

A THEORETICAL STUDY OF DEFECTS  
IN AMORPHOUS SEMICONDUCTORS

by

DAVID HAMILTON VANDERBILT

B.A., Swarthmore College  
(1976)

SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE  
DEGREE OF

DOCTOR OF PHILOSOPHY  
IN PHYSICS

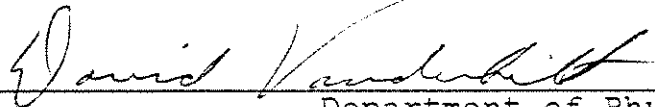
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1981

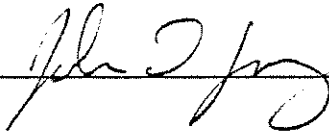
© Massachusetts Institute of Technology 1981

Signature of Author




Department of Physics  
May 1, 1981

Certified by



John D. Joannopoulos  
Thesis Supervisor

Accepted by

  
George F. Koster  
Chairman, Department Committee



A THEORETICAL STUDY OF DEFECTS  
IN AMORPHOUS SEMICONDUCTORS

by

DAVID VANDERBILT

Submitted to the Department of Physics  
on May 1, 1981  
in partial fulfillment of the requirements for the  
Degree of Doctor of Philosophy in Physics

ABSTRACT

The structure and electronic properties of bonding defects in amorphous semiconductors are investigated using a variety of theoretical approaches. First, realistic self-consistent pseudopotential calculations on superlattice structures containing defects are applied to glassy selenium as a model system. It is demonstrated that accurate structural ground state and excitation energies can be obtained; these are used to rule out the existence of certain defects and predict the relaxed structural configuration of others. Information on the electronic structure is then extracted and utilized in the construction of a tight-binding Hamiltonian. The latter is applied to more realistic structures containing truly isolated defects, using cluster-Bethe-lattice techniques. The origin, character, energy location, and localization of the defect states associated with bond coordination defects, and with defect pairs and certain relaxed defects, are discussed. It is shown that a unique  $\pi$ -interaction between non-bonding p-orbitals stabilizes the neutral one-fold defect and gives rise to a well-localized deep gap state. Experimental support for this picture is discussed.

Similar cluster-Bethe-lattice tight-binding calculations are then applied to the increasingly dissimilar cases of heteropolar chalcogenides and polyacetylene. Important similarities and differences emerge. Defects in arsenic selenide are classified, and a new selection rule is formulated which greatly reduces the number of allowed Fermi-level pinning reactions. The defect gap states are found to arise due to unique bond-orbitals that occur at defects. Finally, bonding

defects (chain ends and crosslinks) in polyacetylene are investigated, with particular attention to soliton binding or emission at these sites. Certain other kinds of structural disorder are also studied, and the effect upon the electronic density of states is calculated in each case.

Thesis Supervisor: Dr. John D. Joannopoulos

Title: Associate Professor of Physics

TABLE OF CONTENTS

CHAPTER I INTRODUCTION . . . . .	5
CHAPTER II SELENIUM . . . . .	9
(a) Introduction . . . . .	9
(b) Bethe-Lattice Tight-Binding Calculations . . . . .	16
(c) Self-Consistent Pseudopotential Calculations . . . . .	35
(d) Total Energies . . . . .	43
(e) Summary and Conclusions . . . . .	53
CHAPTER III ARSENIC SELENIDE . . . . .	55
(a) Introduction . . . . .	55
(b) Classification of Defects . . . . .	56
(c) Electronic States . . . . .	66
(d) Defect Creation Energies . . . . .	83
(e) Summary and Conclusions . . . . .	89
CHAPTER IV POLYACETYLENE . . . . .	92
(a) Introduction . . . . .	92
(b) Topological Defects . . . . .	93
(c) Structural Disorder . . . . .	111
(d) Summary and Conclusions . . . . .	123
APPENDIX A BETHE LATTICE FORMALISM . . . . .	125
APPENDIX B SELF-CONSISTENT PSEUDOPOTENTIAL METHOD . . . . .	140
ACKNOWLEDGMENTS . . . . .	166
REFERENCES . . . . .	167



## CHAPTER I INTRODUCTION

For roughly fifty years, our understanding of the existence of energy bands and gaps in the electronic structure of crystalline semiconductors has rested upon Bloch's theorem. It is not surprising, therefore, that the discovery of corresponding features in the spectrum of amorphous materials has led to new theoretical approaches and new physical insights. Perhaps the central insight has been the understanding that it is the short range order (the uniformity in the local bonding environment of each atom), rather than the long range order (periodicity), which primarily determines the existence and character of bands and associated gaps. Accordingly, amorphous semiconductors are presently understood to be materials that possess almost perfect short range order, while lacking long range order.

The prevalence of short range order means that almost all atoms are able to satisfy their preferred coordination, with fairly uniform bond lengths and bond angles, resulting in a "continuous random network." However, it is inevitable that the geometrical constraints involved in packing atoms will occasionally cause the short range order to break down. One therefore expects some density of "defects" in which the preferred bonding geometry is not satisfied for one or more atoms. These intrinsic defects play a central role in the physics of amorphous semiconductors because they can give rise to gap

states which dominate the optical, transport, and magnetic properties of the material.

It is clearly important to obtain a thorough theoretical understanding of such defects. Of course, the absence of translational symmetry in the defect problem means that traditional approaches based upon Bloch's theorem are inapplicable. However, a variety of alternative techniques have become available in the last ten years or so. Of these, two have been found to be particularly useful and have been employed extensively below. One is the construction of artificially periodic ("superlattice") structures containing defects, which may then be solved using traditional realistic (e.g. self-consistent pseudopotential) methods. The second is the cluster-Bethe-lattice method, in which structural models containing isolated defects are solved exactly for a given tight-binding Hamiltonian using Green's function techniques.

This thesis represents a detailed theoretical investigation of defects in three model systems: amorphous selenium (a-Se), amorphous arsenic selenide (a-As<sub>2</sub>Se<sub>3</sub>), and polyacetylene [(CH)<sub>x</sub>]. This choice of materials provides interesting contrasts. The first two are chalcogenide glasses, one elemental and the other heteropolar; the third is an organic polymer. a-Se and a-As<sub>2</sub>Se<sub>3</sub> are experimentally quite similar, although the partial ionic character and the possibility of breaking the chemical ordering (introducing "like-atom bonds") in As<sub>2</sub>Se<sub>3</sub> give rise to a variety of inter-



esting phenomena which are not to be found in pure Se. In polyacetylene, the possibility of breaking the bond-alternation (Peierls) ordering (introducing "solitons") provides still other unique phenomena not found in the other two materials.

Despite these contrasts, there are a number of unifying threads which weave through our study of all three of these systems. The same theoretical methods will frequently be applied, and certain important concepts will re-emerge. (For example, total energy calculations and estimates will play an important role in determining what defect configurations are likely to exist and what their stable configuration will be.) All three materials are amorphous semiconductors with gaps on the order of 1-2 eV. All three are believed to contain coordination defects with associated deep gap states. Perhaps a particularly intriguing common feature is the existence of free spins at charge-neutral defects; this can occur when defects are topologically constrained to be created in pairs, as is the case for certain defects in all three systems.

The plan of the thesis is straightforward. We shall consider each of these materials in turn, with the emphasis upon the physics and the results, in Chapters II-IV. The theoretical methods will be introduced as necessary as we proceed; certain details concerning the Green's function solution of the cluster-Bethe-lattice, and the self-consistent

pseudopotential method, are deferred to Appendices A and B respectively.

## CHAPTER II SELENIUM

### (a) Introduction

From a theorist's point of view, the elemental nature of glassy selenium makes it uniquely suited as a model system for the chalcogenide glasses, a family of materials which includes the arsenic selenide, sulfide and telluride glasses as well. All of these materials have a number of experimental properties in common.<sup>1-6</sup> They all are excellent glass-formers, with glass transition temperatures in the range 300-500°K. They are all p-type semiconductors with the Fermi level pinned near midgap.<sup>1-3</sup> They all show diamagnetic behavior at low temperatures, with the density of free spins estimated at less than  $10^{14}$  cm<sup>-3</sup> on the basis of ESR (Electron Spin Resonance) measurements.<sup>4,5</sup> Moreover, there is a variety of photoinduced phenomena common to these materials, including photoluminescence,<sup>6</sup> photoinduced ESR,<sup>5</sup> and photoinduced IR (infrared) absorption.<sup>5</sup> The photoluminescence is quite similar in all cases, with a large Stokes shift (about half the band gap), an excitation spectrum which falls off strongly above the Urbach edge, and an anomalous quenching of the photoluminescence quantum efficiency beginning at low temperatures ( $\sim 10^{\circ}\text{K}$ ).<sup>6</sup>

Naturally, one wanted to understand these features in terms of a single model for the chalcogenide glasses. As it

happened, the key to the development of this model lay in a long-standing puzzle: how can the Fermi level be pinned in a system which shows no free spins? The traditional picture is that Fermi level pinning is accomplished by the capture of added carriers (say electrons) into gap states at the Fermi level. Once an electron is added to given gap state, however, the repulsive Coulomb correlation energy  $U_c$  will make it unfavorable to place a second electron into the same state. One therefore expects a manifold of states (those whose "first ionization potential" lies between  $\epsilon_F$  and  $\epsilon_F + U_c$ ) to contain free spins. These spins were not observed.

The solution to this puzzle was indicated in an obscure paper by Anderson.<sup>7</sup> He suggested, on the basis of a simple Hubbard Hamiltonian, that a strong electron-phonon coupling in these materials could give rise to a negative effective correlation energy,  $U_{\text{eff}} < 0$ , which would tend to pair spins in the material. The presence of this "negative U" can be understood as follows. First, consider adding a single electron to a state which is initially unoccupied. In a system with a strong electron-phonon coupling, there will be a different equilibrium structural configuration for each occupancy; the system will therefore undergo structural relaxation when the first electron is added. This relaxation always occurs in such a way as to lower the one-electron eigenvalue of the newly-occupied eigenstate. (Let  $U_{\text{ph}}$  be the magnitude of the lowering.) If  $U_c - U_{\text{ph}} \equiv U_{\text{eff}} < 0$ , then as a second electron is added to the

system, it will be energetically favorable to doubly occupy this level before singly occupying a different one. In this way the Fermi level is pinned without accumulating free spins.

