heavier atoms In or Tl.

Such shallow–deep transitions of deep levels are well known in semiconductors and profoundly alter the electrical or optical properties of the host. Perhaps the best-studied case is *N* in $GaAs_{1-x}P_x$, which produces an isoelectronic deep level in the gap for $x > 0.2$, but moves into the conduction band for As-rich material, $x < 0.2$.[10,11] This *N* level plays a prominent role in the electroluminescence of $GaAs_{1-x}P_x$. The *DX* center in $Al_xGa_{1-x}As$ is another case. This center, which is very likely associated with Si or a similar donor,[4,12–14] descends from the conduction band of GaAs into the gap with increasing x, much as Ga does in $Zn_{1-x}Mn_xSe$. The *DX* center is technologically important because it limits the fabrication of high electron mobility transistors (HEMTs).

The experimental evidence supporting the prediction of a Ga shallow–deep transition in $Zn_{1-x}Mn_xSe$ is presently fragmentary and qualitative: Ga dopes ZnSe *n* type but Ga doping of $Zn_{1-x}Mn_xSe$, even with only modest amounts of Mn, has not yielded a significant number of carriers.[15] Systematic studies of $Zn_{1-x}Mn_xSe$ alloys are needed to test the picture presented here.

One possible way to test the theory is to apply hydrostatic pressure to Ga-doped $Zn_{1-x}Mn_xSe$ and monitor the carrier density. For pressures of 30 kbar, the conduction-band edge should move[16,17] above the Ga deep level for $x > 0.07$, causing an abrupt drop in the number of carriers.

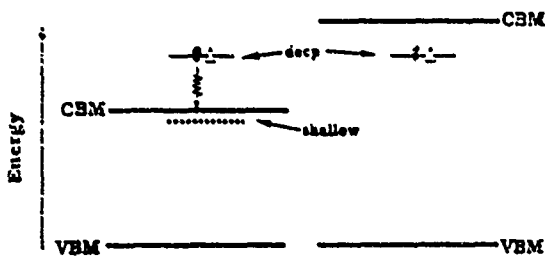


FIG. 1. Schematic illustration of the shallow–deep transition of Ga in $Zn_{1-x}Mn_xSe$. (a) For $x = 0$, the Ga A_1 symmetric deep level is resonant with the conduction band. The extra electron associated with this column III impurity on a column II site spills out of the level and falls to the conduction-band edge, ionizing the Ga. The Coulomb potential of the ionized Ga then binds the electron in a hydrogenic shallow donor state (at zero temperature). (b) For $x > x_c$, the deep state lies in the fundamental band gap with the conduction-band edge at higher energy, and the neutral Ga deep level contains one electron (circle) and one hole (triangle), and can trap an electron of opposite spin, depleting a carrier from the conduction band. The energy scale of this diagram is distorted for illustrative purposes.

 0003-6951/89/252597-02$01.00 © 1989 American Institute of Physics

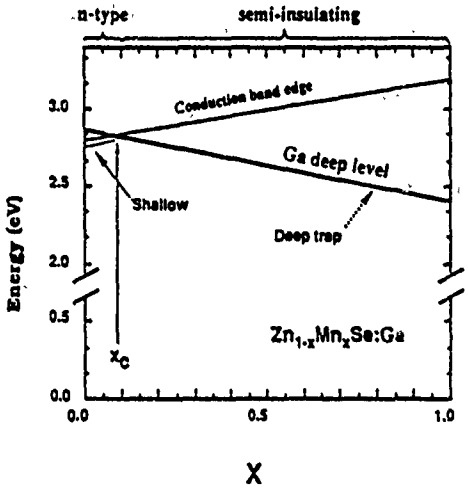


FIG. 2. Predicted energies (in electron volts) of the conduction-band edge (light solid line) and the cation-site Ga deep level (heavy) vs alloy composition x in $Zn_{1-x}Mn_xSe$. For $x<x_c$ the Ga deep level is autoionized and the extra electron is trapped in the 1s hydrogenic shallow level (dashed) at zero temperature. Thermal excitation of this level causes the material to be n type. For $x>x_c$, Ga is a deep trap and tends to make $Zn_{1-x}Mn_xSe$ semi-insulating.

We believe that shallow–deep transitions such as the one described here play a significant role in the doping problems experienced by II-VI semiconductors. Subsequent work will deal with the complexities of p doping and propose schemes, based on the idea of shallow–deep transitions, for achieving p-type conductivity in ZnSe. We hope that the picture we have presented here will be tested experimentally.

We are grateful to the Defense Advanced Research Projects Agency (contract No. N0530-0716-09), and the U. S. Office of Naval Research (contract No. N00014-84-K-0352) for their generous support. We also thank R. Gunshor for stimulating discussions of the doping problem.

[1]D. C. Reynolds and T. C. Collins, Phys. Rev. 188, 1267 (1969); J. L. Merz, K. Nassau, and J. W. Shiever, Phys. Rev. B 8, 1444 (1973); P. J. Dean, W. Stutius, G. F. Neumark, B. J. Fitzpatrick, and R. N. Bhargava, Phys. Rev. B 27, 2419 (1983).

[2]T. Yasuda, I. Mitsuishi, and H. Kukimoto, Appl. Phys. Lett. 52, 57 (1988), and references therein.

[3]J. Nishijawa, K. Itoh, Y. Okuno, and F. Sakurai, J. Appl. Phys. 57, 2210 (1985).

[4]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).

[5]For a review, see J. D. Dow, in *Highlights of Condensed-Matter Theory* (Proceedings of the International School of Physics "Enrico Fermi," Course 89, Varenna, 1983), edited by F. Bassani, F. Fumi, and M. P. Tosi (Societa Italiana di Fisica, Bologna, Italy, and North-Holland, Amsterdam, 1985), pp. 465–494.

[6]The tight-binding sp^3s^* Hamiltonian for ZnSe (see Ref. 7) was modified as follows: The on-site s and p matrix elements of Mn were obtained from the Zn matrix elements by assuming they are proportional to differences in atomic energies, with the proportionality constants β_l taken from Ref. 7. Off-diagonal matrix elements were scaled as the inverse square of the bond length. The d states were added but not coupled to the s and p states because the filled d states lie well below the valence-band maximum, while the empty states are well above the conduction edge and hence, do not play a dominant role in the electronic structure.

[7]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[8]The spin-orbit interaction is included as done by A. Kobayashi, O. F. Sankey, and J. D. Dow, Phys. Rev. B 25, 6367 (1982). For ZnSe we have $\lambda_v = 0.011$ and $\lambda_c = 0.007$, while for MnSe we have $\lambda_v = 0.345$ and $\lambda_c = 0.013$.