Anderson did not specify a microscopic model for the states and relaxations involved, but he appears to have been thinking in terms of a truly random picture of the glass containing a broad distribution of bond lengths and bond strengths, with many or all valence states being negative  $U$  states. Subsequent models based upon this idea have even tried to explain the existence of the band gap on the basis of these two-electron effects.<sup>8-10</sup> However, this picture would require a huge  $U_{ph}$  to account for a band gap of  $\sim 2$  eV, and is hard to reconcile with the uniformity of bond lengths indicated by the sharp first peak in the radial distribution function of the glass.<sup>11</sup>

The next major advance came in a series of papers by Mott, Davis, and Street,<sup>12,13</sup> who recognized that the prevalence of short range order in these materials gave rise to bands and gaps. They attributed the negative  $U$  to some density ( $\sim 10^{18}$   $\text{cm}^{-3}$ ) of intrinsic structural defects which could exist in the glass. They referred to these defects as "dangling bonds," and used  $D^+$ ,  $D^0$ , and  $D^-$  to denote the three charge states of the defect.  $D^+$  and  $D^-$  were assumed to have hydrogenic (acceptor and donor) gap states, while the  $D^0$  would have a midgap state and a free spin. This model proved to be extremely successful in accounting for many of the experimental properties of

chalcogenide glasses mentioned above. In addition to providing a natural explanation for the diamagnetic Fermi level pinning, it explained all of the photoinduced phenomena on the basis of the photoexcitation of  $D^{\circ}$  from  $D^{+}$  and  $D^{-}$ . The photoluminescence results from the radiative recombination  $D^{\circ} \rightarrow D^{+}+e$  or  $D^{-}+h$ , with the half-band-gap Stokes shift a natural consequence of the relaxations at the defect. The photoinduced ESR is just due to the free spin on the  $D^{\circ}$ , and the photoinduced absorption corresponds to  $h+D^{\circ} \rightarrow D^{+}$  or  $e+D^{\circ} \rightarrow D^{-}$ .

Despite the extraordinary success of this model, there was still no clear identification of the microscopic bonding configurations involved, and no clear explanation of why these phenomena should not occur in other (e.g. pnictide or tetrahedral) amorphous semiconductors. By referring explicitly to the coordination chemistry of these lone pair semiconductors, however, Kastner, Adler, and Fritzsche<sup>14-16</sup> were able to address these points. On the basis of simple bond-orbital models, they suggested that the  $D^{-}$  was indeed a dangling bond ( $C_1^{-}$  in their notation), but that the  $D^{+}$  corresponded to a threefold coordinated selenium atom ( $C_3^{+}$ ). These two defects form a "Valence Alternation Pair" (VAP). The Fermi-level pinning reaction  $C_1^{-}+2h \rightarrow C_3^{+}$  occurs via a mechanism in which a dangling Se atom swings over and bonds to an atom of a neighboring chain. They point out that it is really the availability of lone-pair orbitals on the chalcogen atoms which makes this negative U possible. This explains the appar-

ent absence of the negative U in pnictide and tetrahedral systems.

These are the broad outlines of the defect model for chalcogenide glasses. While it is used today to interpret much of the experimental work on chalcogenide glasses,<sup>2-4,12,15,17</sup> it is not universally accepted. As mentioned above, some models propose a smooth distribution of bond strengths and invoke two-electron effects to explain the gap.<sup>8-10</sup> Emin has proposed a model based on small polarons.<sup>18</sup> Several other defect-related models have been suggested,<sup>19-22</sup> including the raft model of Phillips.<sup>19</sup> The controversy surrounding the defect model provides one impetus for a detailed theoretical investigation.

A second impetus is the fact that there has been controversy among the proponents of the defect model themselves about the nature of the defects involved. The simple models in circulation have not been able to address with sufficient accuracy such important questions as the character, localization, and energy position of the defect states. A particular source of controversy has been the nature of the photoinduced neutral ( $D^{\circ}$ ) center.<sup>12,14-16,23-24</sup> Since it governs the photoinduced phenomena, it is crucial to obtain a better understanding of this defect. Mott, Davis, and Street originally assumed that the preferred neutral defect would be a 1-fold atom, weakly bonded to an atom of a neighboring chain. A 3-fold defect was thought to be unlikely because a Jahn-Teller distortion was

expected to weaken one of the three equivalent bonds.<sup>12,23</sup> However, Kastner, Adler, and Fritzsche then pointed out that the electronic energy gained by bond formation would favor the 3-fold site.<sup>14</sup> On the other hand, one expects that inclusion of intercore repulsive terms could easily turn the tides in favor of the 1-fold defect again. On the basis of the simple bond-orbital models, it was impossible to resolve this controversy theoretically. Moreover, although important new experiments have shed light upon the photoluminescence mechanism,<sup>25</sup> the VAP radial distribution function,<sup>25,26</sup> and the density of states in the gap,<sup>26-28</sup> these experiments are generally not sensitive to the bonding configuration of the defects involved.

To obtain a satisfactory understanding of the basic physics underlying the structure, ground-state properties, and elementary excitations of defects, it is necessary to develop a theoretical description which is realistic. Owing to the inherent nonperiodicity of the defect problem, this has been an elusive goal. We have approached the problem by employing a powerful combination of realistic self-consistent pseudopotential (SCPSP) calculations on periodic superlattice structures containing defects and by employing flexible tight-binding (TB) techniques which can be used to solve nonperiodic single-defect structures. The SCPSP calculations employ a realistic Hamiltonian and are self-consistent, but the structure, being periodic, is approximate. Defect states



get broadened into bands, and therefore these results do not lend themselves to direct interpretation. However, the SCPSP results were then used as the basis for the fitting of a TB Hamiltonian, which was then applied to more realistic nonperiodic (cluster-Bethe-lattice) single-defect structures. These realistic tight-binding calculations reproduced all of the essential features of the SCPSP results. Moreover, they provide a good physical description of the nature of the defect states.

Finally, we have resolved the controversy regarding the neutral ( $D^0$ ) center by performing self-consistent pseudopotential calculations of the total energies of various possible bonding configurations for the neutral defect. The calculation of total energies is of crucial importance because it allows us to determine the stable (or metastable) defects which can actually occur in the glass. This can be accomplished simply by searching for the lowest-energy structural configuration.

The remainder of this chapter will be devoted to discussing these results. Since the main goal here is to present the theoretical understanding which has emerged from this work, we begin Sec. II(b) by discussing the nature of the electronic states at unrelaxed onefold and threefold coordination defects. We begin with an extremely simplified but intuitively accessible TB model and then discuss in some detail the new physics which arises when a realistic TB Hamiltonian, fit to a

realistic SCPSP calculation, is used. We then extend the work to include more complicated defect structures, including close defect pairs and certain kinds of relaxations. The method used to obtain the density of states for these defect structures (generalized to the case of nonorthogonal basis orbitals) is given in Appendix A, together with a discussion of an empirical method for estimating defect total energies. Sec. II(c) contains a discussion of the SCPSP results and the fitting of the realistic TB Hamiltonian to these results. In Sec. II(d) we present the results of realistic total energy calculations based upon the SCPSP method. (Some details relevant to the methods of Secs. II(c) and II(d) are deferred to Appendix B.) Finally, in Sec. II(e) we summarize and discuss the implications of the work.

#### (b) Bethe-Lattice Tight-Binding Calculations

Our model for the glass will consist of defects embedded in an otherwise perfectly coordinated continuous random network which will be referred to as the bulk. For the case of glassy Se, this consists of infinite chains (and perhaps some eightfold and larger rings) packed in such a way that bond angles and bond lengths remain approximately those of the crystalline structure. The first step, clearly, is to model the electronic density of states of the bulk. Because the inter-chain bonding is weak, having considerable van der Waals character, and because the gross features of the density of

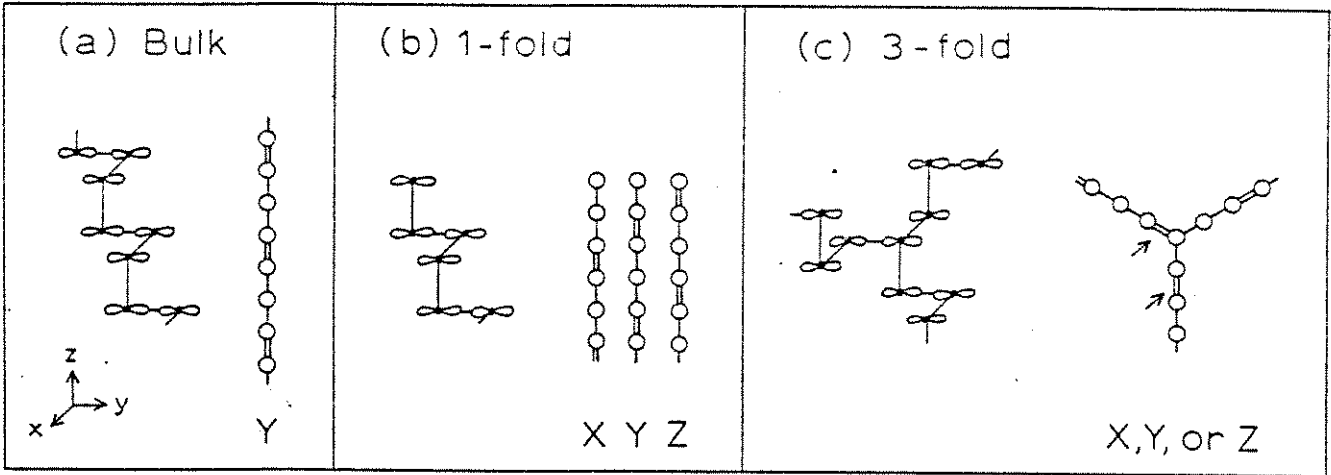
states are expected to be determined by the short-range order (i.e., the local bonding configuration), we model the bulk random network by focusing on a single average chain embedded in the glass. Similarly, an N-fold coordination defect is modeled by attaching N such bulk chains to the defect site. For definiteness we choose a chain identical to that in the trigonal crystal, keeping in mind that bond-angle fluctuations and dihedral-angle disorder may broaden the resulting density of states slightly and introduce narrow band tails. Thus, within a first-neighbor tight-binding point of view, our model for the electronic structure of the bulk has reverted to that of trigonal Se, although the results must be interpreted with the philosophy outlined above in mind.

The electronic structure of these trigonal chains is by now well understood.<sup>29,30</sup> In the valence region, one finds a 4s-like band fairly well separated from three 4p-like bands. The latter consist of filled valence and lone-pair bands and an empty antibonding conduction band.

The origin of the distinct p bands may be understood by referring to Fig. 1(a). Here we have taken an extremely simplified model for trigonal Se which nevertheless contains much of the essential physics. The helical chain is chosen to have bond angles and dihedral angles of exactly  $90^\circ$ , and we use a simple nearest-neighbor tight-binding model with just the three valence p orbitals on each site. Because of the geometry, there are no interactions between the systems of  $p_x$ ,

Figure 1

Structure and interaction diagram (double line  $V_\sigma$ , single line  $V_\pi$ ) for (a) helical chain, (b) onefold defect, and (c) threefold defect.



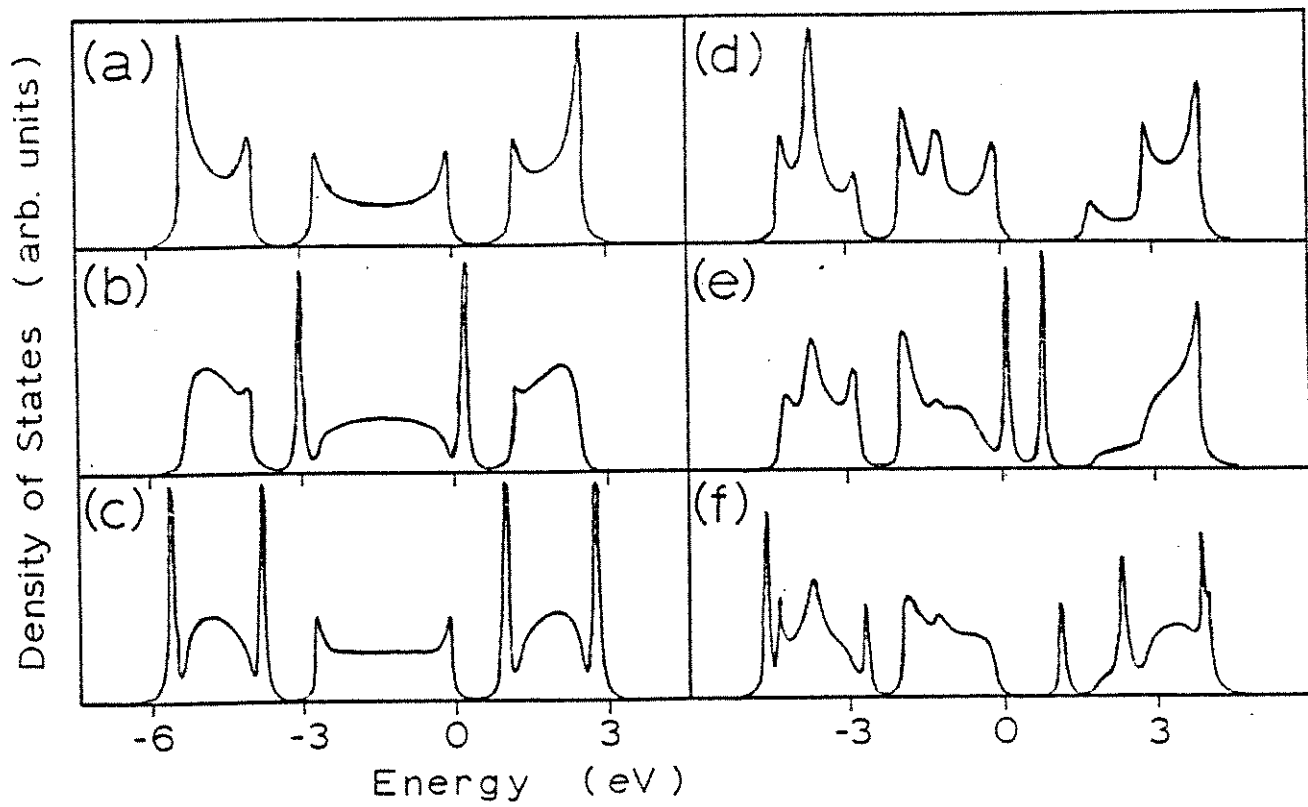
$p_y$ , and  $p_z$  orbitals, so that the three subsystems become decoupled and degenerate. It thus suffices to consider only the system of  $y$  orbitals shown explicitly in Fig. 1(a). The only free parameters in the model are a  $\sigma$ -bonding interaction integral  $V_\sigma$  and a weaker  $\pi$  interaction  $V_\pi$ . If we represent these schematically by double and single lines, respectively, we obtain the interaction diagram shown at right in Fig. 1(a).

The origin of the three  $p$  bands now becomes clear. In the limit  $V_\pi \rightarrow 0$ , the  $\sigma$  bonded pairs of orbitals decouple from the remaining nonbonding orbitals (NBO's or lone pairs). The latter give rise to a discrete lone-pair level at the unperturbed  $p$  energy, while the former produce bonding and antibonding ( $\sigma$  and  $\sigma^*$ ) levels below and above the lone-pair level, respectively. For  $V_\pi \rightarrow 0$  these three discrete levels must broaden into bands; the resulting density of states is easily calculated and is shown in Fig. 2(a). The  $\sigma$  and the lone-pair bands are filled, while the  $\sigma^*$  band is empty. Note that the widths of the bands are determined not simply by a direct  $\pi$  interaction, but rather by a weaker effective interaction of order  $V_\pi^2/V_\sigma$ , because of the alternation of bond orbitals and NBO's (separated by energy  $V_\sigma$ ). This fact will be crucial to the understanding of bond-coordination defects.

Let us consider the simplest such defect, a onefold coordinated (dangling-bond) site terminating a chain in the bulk. This is shown in Fig. 1(b) along with the appropriate interaction diagrams for the systems of  $x$ ,  $y$ , and  $z$  orbitals. Note

Figure 2

Local density of states averaged over sites near the defect. Using a simple tight-binding Hamiltonian: (a) bulk, (b) onefold defect, and (c) threefold defect. Using a realistic tight-binding Hamiltonian: (d) bulk, (e) onefold defect, and (f) threefold defect.





that the three subsystems are still decoupled but but no longer identical; each is terminated in a different fashion. The  $p_z$  system ends on  $\sigma$  and  $\sigma^*$  orbitals which contribute featurelessly to the bonding and antibonding bands. Similarly, the terminal non-bonding  $y$  orbital contributes to the lone-pair band. Of special interest is the system of  $x$  orbitals which terminates on a pair of NBO's connected by a direct  $\pi$  interaction. Recall that this is a stronger interaction than the effective interaction which determines the lone-pair bandwidth. As a result the two NBO's are split into  $\pi$  bonding and  $\pi$  antibonding combinations [henceforth  $\pi(\text{NBO})$  and  $\pi^*(\text{NBO})$ , respectively], giving rise to localized states above and below the lone-pair band edges. This is shown in Fig. 2(b), where the model of Fig. 1(b) is solved exactly using Green's-function techniques. If the defect is neutral, the  $\pi^*(\text{NBO})$  state above the valence-band maximum is half filled (i.e., contains a hole). This defect state will be localized strongly to at most two sites. At this point the existence of localized states at the onefold defect is a consequence of the unique direct  $\pi$  interaction between NBO's on neighboring sites.

Consider now the threefold defect shown in Fig. 1(c). The geometry has been chosen such that there is an exact symmetry of  $120^\circ$  rotation about an axis passing through the threefold site. The decoupled systems of  $x, y$ , and  $z$  orbitals are once again identical, restoring the threefold degeneracy. We observe a behavior analogous to that at the onefold site. This

time there is a pair of  $\sigma^*$  orbitals (and a pair of  $\sigma$  orbitals) shown by arrows which are connected by a direct  $\pi$  interaction. Once again the splitting exceeds the band width and we expect a pair of threefold-degenerate localized states [ $\pi^*(\sigma^*)$  and  $\pi(\sigma^*)$  in our notation] to emerge from the antibonding band edges, and similarly for the bonding band. Figure 2(c) shows that this is indeed the case. The threefold degenerate  $\pi(\sigma^*)$  state below the conduction-band minimum contains a single electron if the defect is neutral. Note that this state will be highly delocalized, sharing its character among at least the seven central atoms near the defect.

At this point we have seen that onefold defects will tend to produce highly localized defects at the lone-pair band edges, and that threefold defects will give rise to much more delocalized (but still nonhydrogenic) gap states at the bonding and antibonding band edges. Before carrying the discussion further, it becomes essential to improve drastically upon the simple model presented above. We will extend the above discussion by using a much more realistic tight-binding model, but in doing so a fundamental question emerges: How can we be sure that a tight-binding model which has been fit to the bulk density of states will be valid in the neighborhood of the defect? The only way to answer this question is to appeal to realistic first-principles calculations on defect structures. This has been done in Sec. II(c), where we have applied self-consistent pseudopotential (SCFSP) calculations to crystalline Se and to

periodic superlattice structures containing defects. We fit a tight-binding Hamiltonian to the bulk crystal and then determine what modifications to the TB model are necessary to give accurate results for defects. We find that it is only necessary to adjust the diagonal Hamiltonian matrix elements on the defect sites themselves. The onefold atom has its self-energy shifted upward by 1.25 eV, while the threefold atom is shifted downward by the same amount.

The SCPSP calculations are used to verify the existence of these self-energy shifts in Sec. II(c), but they can be understood physically in the following way. Suppose we break a single bond in an infinite Se chain, producing two onefold defects. The two electrons which were shared in the  $\sigma$ -bonding orbital each singly occupy a new NBO on the onefold sites. In a simple picture of atomic orbitals, one would then have a hole localized completely on one of the two NBO's on the defect site, and each atom would be individually charge neutral. However, we have shown above that one of the NBO's at the defect interacts strongly with an NBO on the neighboring site, so that in fact the hole is shared between the two sites contributing to the  $\pi^*(\text{NBO})$  defect state. This would lead to a charge of  $\sim \pm 0.5 e$  on the second-to-the last and last atoms, respectively. Now the Coulomb interaction shifts the self energies upward on the terminal atom (and downward on the penultimate one), causing the  $\pi(\text{NBO})$  state to reside mostly on the neighboring site and the  $\pi^*(\text{NBO})$  state to reside mostly on the defect site. The

charge transfers and self-energy shifts now adjust themselves self-consistently. When a realistic calculation is done [including, for example, the fact that the  $\pi^*(\text{NBO})$  defect state decays into the bulk], the net result is that the charge transfers are small and the self-energy shift is large only on the defect site itself [see Sec. II(c)].