[9]The d states of $Cd_{0.8}Mn_{0.2}Te$ and $Zn_{0.8}Mn_{0.2}Se$ are known to lie ~3.5 eV below the valance-band maximum and 3 eV above the the valence-band minimum in $Cd_{0.8}Mn_{0.2}Te$: A. Franciosi, S. Chang, C. Caprile, R. Reifenberger, and V. Debska, J. Vac. Sci. Technol. A 3, 926 (1985); A. Franciosi, A. Wall, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, R. V. Kasowski, R. Reifenberger, and F. Pool, *d-states, Exchange Splittings and Mn Electronic Configuration in $Cd_{1-x}Mn_xTe$* (to be published); A. Wall, A. Franciosi, Y. Gao, J. H. Weaver, M.-H. Tsai, J. D. Dow, and R. V. Kasowski, J. Vac. Sci. Technol. A 7, 656 (1989). A similar situation is expected to hold for MnSe.

[10]D. J. Wolford, B. G. Streetman, W. Y. Hsu, J. D. Dow, R. J. Nelson, and N. Holonyak, Jr., Phys. Rev. Lett. 36, 1400 (1976); W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977); D. J. Wolford, B. G. Streetman, W. Y. Hsu, and J. D. Dow, *Proc. 13th International Conference on Physics of Semiconductors*, edited by F. G. Fumi (Tipografia, Marves, Rome, Italy, 1976), p. 1049.

[11]D. J. Wolford, W. Y. Hsu, J. D. Dow, and B. G. Streetman, J. Lumines. 18/19, 863 (1979).

[12]H. P. Hjalmarson and T. J. Drummond, Phys. Rev. Lett. 60, 2410 (1988), Appl. Phys. Lett. 48, 656 (1986).

[13]S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B 38, 10677 (1988).

[14]D. J. Chadi and K. J. Chang, Phys. Rev. Lett. 61, 873 (1988).

[15]R. Gunshor (private communication).

[16]Y. F. Tsay, S. S. Mitra, and B. Bendow, Phys. Rev. B 10, 1476 (1974).

[17]R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Phys. Rev. B 38, 12549 (1988).

Thin quantum-well superlattices of GaAs and $(GaAs)_{1-x}(ZnSe)_x$ with ZnSe: Possibility of band gaps in the blue-green

Jun Shen and John D. Dow
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556
Shang Yuan Ren
Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556 and Department of Physics, University of Science and Technology of China, Hefei, Anhui, People's Republic of China

(Received 13 March 1989; accepted for publication 12 July 1989)

It is predicted that thin quantum-well superlattices or spike superlattices of GaAs in ZnSe will produce band gaps in the yellow-green, and that $(GaAs)_{1-x}(ZnSe)_x$ spikes will lead to green and blue-green gaps. These thin quantum-well structures should have better doping properties than ZnSe for $x < 0.6$.

I. INTRODUCTION

Diffraction-limited optical storage on compact disks has stimulated the quest for light emitters with short wavelengths in the green-to-blue portion of the spectrum. The semiconducting materials most often considered for such applications are II-VI compound semiconductors[1] and *N*-based III-V semiconductors, such as $In_{1-x}Ga_xN$.[2] The *N*-based III-V's are difficult to grow and dope, and have undesirable large shallow acceptor binding energies, whereas the II-VI's generally suffer from the "doping problem"; it has been very difficult to dope most II-VI's *p* type, with the major exception being ZnTe, which resists *n* doping. While there have been some recent advances on the II-VI doping problem,[3,4] one cannot avoid wishing that a well-understood III-V semiconductor such as GaAs had a band gap in the green or blue, a material that is dopable both *n* and *p* type, and integratable in *p-n* junctions and artificial microstructures. The purpose of this paper is to predict that thin GaAs layers or "spikes" embedded in ZnSe should have band gaps in the yellow-green, and that interdiffusion of the GaAs spikes with the ZnSe should increase the band gap to the blue-green. Furthermore, we shall argue that the $(GaAs)_{1-x}(ZnSe)_x$ spikes for $x < 0.6$ will be largely free of the doping problems that have traditionally limited ZnSe and other II-VI semiconductors.

II. BAND

Our approach is based on the theory of the electronic structure of superlattices and deep levels in superlattices.[5–8] We consider $N_{spike}N_{ZnSe}[001](GaAs)_{1-x}(ZnSe_x)/ZnSe$ superlattices where N_{spike} is small, typically less than 10, and N_{ZnSe} is large, typically 10 or greater. (A 1×10 GaAs/ZnSe superlattice consists of one two-atom-thick layer of GaAs, alternating with 10 layers of ZnSe.) The $(GaAs)_{1-x}(ZnSe)_x$ is treated in the virtual crystal approximation.[9] These superlattices are, for all intents and purposes, lattice matched; hence, their theoretical treatment is the same as for $GaAs/Al_xGa_{1-x}As$ superlattices.[5–10] The GaAs/ZnSe valence-band offset has been measured to be 0.96 eV,[11] and is assumed to be a linear function of alloy composition *x*. The parameters of the superlattice Hamiltonian are known for GaAs and ZnSe,[12] but we update the ZnSe parameters here to account for the determination of the (low-temperature) ZnSe band gap at Γ of 2.8 eV.[13,14] The resulting band-gap predictions for selected values of N_{spike} and *x*, with $N_{ZnSe} = 10$, are given in Fig. 1. The gaps are direct, with the superlattice conduction-band minimum at the Γ point of the mini-Brillouin zone; the quantum well is Type I: The superlattice conduction- and valence-band edges lie at energies within the fundamental band gap of ZnSe. Hence, the spikes should meet the conditions for luminescence: (i) the carriers will thermalize to the $(GaAs)_{1-x}(ZnSe)_x$ layers of the material; and (ii) electrons and holes will have the same wave vectors and therefore satisfy the crystal-momentum selection rule for recombination.

Having a direct band gap is a necessary, but not sufficient, condition for a semiconductive light-emitting diode material. In addition, the semiconductor must be dopable both *n* and *p* type. This is not a problem in bulk GaAs, but is in ZnSe, which is difficult to dope *p* type. The cause of the *p* doping problem in ZnSe is thought to be a distribution of deep hole traps within ~0.3 eV of the valence-band maximum.[15] Therefore, if our spike superlattices are to be doped *p* type, those deep levels must be covered up by the superlattice valence-band so that they cease trapping holes. To achieve this, the superlattice must have its valence band maximum at least 0.3 eV above the valence maximum of ZnSe (see Fig. 2). This implies normally that *x* must be less than 0.5 and the wells must have widths corresponding to $N_{spike} = 2$ or more diatomic layers.