The situation for the threefold defect is again analogous. The defect atom shares three of its electrons in  $\sigma$  bonds; if the fourth electron could be localized to the defect site, each atom would be individually neutral. Instead, the electron is shared in the highly delocalized  $\pi(\sigma^*)$  defect state. The self-consistent process leads to a lowered self-energy on the threefold atom which strongly reduces the amount of charge transfer from the defect site.

When these self-energy shifts are included, we obtain the realistic tight-binding Hamiltonian presented in Table I. The TB model includes s as well as p states with s-p interactions, overlaps between orbitals on neighboring sites, and self-energy shifts of  $\pm 1.25$  eV on onefold and threefold sites, respectively. The defect structures to be considered are identical to those of Fig. 1 except for the distortion necessary to obtain the experimental bond angle of  $102.54^\circ$  and the dihedral angle of  $100.27^\circ$ .<sup>31,32</sup>

These structures have been solved within the TB Hamiltonian. The introduction of overlaps introduces subtleties which are discussed in Appendix A. The results are shown

Table I

Tight-binding Hamiltonian for selenium.  $E$  is diagonal matrix element,  $V$  is nearest-neighbor interaction, and  $S$  is nearest-neighbor overlap.  $\Delta_m$  is change in  $E_s$  and  $E_p$  at  $m$ -fold site.

---

$E_s$	$E_p$	$\Delta_1$	$\Delta_3$
-13.00 eV	-1.10 eV	+1.25 eV	-1.25 eV
$V_{ss}$	$V_{sp}$	$V_{pp\sigma}$	$V_{pp\pi}$
-2.27 eV	-2.07 eV	-2.97 eV	-1.19 eV
$S_{ss}$	$S_{sp}$	$S_{pp\sigma}$	$S_{pp\pi}$
0.08	0.08	0.13	0.15

---

for the bulk chain, onefold defect, and threefold defect in Figs. 2(d), 2(e), and 2(f), respectively. The effect of the self-energy shift can be seen dramatically in Fig. 2(e). The  $\pi(\text{NBO})$  and  $\pi^*(\text{NBO})$  states both have substantial character on the defect site and so are strongly influenced. The former moves up in energy and becomes a resonance just above the lone-pair-band minimum, while the latter moves deep into the gap, forming a state near midgap. This state is highly localized to the defect atom, having 68% of its character there and 21% on the neighboring site. Remarkably, there is a second state emerging into the gap just above the valence-band maximum. This corresponds to the terminal nonbonding  $\gamma$  orbital of Fig. 1(b) which is also severely influenced by the self-energy shift.

Consider now the threefold site, Fig. 2(f). Recall that the simple model of Fig. 2(c) gave rise to threefold-degenerate defect states above and below both the bonding and antibonding bands. Now taking the bond angles and dihedral angles different from  $90^\circ$  breaks the threefold degeneracy of these states. The  $\pi(\sigma^*)$  state in the fundamental gap, for example, splits into a nondegenerate (A-like) state deeper in the gap and a twofold-degenerate (E-like) resonance above the conduction-band minimum. The self-energy shift tends to move these states further downward in energy, but the effect is not dramatic because no state is primarily localized to the defect atom. The defect state occurs 0.55 eV below the conduction-band edge

and is highly delocalized, having no more than about 15% of its character on any one site.

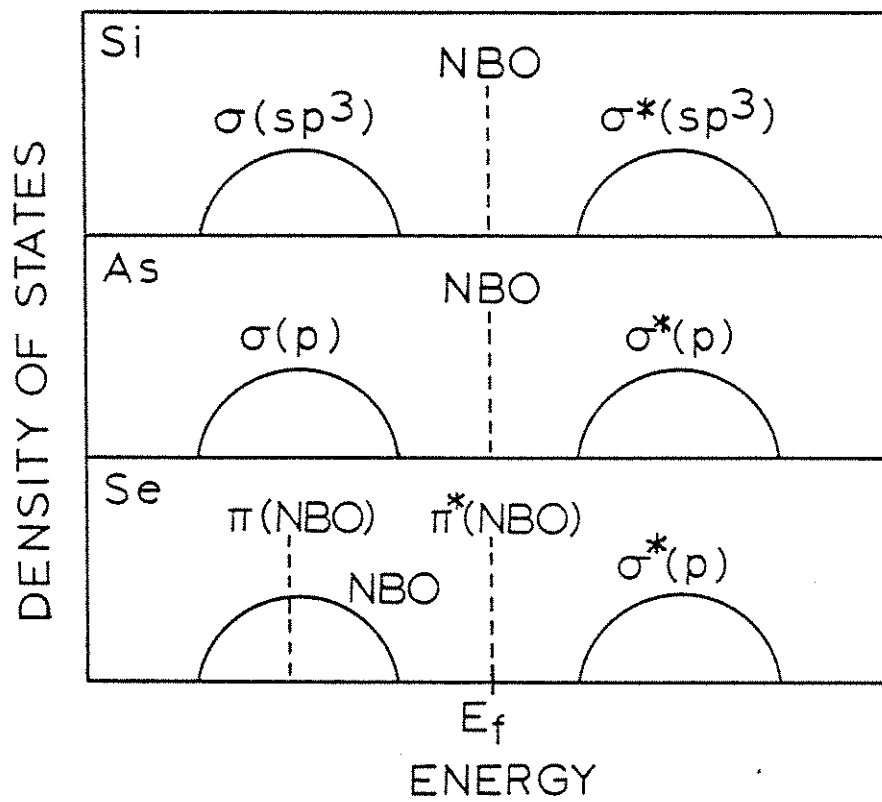
In order to appreciate the uniqueness of these defects in chalcogenides, consider for a moment the corresponding defects in pnictide or tetrahedral semiconductors, e.g., a-As or a-Si. First of all, overcoordinated sites are not expected at all in Si. If they exist in a-As, their properties would be determined by the exceptional  $sp_3$  hybridization at the defect site. Pollard and Joannopoulos<sup>33</sup> have shown that the central  $\sigma^*$  bond orbitals, composed of  $sp_3$  and  $p$  orbitals on the fourfold and neighboring sites, respectively, are split by the dehybridization interaction into a  $T_2$ -like resonance in the conduction band and an A-like gap state near the conduction-band edge. The gap state is essentially localized to the fourfold site and its immediate neighbors and has no  $p$  character on the defect site. This state is thus completely different from the overcoordinated a-Se defect state in origin, character, and symmetry. Thus the chalcogenides are the only systems where one finds such an elegant correspondence between the properties of under- and overcoordinated defects.

Secondly, even the undercoordinated defects are unique in the chalcogenides. The situation is sketched in Fig. 3. In a-As or a-Si, one finds neighboring  $p$  or  $sp_3$  orbitals split into  $\sigma$  and  $\sigma^*$  bond orbitals, which are then broadened into valence and conduction bands. A dangling bond simply gives rise to a  $p$  or  $sp_3$  orbital which is not bonded away by the  $\sigma$

Figure 3

Sketch of the gap region for undercoordinated defects in group IV, V, and VI semiconductors. The valence bands are shown schematically. Dashed lines correspond to defect states or resonances.





interaction and which, therefore, remains sitting near midgap. Because there are no such nonbonded orbitals on neighboring sites with which to interact, these defect states are initially strongly localized to the defect atom. Thus no self-consistent charge transfers or self-energy shifts are expected. This is in complete contrast to the case of the chalcogenides, where the availability of neighboring NBO's and the unique  $\pi$  interaction give rise to the rich structure described above.

Bishop, Strom, and Taylor,<sup>5</sup> have identified the photinduced electron spin resonance (ESR) center in selenium as a gap state localized primarily to a single nonbonding orbital. This fits in nicely with our picture of the onefold defect and indicates therefore that onefold defects may be preferred to threefold defects in the glass. This would be surprising, however, in the light of the suggestion by Kastner et al. that the lowest-energy neutral defect would be threefold coordinated.<sup>14</sup> This was based on the observation that the threefold defect has an extra bond compared to the onefold, and that consequently the electronic energy of bond formation would favor the threefold. From elementary models based on discrete bond-orbital energy levels, the energies of the onefold and threefold defects were estimated to be ~3-4 eV and ~1 eV, respectively.<sup>15</sup>

However, such an analysis omits the ion-ion repulsion which stabilizes the bonds and fails to take into account the broadening of molecular-orbital levels into bands and the

existence of gap states and resonances near defects. The calculated density of states in Fig. 2(e) or 2(f) contains all the information necessary to sum correctly the latter one-electron energies, and the additional repulsive term can be empirically modeled as a constant correction per bond. This is carried out in detail in Appendix A.

The resulting total energies of the onefold and threefold defects are estimated at 1.17 and 1.56 eV, respectively. Three effects are responsible for this remarkable lowering of the energy of the onefold site. Firstly, the ion-ion repulsion largely compensates for the loss of electronic binding energy when a bond is broken. Secondly, the hole which is constrained to lie at the NBO level in the simple models rises from midband to the valence-band and beyond in our calculation. Thirdly, the partial  $\pi$  bond between the defect site and its neighbor along the chain stabilizes this defect even further. The net result is that the onefold defect has its energy lowered by about 3 eV. The first and the second effects have analogs for the case of the threefold defect which raise and lower the total energy, respectively. The net result is a slight increase of  $\sim 0.5$  eV in the defect formation energy.

The energy estimates given here apply to neutral defects. In order to address such questions as the negative U or the luminescence Stokes shift, it will be necessary to identify the lowest-energy charged defects as well. On the one hand, it may be that the defect retains its basic bonding configuration upon

change of charge states, with only modest changes in bond lengths and bond angles (bond relaxations). On the other hand, the large Stokes shift would argue that a bond-switching or coordination can take place. This would be the case, for example, if the lowest-energy positively charged defect were threefold coordinated.

Because bond relaxations away from the crystalline bond lengths and angles could cause reductions in the neutral-defect energies (on the order of a tenth of an eV), and because of our rough model for  $\Delta U_0$ , the energies reported above are not final and merely point out the need for more realistic calculations. In particular, lower defect energies are needed if  $\sim 10^{17} \text{ cm}^{-3}$  defects are to be frozen in at the glass transition temperature. We have recently extended the self-consistent pseudopotential calculations to the evaluation of neutral-defect total energies and relaxation energies [see Sec. II(d)]. We find sufficiently low energies to give  $\sim 10^{15} \text{ cm}^{-3}$  coordination defects. However, here we wish to emphasize the inadequacy of simple models which are based on bond-orbital levels, or which neglect repulsive interatomic terms, and to point out the plausibility of having the onefold be the favored neutral defect in the glass.

So far, the discussion has been limited to individual onefold and threefold sites, the simplest structural defects. However, once the TB Hamiltonian has been established, it is straightforward to make the extension to more complicated

defects and interacting defect pairs. In particular we shall discuss nearest-neighbor onefold and threefold pairs ("intimate valence alternation pairs," or IVAP's),<sup>14</sup> the transition from threefold to onefold coordination, and dihedral-angle variations for the simple defects.

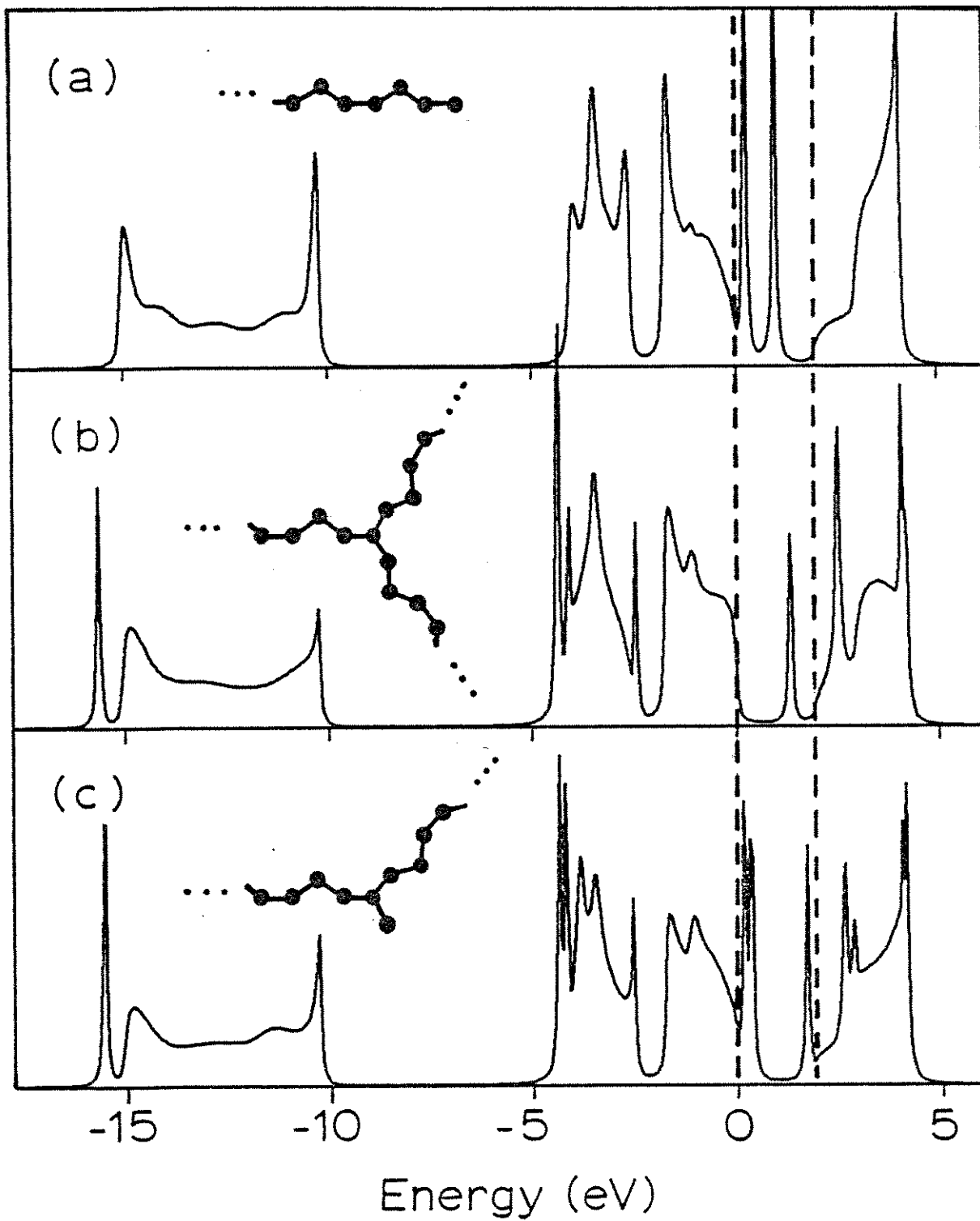
Because onefold and threefold defects are expected to be oppositely charged, the resultant Coulomb attraction will tend to favor close defect pairs.<sup>14</sup> The extreme form of this pairing is the IVAP shown in Fig. 4(c) [together with isolated onefold and threefold defects in figs. 4(a) and 4(b) for comparison]. Here the onefold and threefold defects are nearest neighbors. The defect states associated with the two defect sites remain, but are shifted closer to the band edges. For the onefold defect, this occurs because the strong direct  $\pi$  interaction which previously pushed the defect state deep into the gap, disappears due to the lack of an available NBO on the neighboring threefold site. Similarly, the electron trap state becomes more shallow in the IVAP because the strong direct  $\pi$  interaction connects only one pair of  $\sigma^*$  orbitals, rather than three, in the neighborhood of the threefold site.

For an overall neutral IVAP the defect states near the valence-band edge are occupied and the state near the conduction-band edge is empty. Thus, when a pair of isolated neutral onefold and threefold defects are brought together in this fashion, an electron transfer from the threefold to onefold state occurs. Surprisingly, all sites remain individually

Figure 4

Tight-binding density of states for (a) onefold defect, (b)threefold defect, and (c) IVAP. The structures are shown schematically. Dashed lines represent the gap edges.

D. O. S.



almost neutral. This is a remarkable consequence of the self-energy shifts, together with the fact that the defect states are not strictly localized to the defect site.

From experiments on transport properties during thermal cycling below the glass transition temperature, Abkowitz<sup>34</sup> has found a close correlation between a population of electron traps 0.33 eV below the conduction-band edge and hole traps 0.25 eV above the valence-band edge. In light of our results, these experiments can be plausibly explained by assigning both states to IVAP's, whose density in the glass presumably varies with temperature according to the free energy of defect formation.<sup>14</sup> The method of Appendix A gives a total energy of 1.22 eV for this defect. (Recall that the total energies of the onefold and threefold defects were 1.17 and 1.56 eV, respectively.) This indicates that the IVAP (which is really a defect pair) is comparable in energy to an isolated onefold or threefold site and may therefore occur at comparable densities. The energy of 1.22 eV is still too large to allow a sufficient number of defects to be frozen in at the glass transition, but it could be reduced substantially by minimizing with respect to bond relaxations in the defect vicinity. In any case, a quantitative comparison with the results of Abkowitz would require a more realistic calculation than we have performed.

An interesting feature of glassy Se is the possibility of interconversion of onefold and threefold defects. Figure 5



shows the density of states in the defect region as a onefold atom swings over and bonds with an atom of a neighboring chain, creating a threefold defect. As one creates the threefold defect, two NBO's interact and contribute a  $\sigma$  and  $\sigma^*$  state. One of these can be clearly seen moving from the lone-pair to the bonding band in Fig. 5. In the fundamental gap, however, the deep defect state simply changes from being an NBO-derived state to being a  $\sigma^*$ -derived state while shifting only very slightly in energy.

Finally, in Fig. 6 we show the energy location of the deep defect level as a function of dihedral-angle variations for isolated onefold and threefold defects. This case has been chosen as a likely example of possible relaxations at the defect site because the constraint on dihedral angle is not expected to be as strong as that on bond length or bond angle. Note that the onefold level is not strongly affected, while the threefold level is quite sensitive. In both cases the eigenvalue has a minimum near the crystalline dihedral angle. For neutral defects these levels are singly occupied, suggesting that the dihedral angle probably does not deviate greatly from the unrelaxed value.

### (c) Self-Consistent Pseudopotential Calculations

To provide a realistic foundation for the tight-binding Hamiltonian of Table I, we have carried out self-consistent

Figure 5

Change in the density of states as a onefold site bonds to a nearby twofold site.  $V_{12}$  is the onefold-twofold interaction in units of the normal nearest-neighbor interaction. (a) Onefold defect and intact chain, (b) weak interaction, (c) moderate interaction, and (d) threefold defect.