The energy band lineups are displayed in Fig. 3 for 1×10 GaAs/ZnSe. Note that most of the band offset is in the valence band, and that the energy of the conduction band edge, when measured relative to the ZnSe valence-band edge (Fig. 2), is relatively independent of layer thickness and alloy composition. Therefore, by altering the alloy composition of the $(GaAs)_{1-x}(ZnSe)_x$ we are, in effect, tuning the absolute valence-band energy. With this approach we can cover up deep levels that might otherwise trap holes and adversely affect the *p* doping of the material. Once covered, any holes in the deep levels are autoionized and become carriers. This method for eliminating deep traps has been proposed for doping ZnSe *p* type, and explains the relative ease with which ZnTe can be doped *p* type (in contrast with most

 0021-8979/90/083761-03$03.00 © 1990 American Institute of Physics

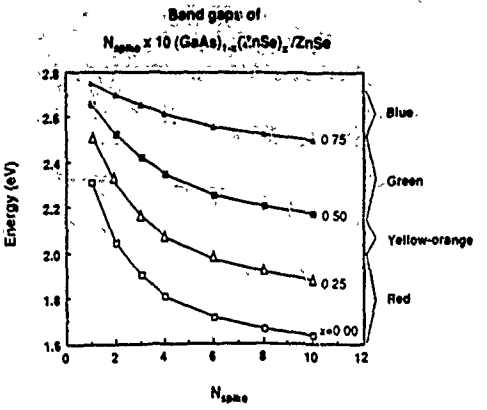


FIG. 1. Predicted band gaps (in eV) of $N_{spike} \times 10$ $(GaAs)_{1-x}(ZnSe)_x$/ZnSe superlattices vs N_{spike}. Gaps for $x = 0$, 0.25, 0.5, and 0.75 are denoted by open squares, open triangles, closed squares, and closed triangles, respectively.

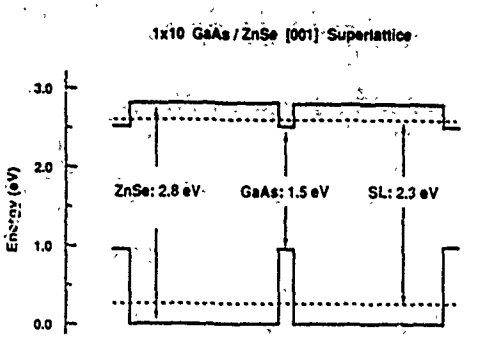


FIG. 3. Energy-band alignments for a 1×10 GaAs/ZnSe [001] superlattice. The band gaps of ZnSe, GaAs, and the superlattice (SL) are 2.8, 1.55, and 2.32 eV, respectively. The superlattice band edges are denoted by dashed lines. The zero of energy is the valence-band maximum of bulk ZnSe.

other II-VI semiconductors).[15]

The ZnSe, when interdiffused with the GaAs, will provide GaAs with both donors and acceptors, and in the ideal limit will yield perfectly compensated material. By carefully controlling the growth conditions, it should be possible to dope the spike either *n* type or *p* type with Se or Zn, which are shallow acceptors and donors in GaAs. It should even be possible, with masking, to fabricate a *p-n* junction in a single spike.

Impurities that normally do not adversely affect GaAs can be problems in spike GaAs. Because the GaAs band gap has been widened so much, impurities that were shallow donors or acceptors in GaAs might become deep traps in the spike superlattices. Such shallow-deep transitions have been observed in III-V alloys. For example, an anion site *N* impurity in $GaAs_{1-x}P_x$ produces a deep level resonant with the conduction band in GaAs, which becomes uncovered as the gap widens in those alloys,[16] so that *N* becomes a deep trap. Similar behavior has been observed for the donor Si_{Ga} (and the related *DX* center) in $Al_xGa_{1-x}As$.[17] However, we do not find such behavior for Si or *N* in our spike superlattices, mostly because the conduction-band edges in these superlattices are so independent of layer thickness and alloy composition. Nevertheless, some levels near the valence-band maximum do undergo such a shallow-deep transition. For example, Zn_{anion} is a deep trap in the quantum well of a 1×10 GaAs/ZnSe superlattice (see Fig. 4) but lies in the valence band of bulk ZnSe. Finally, the native point defects, the antisite defects Se_{Ga}, Zn_{As}, and As_{Ga}, and the vacancies Va_{Ga} and Va_{As} can produce deep levels in the superlattice gap, as shown in Fig. 4, and control of their concentrations may be required.

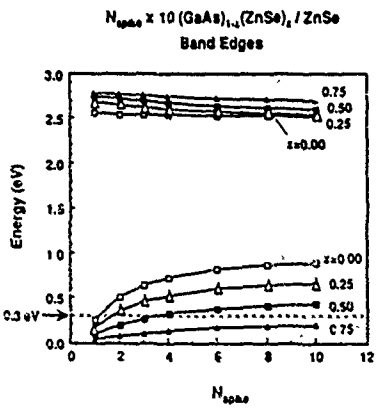


FIG. 2. Predicted conduction- and valence-band edges (in eV) for $N_{spike} \times 10(GaAs)_{1-x}(ZnSe)_x$/ZnSe superlattices, for various *x*, using the notation of Fig. 1. Note that the major variations are in the valence band. The zero of energy is the valence-band maximum of bulk ZnSe. The materials with valence-band maximum above 0.3 eV should not be plagued by the *p* doping problems experienced by ZnSe (see Ref. 15).

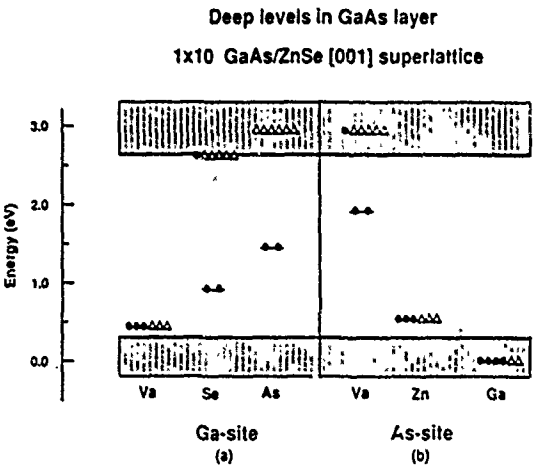


FIG. 4. Native defect deep levels in or near the band gap of 1×10 GaAs/ZnSe for (a) the cation site and (b) the anion site. Va denotes a vacancy. Open triangles (closed circles) denote holes (electrons) occupying the deep levels for a neutral defect. For the As-site vacancy, the electron in the deep level above the conduction band edge is autoionized and falls to the conduction-band edge, becoming a shallow donor electron. Similarly, the Ga_{As} holes bubble up to the valence-band maximum. The zero of energy is the valence-band maximum of bulk ZnSe.

III. SUMMARY

The theory presented here should be quite reliable. A similar theory of 1×1 GaAs/AlAs [001] superlattices predicted a band gap of 2.1 eV, in comparison with the experimental value of 2 eV.[18] Thus, the band-gap predictions are likely to have an uncertainty only of order 0.1 eV.

Another theoretical study of GaAs/ZnSe [001] superlattices, based on our electronic structure model[12] (but without the new parameters[14]) has been presented,[19] with similar results to ours for $x = 0$. However, that theory was directed toward the core exciton problem rather than short wavelength light emission, and did not consider the possibility of modifying the GaAs in order to achieve blue-green luminescence.

We hope that these theoretical predictions will stimulate efforts to grow ZnSe with GaAs spike quantum wells. While such growth will require great skill, it should be feasible.

Finally, we note that even larger band gaps should be achievable with $(GaAs)_{1-x}(ZnSe)_x$ spikes in $Zn_{1-y}Mn_ySe$.[20]

ACKNOWLEDGMENTS

We are grateful to the U. S. Office of Naval Research (Contract No. N00014-84-K-0352) and the Defense Advanced Research Projects Agency (Contract No. N0530-0716-09) for their generous support.

[1]D. C. Reynolds and T. C. Collins, Phys. Rev. 188, 1267 (1969); J. L. Merz, K. Nassau, and J. W. Shiever, Phys. Rev. B 8, 1444 (1973); P. J. Dean, W. Stutius, G. F. Neumark, B. J. Fitzpatrick, and R. N. Bhargava, Phys. Rev. B 27, 2419 (1983).

[2]D. W. Jenkins and J. D. Dow, Phys. Rev. B 39, 3317 (1989), and references therein.

[3]T. Yasuda, I. Mitsuishi, and H. Kukimoto, Appl. Phys. Lett. 52, 57 (1988), and references therein.

[4]J. Nishizawa, K. Itoh, Y. Okuno, and F. Sakurai, J. Appl. Phys. 57, 2210 (1985).

[5]S. Y. Ren, J. D. Dow, and J. Shen, Phys. Rev. B 38, 10677 (1988).

[6]S. Y. Ren and J. D. Dow, J. Appl. Phys. 65, 1987 (1989).

[7]R.-D. Hong, D. W. Jenkins, S. Y. Ren, and J. D. Dow, Proc. Mater. Res. Soc. 77, 545 (1987).

[8]J. D. Dow, S. Y. Ren, and J. Shen, "Deep impurity levels in semiconductors, semiconductor alloys, and superlattices," NATO Advanced Science Institutes Series B 183: *Properties of Impurity States in Superlattice Semiconductors*, edited by C. Y. Fong, P. Batra, and S. Ciraci (Plenum Press, New York, 1988), pp. 175–187.

[9]H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980); see also, W. Y. Hsu, J. D. Dow, D. J. Wolford, and B. G. Streetman, Phys. Rev. B 16, 1597 (1977).

[10]Because the superlattice does not have a common anion, we use a slightly different treatment of interfacial off-diagonal matrix elements from that in Ref. 5 when the interface is between Zn and $(As_{1-x}Se_x)$ or between Se and $(Ga_{1-x}Zn_x)$: The off-diagonal matrix elements are taken to be the averages $[(1+x)M(ZnSe) + (1-x)M(GaAs)]/2$, where M(ZnSe) indicates the relevant matrix element for bulk ZnSe.

[11]S. P. Kowalczyk, E. A. Kraut, J. R. Waldrop, and R. W. Grant, J. Vac. Sci. Technol. 21, 482 (1982). Our results are not very sensitive to the precise value of the band offset: If the measured value of 0.96 eV were to be increased by a small value ϵ, then the band gap calculated in this paper for a 1×10 GaAs/ZnSe superlattice would decrease by 0.02 ϵ. Similar changes are to be expected in the band gaps of superlattices with other periods and compositions.

[12]P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44, 365 (1983).

[13]D. Theis, Phys. Status Solidi B 79, 125 (1977).

[14]In the notation of Ref. 12, the revised parameter is $E(s,c) = +0.1602$ eV.

[15]J. D. Dow, R.-D. Hong, S. Klemm, S. Y. Ren, M.-H. Tsai, O. F. Sankey, and R. V. Kasowski (to be published).

[16]D. J. Wolford, W. Y. Hsu, J. D. Dow, and B. G. Streetman, J. Lumin. 18/19, 863 (1979).

[17]H. P. Hjalmarson and T. J. Drummond, Appl. Phys. Lett. 48, 657 (1986).

[18]A. C. Gossard, P. M. Petroff, W. Weigmann, R. Dingle, and A. Savage, Appl. Phys. Lett. 29, 323 (1976).

[19]D.-L. Shen, K.-M. Zhang, and R.-L. Fu, Appl. Phys. Lett. 53, 500 (1988).

[20]J. D. Dow, J. Shen, and S. Y. Ren (to be published).

Self-consistent antiferromagnetic ground state for La_2CuO_4 via energy-band theory

Robert V. Kasowski
E. I. du Pont de Nemours and Company, Central Research and Development Department, Experimental Station, Wilmington, Delaware 19898

M.-H. Tsai and John D. Dow
Physics Department, University of Notre Dame, Notre Dame, Indiana 46556
(Received 18 November 1988; revised manuscript received 29 December 1989)

We have used the pseudofunction method to compute self-consistent spin-polarized energy bands for La_2CuO_4. The ground state is found to be semiconducting and antiferromagnetic (AF) with a moment of $0.35\mu_B$. The net moment resulting from the occupation of 82 itinerant bands is localized on the Cu atoms only. Our results agree with experiments in which La_2CuO_4 is found to be semiconducting and AF with a moment on the Cu atoms of $(0.35\pm0.05)\mu_B$ (neutron scattering) or in the range $0.2\mu_B$ to $0.6\mu_B$ (muon-spin resonance). Sr is found to produce holes on the out of plane O contrary to the assumption used in many strong-correlation theories of superconductivity.

The pseudofunction (PSF) method[1] with a local-density potential[2] has been successfully used to compute spin-polarized energy bands for Ni and MnTe.[3] The good agreement with experiment encouraged the use of this method for the new superconducting materials such as Sr-doped La_2CuO_4, hereby denoted in short as (2:1:4). In this paper, we present self-consistent spin-polarized local-density energy-band results for pure and Sr-doped 2:1:4. The ground-state energy bands are semiconducting and antiferromagnetic (AF) with a moment of $0.35\mu_B$ on the Cu atoms. One Cu-O hybrid state is split off by the AF correlation to give a semiconductor. The moment is nonintegral because of the strong hybridization between O and Cu at the Fermi surface. A magnetic ground state is found to be stable even in the metallic state when Sr is substituted for La. For example, an AF state with a moment of $0.09\mu_B$ is obtained when half of the La atoms are replaced by Sr atoms. Thus, contrary to widespread belief, band theory quantitatively accounts for the magnetic properties of 2:1:4 and suggests that the magnetic state is not observed in doped samples as the disorder overcomes the magnetic order.