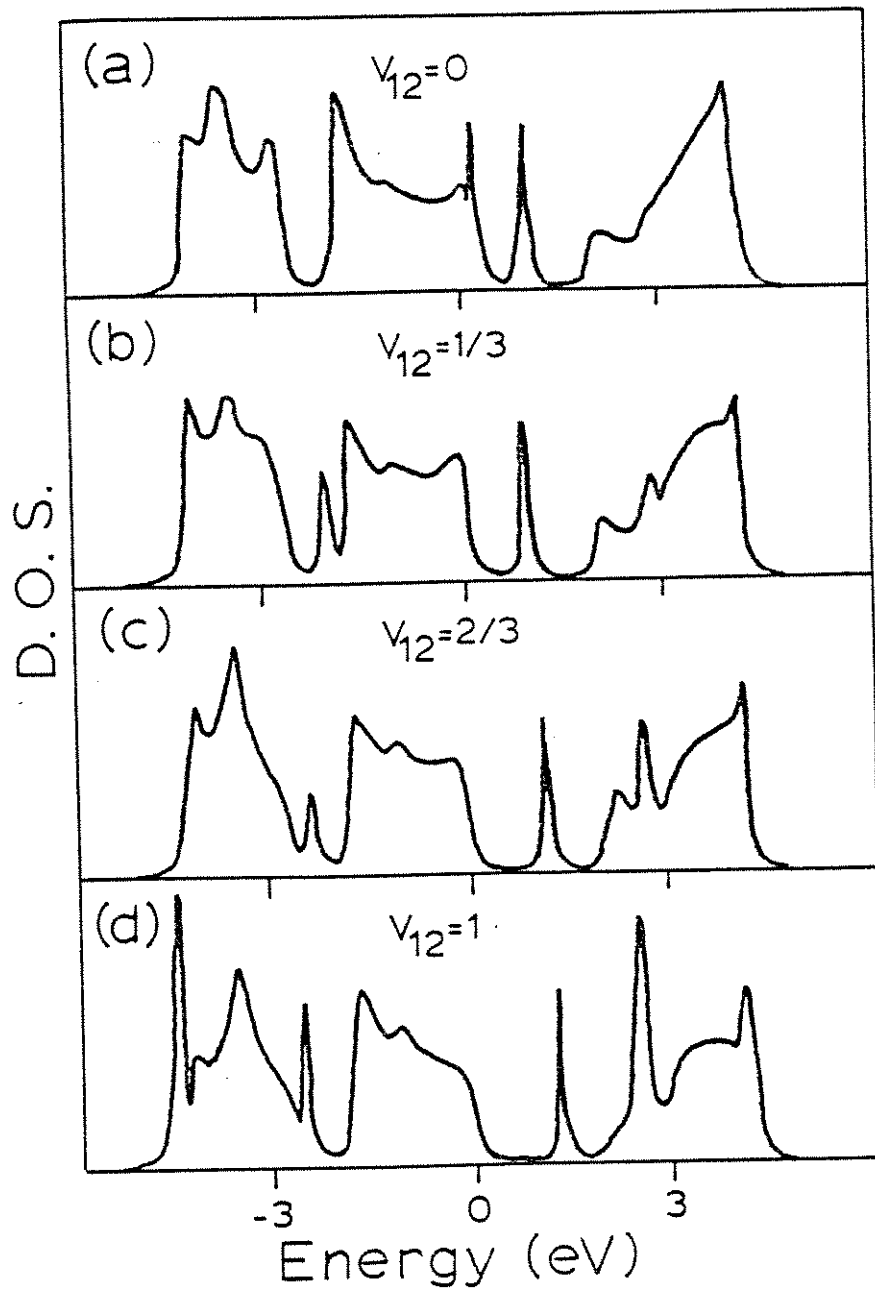
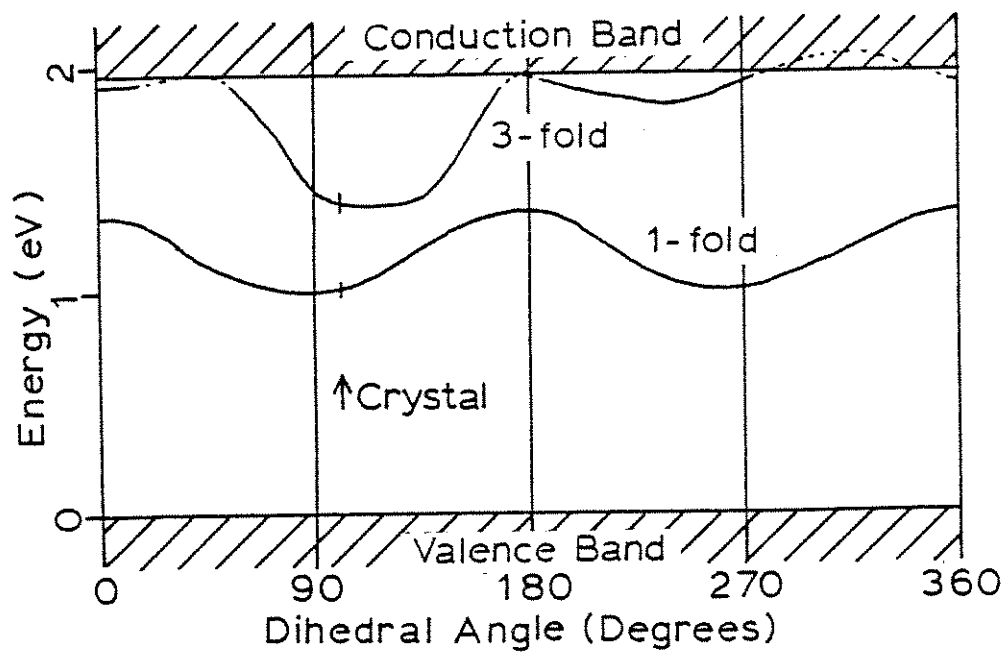


Figure 6

Energy levels of defect gap states as a function of dihedral angle on (i) the next-to-last bond for the onefold defect, and (ii) the three bonds connecting to the defect site for the threefold defect.



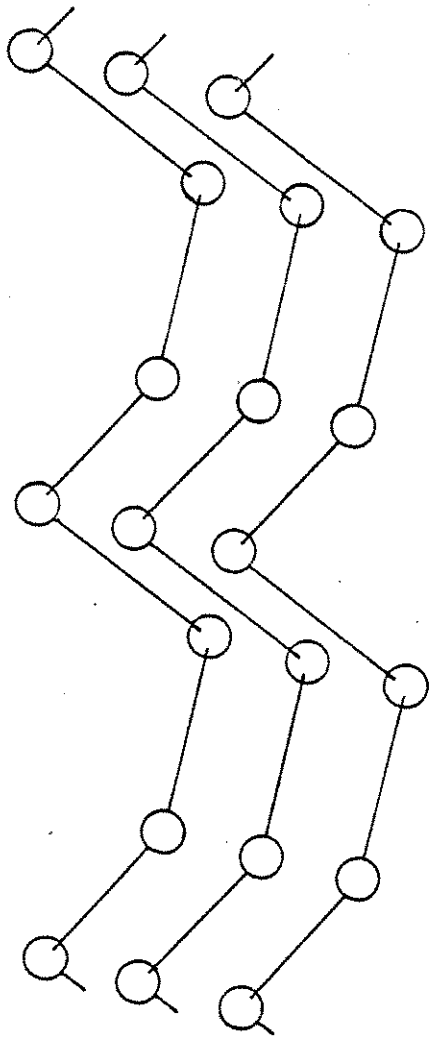
pseudopotential (SCPSP) calculations on a variety of periodic structures containing bond-coordination defects. We describe the structures which have been solved, briefly discuss the method, review the results of these calculations, and present the fitting of the TB Hamiltonian to the SCPSP results. A number of details regarding the method have been deferred to Appendix B.

In Fig. 7 we show trigonal Se and two periodic structures containing defect pairs. The defect structures have unit cells containing six atoms, instead of the usual three, and are made simply by adjusting some of the atomic positions so as to break bonds or form new bonds. The structures are chosen so that all of the bond lengths and bond angles match exactly those of trigonal Se (which we take to be  $2.390 \text{ \AA}$  and  $102.48^\circ$ , respectively).<sup>31,32</sup> In model A, each unit cell contains a onefold defect and a threefold defect as third neighbors along a chain; model B represents the IVAP. We have also solved structures, shown in Sec. II(d), which contain 11 atoms per unit cell such that each cell contains two identical onefold defects (or two identical threefold defects). The latter results merely verify that the coexistence of onefold and threefold defects together in model A has no influence on the important physics (e.g., the existence of self-energy shifts). For our purposes here, it will be sufficient to focus only on a comparison between the results for the trigonal crystal and for model A.

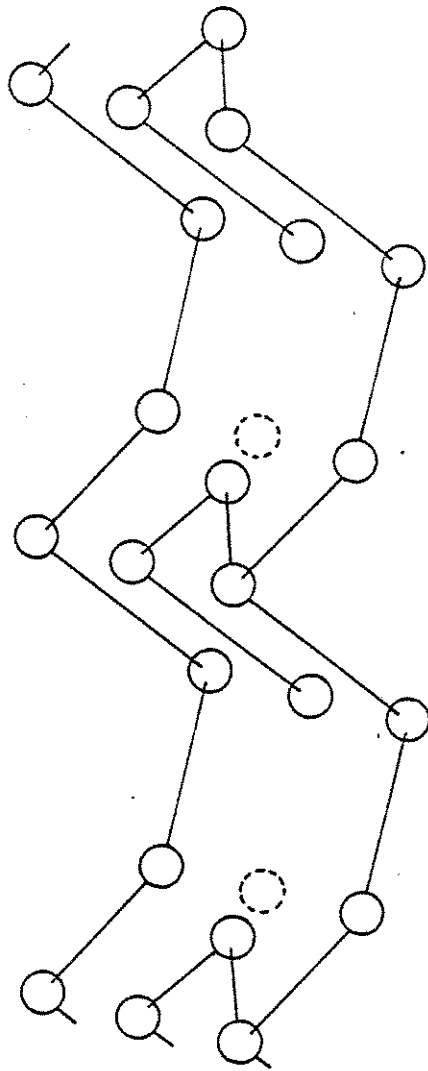
Figure 7

Periodic structures to which the self-consistent pseudopotential method (SCPSP) can be applied. Dotted circles represent unperturbed lattice positions.