The spin-polarized band model with a local-density potential has been used by several investigators to attempt an understanding of the AF and semiconducting behavior of 2:1:4.[4] These spin-polarized calculations with other methods find the ground state to be nonmagnetic and metallic contrary to experiment. We believe the reason that these attempts did not yield the correct ground state is that the bands at E_f are too broad. Specifically, for nonmagnetic calculations, the 2:1:4 bands are doubly degenerate on the hexagonal face of the orthorhombic unit cell. Semiconductivity can be obtained only when this degeneracy is lifted. These bands are only 0.17-eV wide for the PSF method.[5] The other techniques find these bands to be approximately 0.5-eV wide. Thus, the magnetic correlation needed to split a 0.5-eV-wide band is far stronger than that needed for a 0.17-eV band. Finally, we show that the nonspherical part of the crystal potential need be expanded in at least 15 625 plane waves to obtain the AF ground state. We believe this to be the key to the differences with other methods.

There are many proposals of superconductivity based on the coupling via magnetic excitations both local and nonlocal. Such theories of superconductivity are well documented[6] in the literature and will not be referenced here. We will only point out that we find the d_{z^2} band to be at the Fermi energy as postulated by Lee and Ihm.[7] They accounted for superconductivity with a two-band model (broad band and narrow d_{z^2} band) which we find to be appropriate. A more detailed two-band model[8] using local densities of states from the band theory for the broad Cu-O band and a Cu-O band in the z direction can phenomenologically account for the magnetic and superconducting properties of pure and doped 2:1:4.

The spin-polarized PSF method has been tested for ferromagnetic Ni and antiferromagnetic MnTe.[3] The results for Ni agree with those in the literature. For MnTe, we find the antiferromagnetic state to be more stable than the ferromagnetic state in agreement with experiment. Calculations for MnTe with the linearized augmented plane-wave (LAPW) method[9] also find that the antiferromagnetic state is more stable than the ferromagnetic state. K on Si is another example of success for the PSF method. The PSF method predicted the K-Si bond length to be 3.3 Å (Ref. 10) whereas the pseudopotential method predicted the bond length to be 2.59 Å.[11] Subsequent, surface-extended x-ray-absorption fine-structure (SEXAFS) experiments[12] give a bond length of 3.15 Å which is very similar to the PSF result.

The PSF method has a local-orbital basis set employing the full potential including the core states (see the Appendix for details). The computation is done in Fourier space for computational efficiency just as the time-consuming parts of matrix element evaluation is done with the fast Fourier transform. For 2:1:4 the potential

is expanded in 15 625 plane waves. The very accurate expansion of the potential is essential in obtaining the 0.17-eV-wide band at the Fermi energy. A less accurate potential gives broader bands. Nine s, p, and d basis functions were used on the La and Cu atoms with s,p sets on the O atoms. The basis functions are continually charged during iteration to self-consistency in order to optimize the description of the charge density. Identical results are obtained whether or not the O atoms were started with a spin state. The calculation was first brought to self-consistency with only a single K point in the iteration. A four-point sample was next used. The final iterations were done with eight k points. The average deviation in the potential between the input and output potential at 10% mixing was 1.1 mRy, with a maximum deviation of 7 mRy. The moment of the Cu sites increased by about $0.02\mu_B$ on going from a four-point sample to an eight-point sample. Subsequently, the calculations have been redone in entirety with a general set of eight k points. The moment is unchanged to $\pm 0.01\mu_B$ and the bands are virtually identical.

In Fig. 1, the spin-polarized bands near E_f are plotted for 2:1:4. The band gap between filled and empty states varies from approximately 2 eV at Γ to 0.06 eV on the hexagonal face. The bands are relatively flat in the Z direction because of the layered structure and are not shown. For the hexagonal face, the bands are narrow and doubly degenerate for a self-consistent field (SCF) non-spin-polarized calculations.[5] For spin-polarized bands, the gap varies from 0.35 eV at X and M to 0.06 along the M to N and N to S directions. The necessity that bands be narrow on this hexagonal face thus becomes obvious since the spin effect which gives the semiconducting gap varies from 0.35 to 0.06 eV. This band is 0.17-eV wide for the PSF method[5] and approximately 0.5 eV for other methods.[4] At internal points of the Bril-

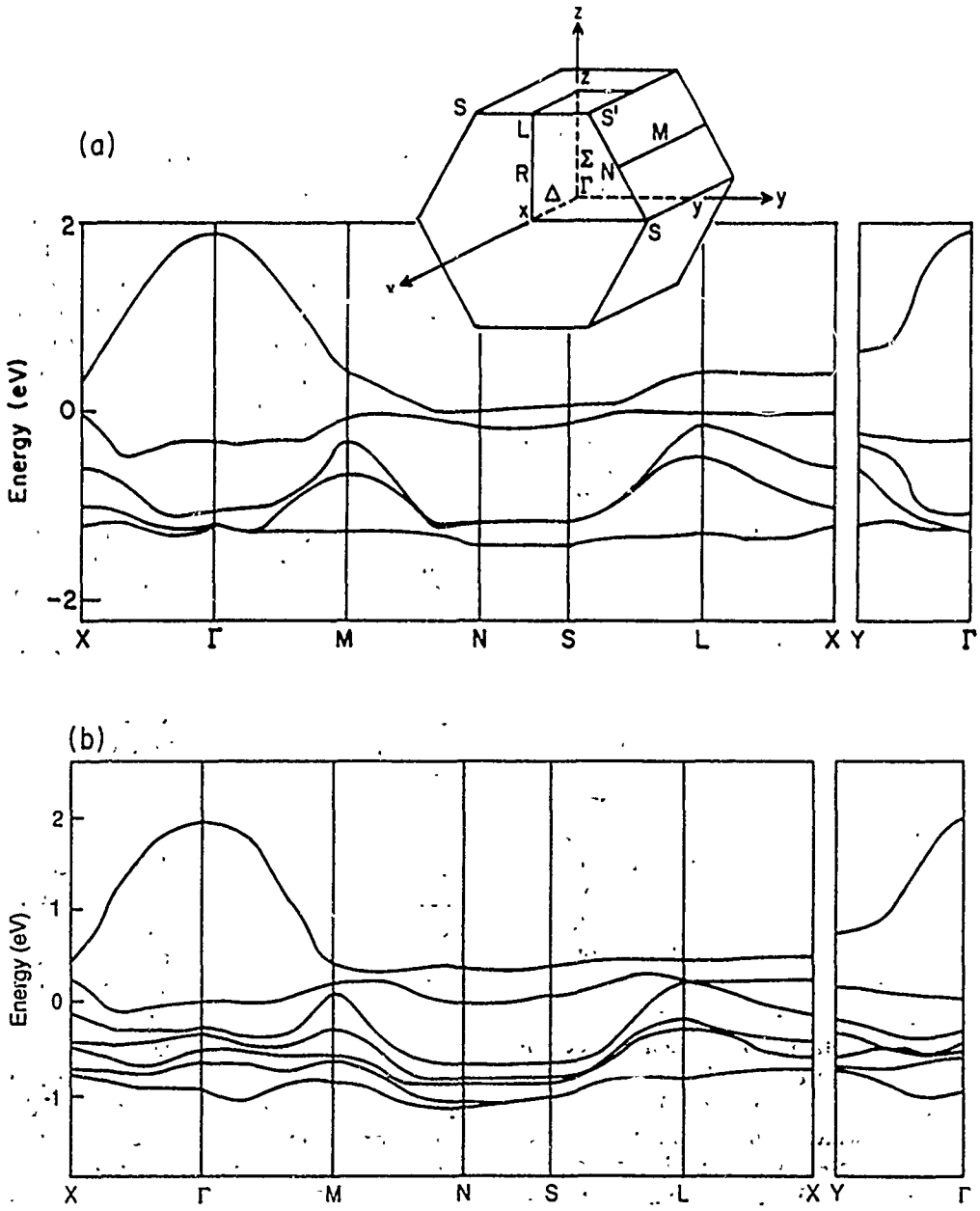


FIG. 1. Spin-polarized energy bands near the Fermi energy for (a) La_2CuO_4 and (b) $LaSrCuO_4$.

louin zone, the gap is large. This gap existed even for non-spin-polarized energy bands. Thus, the spin-polarization effect is only opening a gap at a small percentage of the points within the Brillouin zone. An approximate measure of the number of points at which antiferromagnetism opens a gap is a skin depth on the hexagonal face divided by the volume which would be a small number. This is probably a unique feature of the layered Cu-O materials because the number of electrons is important.

The indirect gap in Fig. 1(a) is as little as 0.01 eV. We were concerned that if the indirect gap were less than zero and a metallic state was obtained, the moment on the Cu atoms might become zero. Thus, the bands were iterated to SCF with an extremely poor sampling of k space using the points S, L, N, M, and X. The bands remain essentially unchanged except that the bands at L overlap the bands at N by 0.03 eV to give a metal. The moment on the Cu atoms increases to $0.39\mu_B$. This result indicates that our moment is not sensitive to small changes in band structure which produce metallic overlap. Alternatively, we believe nesting does not have to be exact to obtain the magnetic ground state as would occur at small Sr-dopant concentrations. Since band theory cannot be performed for small deviations from stoichiometry, we suggest that the above computations for a metallic state suggest that the antiferromagnetic state is stable for small dopant concentrations. Small dopant concentrations of Sr probably would not yield a metallic state because of the narrow bands at E_f and small overlap between carriers. Disorder would also create a mobility gap at low carrier concentrations.

It is important to understand the magnetic behavior as a function of Sr doping. The optimal superconducting transition temperature is observed at the large doping level of 0.15 Sr. Calculations have been performed on 2:1:4 where one-fourth and one-half of the La atoms are substituted by Sr. The dependence of the magnetic moment as a function of Sr is thus obtained.

The AF moment is obtained by integrating over all spheres. A value of $0.35\mu_B$ is obtained. An important understanding can be obtained by computing the charge within spheres about the different sites. In Table I, the amount of charge within the Cu sphere (R=0.98 Å), La sphere (R=1.49 Å), and O sphere (R=0.86 Å) is tabulated for spin-up and spin-down bands for both the pure and doped 2:1:4. The O atoms are labeled according to the type of coplanar metal atom. For 2:1:4, the first Cu atom has a net spin down of 0.35 electrons (4.33−4.69). The O atoms are identical in spin density indicating that the spin is localized to the Cu atom. Even though the band picture is an itinerant picture, moments localized to the metal atoms are obtained. For 25% Sr doping, the moments on the Cu are $0.35\mu_B$ and $0.32\mu_B$, which is indicative of small ferrimagnetism. At 50% doping, moments of $0.09\mu_B$ are obtained on the Cu atoms with antiferromagnetic behavior. The O atoms within the Cu-O basal plane do not have a moment at any level of Sr doping. For 25% Sr, the O atoms adjacent to the Sr and Cu(2) develops a small moment of $0.06\mu_B$. Thus, the moment on the Cu atoms will remain at a large value for Sr doping up to 25% with a slow decrease at higher levels. The O atoms may also participate depending on the doping level.

Spin-polarized energy bands calculated for $LaSrCuO_4$ are shown in Fig. 1(b). The bands are quite similar in appearance to those for pure 2:1:4. Because one-half of a band has been depopulated, the Fermi energy falls in the middle of the Cu-O band associated with the d_{z^2} orbital and O perpendicular to the plane.

Evidence for both spin and charge fluctuations via the apex O can be found in Table I. As the charge density and net spin on the O atoms in the base are independent of Sr content, we conclude that the basal O atoms do not participate in fluctuations. First, we consider fluctuations for 25% doping since this is closer to experiment at 0.15 Sr. Pure 2:1:4 has a charge of 9.03 (4.34+4.69) on the Cu atoms and 2.62 on all O. For 25% Sr, the net charge on Cu atoms is reduced to 8.98 electrons, an increase in Cu valence of 0.05 electrons. The O(La) and O(Sr) have moments of $0.06\mu_B$ by reducing its spin-up character by 0.06 e. One could envision fluctuations among the O(La) and O(Sr) atoms via bands at E_f that Sr has caused to be depopulated. Fluctuations require empty states near E_f. Table I shows these states to exist on the out of plane O.

TABLE I. Partial charge densities. The arrow indicates the spin direction.

	La_2CuO_4			$La_{1.5}Sr_{0.5}CuO_4$			$LaSrCuO_4$	
	↑	↓		↑	↓		↑	↓
Cu_1	4.34	4.69	Cu_1	4.31	4.67	Cu_1	4.33	4.42
Cu_2	4.69	4.34	Cu_2	4.66	4.34	Cu_2	4.66	4.57
La	0.46	0.46	La	0.46	0.46	La	0.45	0.45
La	0.46	0.46	La	0.44	0.44	La	0.45	0.45
La	0.46	0.46	La	0.45	0.45	La	0.33	0.33
La	0.46	0.46	Sr	0.34	0.34	Sr	0.33	0.33
O (base)	2.62	2.62	O (base)	2.61	2.62	O (base)	2.61	2.61
O (base)	2.62	2.62	O (base)	2.61	2.62	O (base)	2.61	2.61
O (La)	2.62	2.62	O (La)	2.56	2.62	O (Sr)	2.49	2.49
O (La)	2.62	2.62	O (Sr)	2.56	2.62	O (Sr)	2.49	2.49
O (La)	2.62	2.62	O (La)	2.61	2.57	O (La)	2.64	2.64
O (La)	2.62	2.62	O (La)	2.62	2.60	O (La)	2.64	2.64

A similar picture is obtained for 50% Sr doping. The net charge on the Cu sites is 8.75 (4.33+4.42) on Cu(1) and 9.23 (4.66+4.57) on Cu(2) as compared to 9.03 (4.34+4.69). Another important feature of Table I is that the O that is the nearest neighbor of Sr losses 0.13 electrons (2.62−2.49) within its sphere. This O is also bonded to the Cu that loses 0.22 (8.98−8.75) electrons within its spheres. Thus, the Sr doping, at least at this level, causes a hole to form on the Cu-O bond that is the nearest neighbor to the Sr atom. Similar but less dramatic results occur for 25% doping. This simply reflects that the band at the Fermi energy is composed mostly of the Cu-O character in the z direction, not the planar directions.