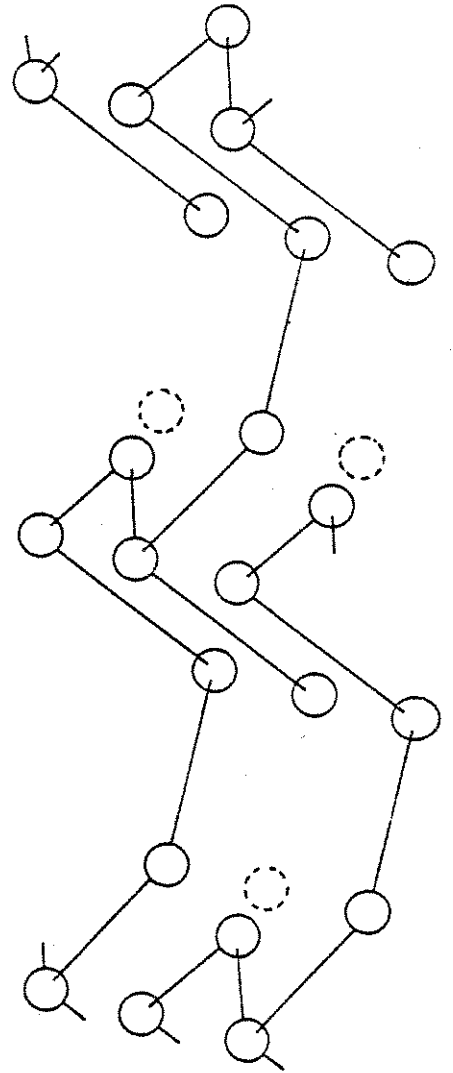
TRIGONAL SE



MODEL A



MODEL B



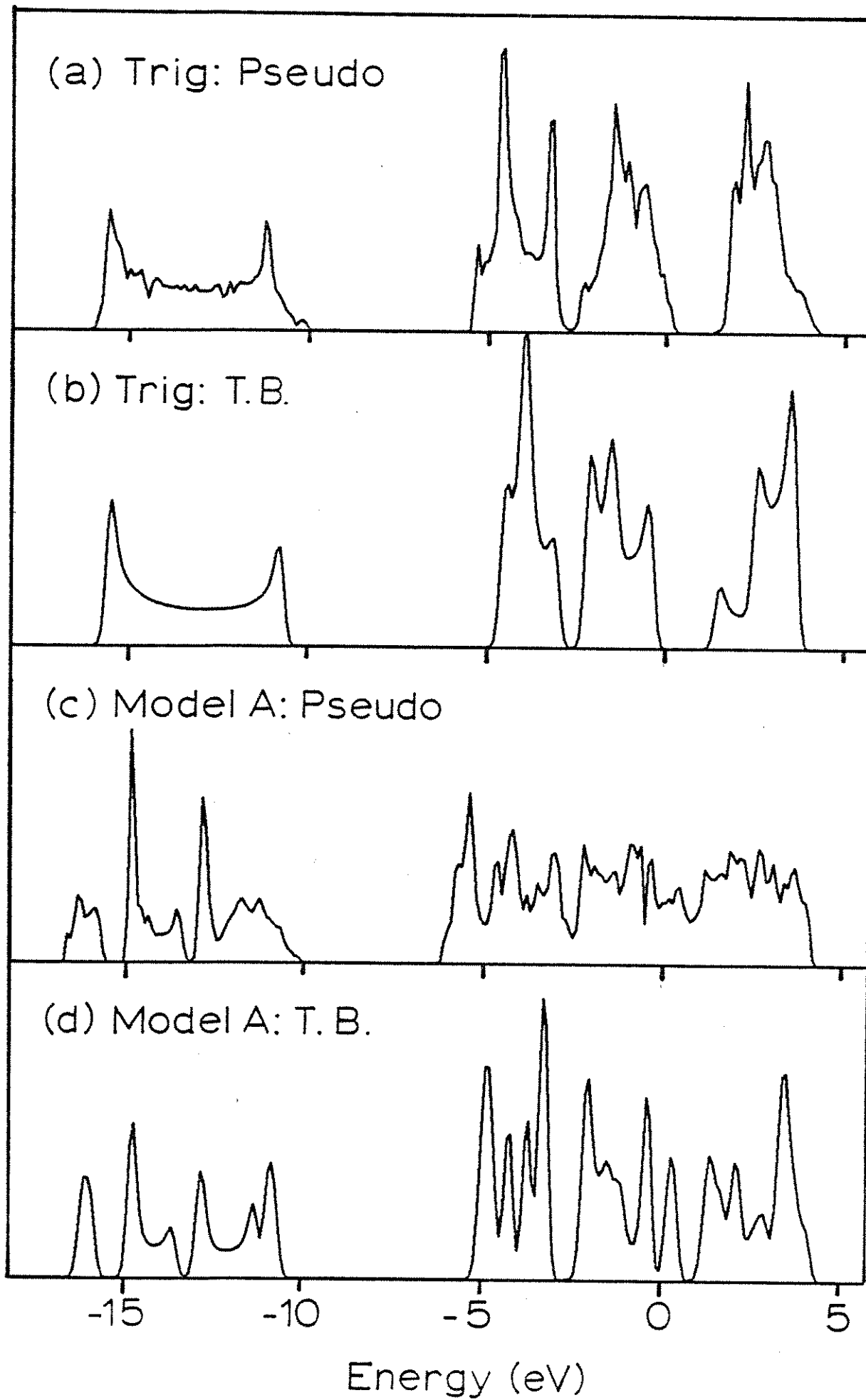


The self-consistent pseudopotential approach has been described in detail elsewhere.<sup>35</sup> Here the selenium cores are represented by a local ionic pseudopotential which has been chosen once and for all such that when self-consistently screened, the valence levels match the published Herman-Skillman eigenvalues for the atom.<sup>36</sup> (Throughout this section, screening is calculated by using Slater  $\rho^{1/3}$  exchange, with a coefficient  $\alpha=1$ , for the exchange-correlation term. For a discussion of this choice of screening, and the choice of local vs. non-local pseudopotentials, see Appendix B.) Once fit to the atom in this way, the pseudopotential is used without modification for all subsequent crystal calculations; in this sense, these are first-principles calculations.

The crystalline pseudopotential is just the sum of the ionic pseudopotentials on each atomic site plus the screening from the valence electrons. These electrons are represented using a plane wave basis. All plane waves with  $q < 3.67 \text{ \AA}^{-1}$  have been included, and those with  $q < 5.93 \text{ \AA}^{-1}$  have been included in Lowdin perturbation theory<sup>37</sup> when solving for the wave functions. For trigonal Se, this corresponds to 70 and 300 plane waves, respectively; for model A, 140 and 600. The charge densities are averaged over the Brillouin zone using Chadi's special-points scheme,<sup>38</sup> with 3 and 8 k-points in the irreducible Brillouin zone for trigonal Se and model A, respectively.

Figure 8

Theoretical electronic density of states for periodic structures: (a) SCPSP applied to trigonal Se, (b) tight-binding fit for trigonal Se, (c) SCPSP applied to model A, and (d) tight-binding fit for model A.



The density of states for trigonal Se is shown in Fig. 8(a). One finds filled s, p bonding, and p lonepair bands, and an unfilled p antibonding band. In Fig. 8(b) we show the density of states for the tight-binding Hamiltonian of Table I. The fitting has been done by artificially reducing the SCPSP band structure to a 1-d band structure  $E=E_n(k_{\parallel})$ , by averaging over all  $k_{\perp}$  for each  $k_{\parallel}$  (the orientations are with respect to the chain axis). The TB Hamiltonian was then adjusted to match the band edges of the 1-d band structure as closely as possible. As can be seen from Figs. 8(a) and 8(b), this results in satisfactory agreement between the full SCPSP and TB densities of states.

Figure 8(c) shows the SCPSP density of states for model A. Because of lowered symmetry, the s-like band has broken into three subbands. In the p region, gap states have been broadened into defect bands (by interdefect interactions), and the gaps have thus been obscured. In Fig. 8(d) we show the corresponding TB calculation on the same structure. Now the self-energy shifts in the Hamiltonian of Table I come into play. These were chosen by actually integrating in real space to calculate the diagonal matrix elements of the converged self-consistent potential between atomic (Herman-Skillman)<sup>36</sup> orbitals centered on each atomic site. The calculated self-energy shift on the onefold site was 1.27 eV, on the threefold site -1.24 eV, and the other four sites ranged from -0.38 to 0.28 eV. For simplicity we have ignored the

self-energy shifts on twofold coordinated sites and approximated those on the defect sites as  $\pm 1.25$  eV. A comparison of Figs. 8(c) and 8(d) indicates that this fitting is adequate. Note, in particular, the agreement in the s region; without the self-energy shifts, there was little resemblance here between the SCPSP and TB results.

An approximate measure of the charge on each site in model A was obtained by integrating the SCPSP valence charge density over a sphere, with radius equal to half the bond length, centered about each atomic site. (Such an estimate is somewhat arbitrary, but should at least reveal any large charge transfers.) Each site was found to be individually neutral to within  $\sim 0.06$  e. This is surprising because a charge transfer from the threefold to the onefold site would have been expected on the basis of simple chemical arguments. The lack of charge transfer is a result of the compensating effect of the self-energy shifts, which tend to attract electrons to the threefold site and repel them from the onefold site.

#### (d) Total Energies

As pointed out in Sec. II(a), there has remained considerable controversy within the defect model, regarding the nature of the lowest-energy neutral defects.<sup>12,14-16,23-24</sup> Since Se is normally doubly coordinated, the interesting defects are the 1-fold (dangling bond) and 3-fold sites. Here, we resolve this controversy by performing self-consistent

pseudopotential (SCPSP) calculations of the total energies of various possible bonding configurations for the neutral defect.

This pseudopotential total energy calculation is the first of its kind to be applied to defects. Initial estimates<sup>14</sup> of defect energies were based upon simple models which considered only discrete bond-orbital energy levels. The more sophisticated approach of Sec. II(b), which sums over all filled band and gap states and includes repulsive intercore terms in an empirical fashion, demonstrates the inadequacy of the simple models. It becomes evident that a truly realistic calculation is needed. Recently, there have been promising reports in the literature indicating tentatively that a SCPSP approach could be developed into a viable tool for calculating total energies, and thereby equilibrium bond lengths, relaxation energies, and vibrational frequencies, in bulk and surface silicon.<sup>39-44</sup> We have therefore developed a SCPSP scheme to calculate defect energies in Se. To examine the accuracy of our approach we must first study trigonal Se as a test case.

We choose a local ionic pseudopotential so that the energies and wave functions of the Se atom are very accurately reproduced.<sup>45</sup> It is then used without modification for all subsequent self-consistent crystal or superlattice calculations. Screening is computed using Slater  $\rho^{1/3}$  exchange-correlation, with a coefficient of  $\alpha=2/3$ . For a fur-

ther discussion of the choice of pseudopotential and exchange-correlation, refer to Appendix B.