The difference between the 25% and 50% behavior is related to how the Sr is substituted for La. For 25%, the Sr is substituted only on one side of the Cu-O plane. For 50%, the Sr is placed symmetrically above and below the plane.

For very low levels of doping, a rigid-band model indicates that only bands on the hexagonal face would be depopulated and these have very important contributions from Cu-O bonds within the plane of Cu-O. Thus, at low levels of doping, states on the hexagonal face are depopulated and holes would be observed in the Cu-O plane. At large dopant levels of 15%, we would predict that the holes are in the Cu-O bond perpendicular to the plane. The exact spatial distribution would depend on the symmetry of Sr dopants relative to the plane.

It is well established experimentally that 2:1:4 is antiferromagnetic with semiconducting behavior.[13] Substitution of Sr for La leads to superconductivity. Neutron studies[14] indicate that the Néel state is destroyed by doping and that the spin-spin correlation length is reduced from 35 to 8 Å as the Sr concentration is varied from 0.02 to 0.18. More importantly, the local order is found to be the same for the doped and pure 2:1:4. Thus, it is stated[14] that no important differences in magnetic scattering were found in the normal and superconducting states. It is important to note that the interpretation of the neutron data is that doping affects the spin-spin correlations leaving the moment essentially unchanged from that of the pure state. Birgeneau *et al.*[14] caution that an unambiguous interpretation will require samples where the Meissner fraction far exceeds the 15% observed in the currently available samples. Birgeneau *et al.* also indicate that there are unexplained differences between samples.

Our calculations are in agreement with the above picture. A moment localized to the Cu atoms is obtained theoretically. This moment is practically unchanged up to a Sr concentration of 25%. We can similarly claim randomly placed Sr will disorder the moments. Our computed moment of $0.35\mu_B$ is quite similar to the experimental value from neutron [$(0.35\pm0.05)\mu_B$ (Ref. 15(b)] and muon-spin resonance [$(0.2-0.6)\mu_B$].[15(a)] It is certain that the moment is much less than $1.0\mu_B$ which would be obtained for spin 0.5.

Finally, whether or not the antiferromagnetic ground state is obtained is closely related to the number of plane waves with which the nonspherical potential is expanded. No AF ground state was obtained when the potential was expanded with 4913 or 9261 plane waves. An AF ground state was found for 24 389 plane waves. This calculation was not run to SCF due to extreme computer cost. The moment was $0.31\mu_B$ and increasing when the job was terminated. Since the number of basis functions was constant, the representation of the potential is critical.

There are other interpretations of the AF state of 2:1:4. The resonating valence bond picture[16] assumes a spin of $S=0.5$ on the Cu sites. An alternative strong-correlation picture[17] is that doping produces holes of spin 0.5 on the O sites in the plane. These local moment theories claim to account for much of the data but with adjustable parameters. Clearly a fundamental assumption is that Sr doping produces holes on the O sites within the Cu-O plane. Our energy-band results would place holes in the plane only for low doping levels. For larger values of doping as in the superconducting region, our energy-band results show holes on the O in the La plane not the Cu-O plane.

Spin and charge fluctuations have been mentioned as a possible mechanism for superconductivity. Our band-theory picture indicates a substantial coupling between planes via the O holes. Fluctuations within the plane can occur as a result of the fluctuations on the O. These fluctuations produce an attractive pairing interaction leading to a $T_c=280$ K for 0.15 Sr doping of 2:1:4.[18] An important feature of our band results is the capability of coupling within and between planes.

To support our claim that the width of the bands at E_f are sensitive to how well the potential is represented, a comparison is made for two different representations of the potential with the wave functions remaining identical. Figure 2 shows the narrow band for 4913 and 15 356

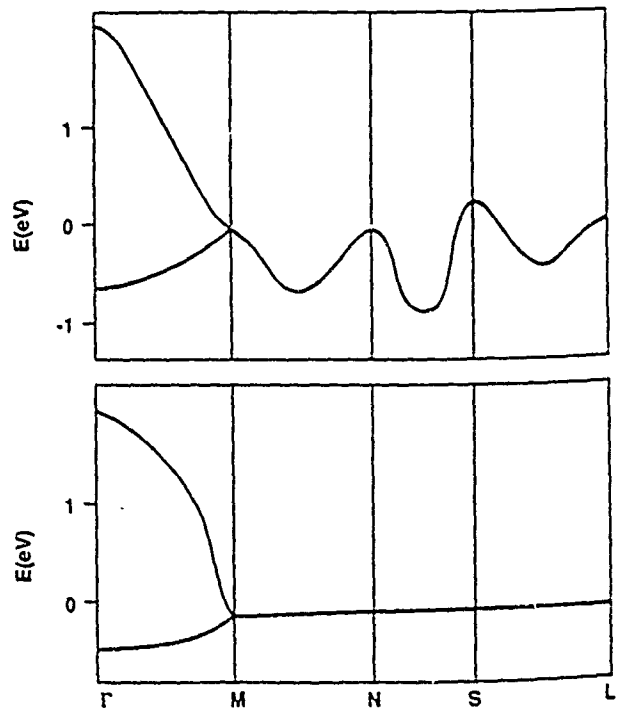


FIG. 2. Non-spin-polarized energy bands at E_f for 4913 and 15 356 plane waves in potential.

plane waves in the potential, respectively. The band is narrow if the potential has many plane waves. Only the band at E_f has been included for clarity. The band for 4913 plane waves is similar to that in the linear muffin-tin orbital (LMTO) calculations. The agreement cannot be exact because the LMTO calculation relies on the approximation of overlapping spheres and the PSF calculation relies on truncating the potential to 4913 plane waves. In further support of the accuracy of the PSF method, previously we have shown that the correct thermodynamic properties are computed for bulk Si (Ref. 19) and that the correct bond length can be computed for a CO molecule.[1] We have also computed the correct surface structure of Si.[1]

In summary, the itinerant energy-band model yields a band structure which has a moment of approximately the correct magnitude on the Cu site. Calculations for Sr doping support the notion of spin or charge fluctuations on the Cu and O sites. No evidence for localized holes of spin 0.5 on the planar O sites is obtained. Our results support the interpretation put forth by Birgeneau *et al.* that Sr disorders long-range spin-spin correlations leaving the local Cu moment relatively unchanged.

ACKNOWLEDGMENTS

We wish to thank W. Y. Hsu, S. T. Chui, and A. Rajagopal for helpful discussions.

APPENDIX

The PSF method has been described in detail[1] with application to CO and Si. Here, we will provide an outline of the method and all information that would be required to repeat our calculations. The unit-cell parameters and atom positions have been determined by Jorgenson *et al.* The unit cell is divided into muffin-tin and interstitial regions. The muffin-tin radii R_{MT} are 2.81 a.u. for La, 1.84 a.u. for Cu, and 1.62 a.u. for O. The potential is divided into spherically symmetric potentials within the muffin tins and a nonspherical potential that extends throughout the unit cell including the spheres. To form the PSF wave functions, we construct linearized muffin-tin orbitals exactly as in the LMTO or LAPW methods using the radial solutions ϕ and $\dot{\phi}$. $\dot{\phi}$ is the energy derivative of ϕ, the solution of the muffin-tin potential for a given energy. The energy parameter is chosen to be an average band energy as in the linearized methods.

Beyond this point, the PSF method differs from other methods. First, we define two new radii R_{in} and R_{out} for all the muffin tins. R_{in} is chosen to be $0.95\,R_{MT}$. R_{out} is chosen to be $1.25\,R_{MT}$. The PSF wave function at site n for angular momentum lm is now defined as

$$\psi_{\mathrm{PSF}_{nlm}}=\begin{cases} a_{nlm}j_{lm}(k_1,r_n)+b_{nlm}j_{lm}(k_2,r_n), & r<R_{\mathrm{in}}\ , \\ N_{lm}(\phi(E,r_n)+\omega_{lm}\dot{\phi}_{lm}(E,r_n)), & R_{\mathrm{in}}<r<R_{\mathrm{out}}\ , \\ -K_{lm}(k,r_n), & R_{\mathrm{out}}<r\ . \end{cases}$$

The parameters ω_{lm} and N_{lm} are chosen for continuity of the function and its radial derivative at R_{out}. The Neuman function K_{lm} has an energy parameter of -0.3 Ry. Within the distance R_{in}, the parameters a_{nlm} and b_{nlm} are also chosen for continuity of the PSF function at R_{in}. The energies for the Bessel functions j_{lm} are $k_1^2=0.1$ Ry and $k_2^2=0.2$ Ry. The above procedure is followed for the s, p, and d orbitals on the La and Cu sites and the s and p orbitals on the O sites. The results are not sensitive to the choice of energy parameters k_1 and k_2 within the radius R_{in}. The tail-parameter energy -0.3 Ry is the result of calculations on many different systems. The resultant energy bands are not very sensitive to the tail parameter if a value near -0.3 Ry is used. The above-defined PSF orbitals are smooth throughout space thus enabling straightforward Fourier transforms as previously shown.[1] The criterion for the number of plane waves in the transform of the PSF orbitals is that R_{MT} times the maximum reciprocal-lattice vector be 7.0 in the three reciprocal-space directions thus defining a polyhedron in reciprocal space instead of a sphere.

The PSF orbitals are now defined by a linear combination of plane waves. The PSF orbitals are orthogonalized to the core states by augmentation as in the LAPW method. The solution of the Hamiltonian is fully described in Ref. 1 and is not repeated here. The charge density is in terms of spherically symmetric parts within the muffin tins and a Fourier series representing the nonspherical contribution. This plane-wave sum contains approximately eight times the number of plane waves in the wave function.

The self-consistency loop follows that of the standard band theory. The charge density is used to define a new potential which is mixed with the old potential. New PSF orbitals are constructed and the Hamiltonian is solved again until the input and output potentials are the same to a small deviation. The PSF orbitals change as one iterates to self-consistency. The PSF orbital molds to the charge density and changes as the charge density changes especially in the bond formation region between atoms. The applications described here and in Ref. 1 indicate that the plane number waves in the potential is vitally important for computation of properties. This dependence is shown in detail for a CO molecule.[1] In this paper we have shown that a bandwidth similar to that for other methods is obtained for a small number of plane waves.

[1] R. V. Kasowski, M.-H. Tsai, T. N. Rhodin, and D. D. Chambliss, Phys. Rev. B 34, 2656 (1986).
[2] U. von Barth and L. Hedin, J. Phys. C 5, 1629 (1982).
[3] M.-H. Tsai, John Dow, and R. V. Kasowski (unpublished).
[4] T. C. Leung, X. W. Wang, and B. N. Harmon, Phys. Rev. B 37, 384 (1988); P. A. Sterne, C. S. Wang, and W. M. Temmerman, *ibid.* 37, 7472 (1988); G. Y. Guo, W. M. Temmerman, and G. M. Stocks, J. Phys. C 22, 120 (1988).
[5] R. V. Kasowski, W. Y. Hsu, and F. Herman, Solid State Commun. 63, 1077 (1987).

[6]*Novel Superconductivity,* edited by S. A. Wolf and V. Z. Kresin (Plenum, New York, 1987).
[7]D. H. Lee and J. Ihm, Solid State Commun. 62, 811 (1987).
[8]A. K. Rajagopal and R. V. Kasowski (unpublished).
[9]S. H. Wei and A. Zunger, Phys. Rev. Lett. 56, 2391 (1986).
[10]R. V. Kasowski and M.-H. Tsai, Phys. Rev. Lett. 60, 546 (1988).
[11]S. Ciraci and Inder P. Batra, Phys. Rev. Lett. 56, 877 (1986).
[12]T. Kendelewicz, P. Soukiassian, R. S. List, J. C. Woicik, P. Pianetta, I. Lindau, and W. E. Spicer, Phys. Rev. B 37, 7115 (1988).
[13]J. D. Jorgensen *et al.*, Phys. Rev. Lett. 58, 1024 (1987).
[14]R. J. Birgeneau *et al.* (unpublished).
[15](a) Y. J. Uemura *et al.*, Phys. Rev. Lett. 59, 1045 (1987); (b) G. Shirane *et al.*, *ibid.* 59, 1613 (1987); (c) D. Vaknin *et al.*, *ibid.* 58, 2802 (1987).
[16]P. W. Anderson, G. Baskaran, Z. Zou, and T. Hsu, Phys. Rev. Lett. 58, 2790 (1987).
[17]Amnon Aharony *et al.*, Phys. Rev. Lett. 60, 1330 (1988).
[18]S. T. Chui, R. V. Kasowski, and W. Y. Hsu, Europhys. Lett. 9, 385 (1989).
[19]M.-H. Tsai, J. D. Dow, and R. V. Kasowski, Phys. Rev. B 38, 2176 (1988).