The total energy is calculated as a sum of three terms. These are: the sum of filled one-electron levels, the correction for Hartree and exchange-correlation overcounting, and the repulsive interaction between ion cores (see Appendix B). The latter two terms are crucial in order to obtain correct and meaningful results. To determine the lowest energy (equilibrium) configuration of trigonal Se we have studied a variety of structural parameters. The results are very satisfying, giving a bond length of 2.33 Å, a bond angle of 104.5°, and a frequency for the Raman active  $\Gamma_1(A_1)$  phonon mode of 204 cm<sup>-1</sup>. These are to be compared with experimental values of 2.39 Å, 102.4°, and 235 cm<sup>-1</sup>, respectively.<sup>32,46</sup> If the energy of the isolated constituent atoms is calculated within the same formalism for comparison we find the binding energy of the crystal to be 3.60 eV/atom. However, the isolated atoms can lower their energy by relaxing into the spin-polarized <sup>3</sup>P state (Hund's rule), gaining ~0.76 eV each (see Appendix B). The resulting crystal binding energy is 2.84 eV/atom. in rough agreement with the experimental value of 2.25 ±10 eV.<sup>47</sup> The calculated equilibrium configuration is used as the unperturbed structure for calculations of defect energies.

The extension to the study of defects is made by choosing a supercell, or a large unit cell, containing the defect of interest. The choice of unit cell is dictated by the necessity

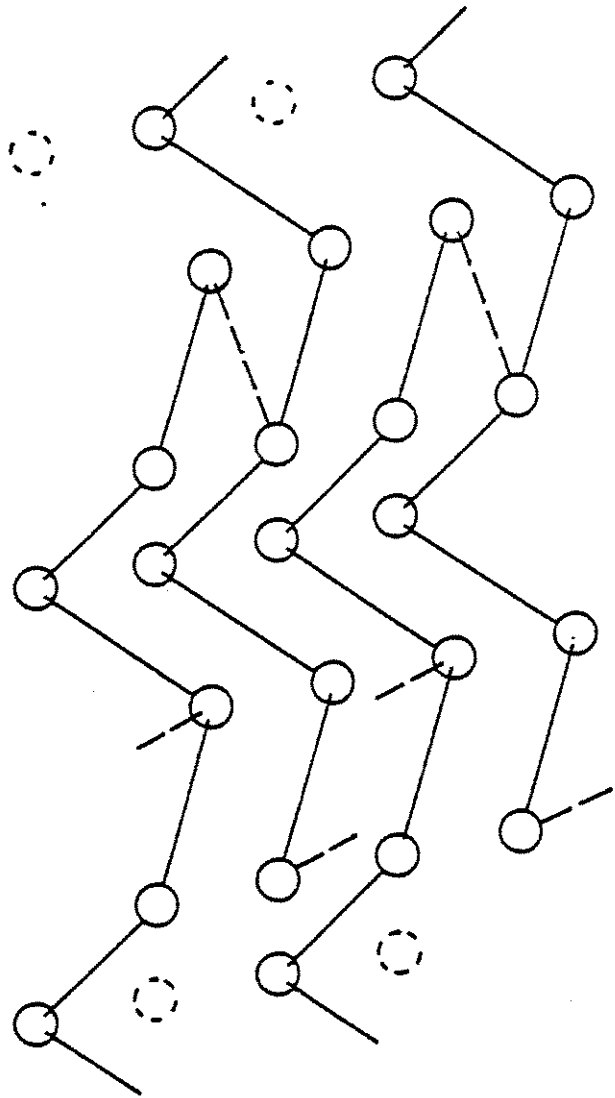
of keeping the defects as far from each other as possible (to minimize interactions); maintaining the crystal bond length and bond angle throughout; and keeping the cell small enough or symmetrical enough to make the calculation tractable. This is accomplished for trigonal Se by removing every twelfth atom of each chain in such a way that the positions of the resulting vacancies in neighboring chains are staggered. There are now two 1-fold defects in each unit cell. If each 1-fold atom is allowed to swing over and bond to an atom of a neighboring chain, the structure of Fig. 9 results containing two 3-fold defects per unit cell. Finally, the total energy of a single defect is calculated as half the energy per unit cell of the super-lattice structure, minus the energy of the same number of normally coordinated atoms as given by an identical calculation on the trigonal structure.

The relevance of such a structure to the case of defects in the glass might be questioned on two accounts. The first concern is that the super-lattice structures contain a high density of defects, and the interactions between defects cannot be eliminated entirely. Such interactions can take two forms: direct interactions caused by overlap of atomic orbitals on different sites, and effective interactions which propagate along chains. However, in neither structure do the defects ever lie closer than the crystal third neighbor distance (1.59 bond lengths), nor are they ever separated by fewer than three intervening atoms along a chain. The average splitting of the



Figure 9

Superlattice structure containing two 3-fold defects per unit cell. Starting with the trigonal structure, atoms are removed (dashed circles) to form 1-fold defects, and then extraordinary bonds are formed (dashed lines). Some of the extraordinary bonds connect to atoms in neighboring layers, not shown.

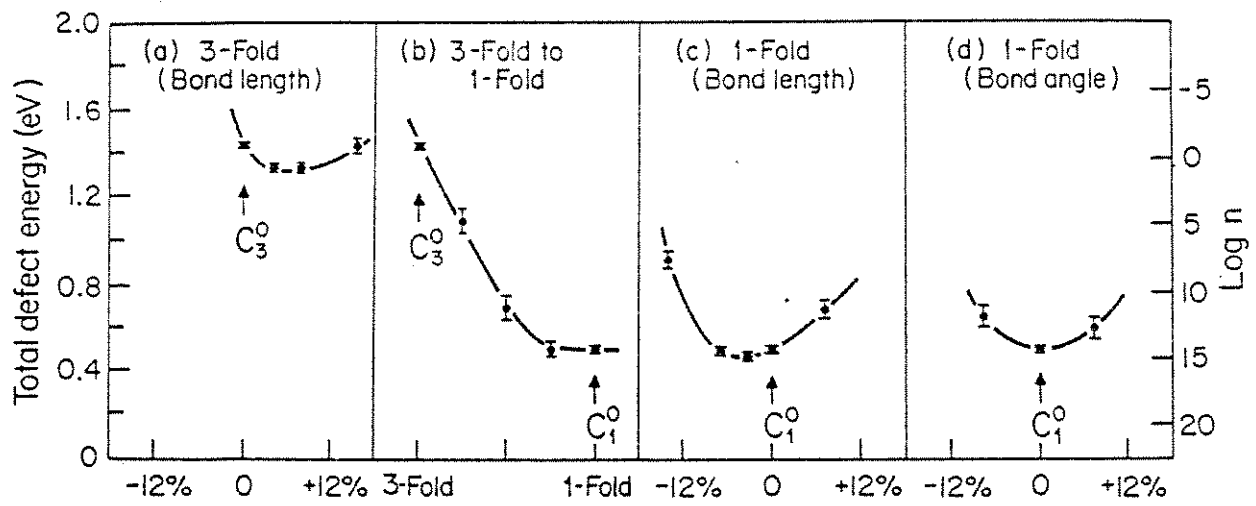


gap levels due to these interactions is  $\sim 0.3$  eV, and such levels are summed over, eliminating to first order any error from this source in the total energies. The second concern is that the super-lattice structures derive from a vacancy in the trigonal crystal, rather than a truly glassy structure. However, the local chemistry is expected to be the principle factor in determining the total energy, so that these models can be taken as giving reasonable estimates for defect energies in the glass.

In Fig. 10(a) we show the energy of a single 3-fold defect as the lengths of the three bonds adjoining the defect are varied equally. We find a minimum when the bonds have been extended by  $\sim 6\%$ , at a defect energy of  $\sim 1.3$  eV. However, consider Fig. 10(b) where only one bond is lengthened, to study the transition from a 3-fold to a 1-fold defect. We notice that the 3-fold defect is not even metastable, but immediately relaxes to the 1-fold configuration via the breaking of one of the three bonds. Moreover, we find no evidence for the proposed<sup>23</sup> energy minimum midway between the 1-fold and 3-fold configurations. These are perhaps the central results. The energy gained is a very substantial 0.8 eV. Figure 10(c) and 10(d) show that relaxations of the 1-fold defect play a very minor role, saving an extra 0.03 eV by a 4% contraction of the bond adjoining the defect site. They also verify that the relaxed 1-fold defect is indeed at a local energy minimum in configuration space.

Figure 10

Calculated total energies for neutral defects relative to the normally coordinated bulk. (a) 3-fold defect with bond length relaxations. (b) Bond weakening transition from 3-fold to 1-fold. Abcissa is length of dashed bonds in Fig. 9; at point marked 3-fold (1-fold) it is crystal first (second) neighbor distance. (c) Length relaxation of the last bond of the chain, and (d) relaxation of the bond angle between the last two bonds of the chain, for 1-fold defect.  $C_3^0$  and  $C_1^0$  correspond to unrelaxed 3-fold and 1-fold configurations (i.e., with trigonal Se bond lengths and bond angles). Scale at the right gives equilibrium defect density in  $\text{cm}^{-3}$  at  $T_g$ .



Of special note is the energy of the 1-fold defect, only  $\sim 0.5$  eV. Considering that previous estimates have ranged as high as  $\sim 3-4$  eV,<sup>15</sup> this is extraordinarily low. Two factors are responsible for this difference. Firstly, the repulsive energy of inter-core interactions is reduced when breaking a bond. Secondly, an inspection of the defect states reveals a subtle point: the dangling lone-pair electron on the 1-fold site can form a partial yet strong  $\pi$ -bond with the lone-pair electrons of the neighboring atom along the chain. This effect can lower the energy by as much as 2 eV. Moreover, one might also expect this to occur for positively charged defects. Thus, even the conventional wisdom that the positively charged defects are overcoordinated may break down in these materials. Finally, it is interesting to note that the low energy to create the 1-fold defect is also in line with recent magnetic susceptibility measurements on liquid Se, which indicate an activation energy of 0.66 eV for creating unpaired spins in the melt.<sup>48</sup>

The scale at the right of Fig. 10 converts defect energies to densities,<sup>14</sup> assuming defects are frozen in at the glass transition temperature  $T_g = 310^\circ\text{K}$ . This gives  $\sim 10^{15}$   $\text{cm}^{-3}$  neutral dangling bonds (and therefore spins) in the material. Of course, if the system exhibits an "Anderson negative U"<sup>7</sup> as proposed in the defect model, charge transfer between defects would occur, removing the neutral defects and creating larger numbers of charged defects (VAP's).<sup>14</sup> Whether this occurs, and

what density of diamagnetic defects would result, are questions which await the formidable task of applying the above techniques to charged defects. In any case, our calculation sets a lower limit on the number of bond-coordination defects, thereby supporting the suggestion that structural defects exist in high densities in these materials<sup>12,14</sup> and lending plausibility to the defect model.

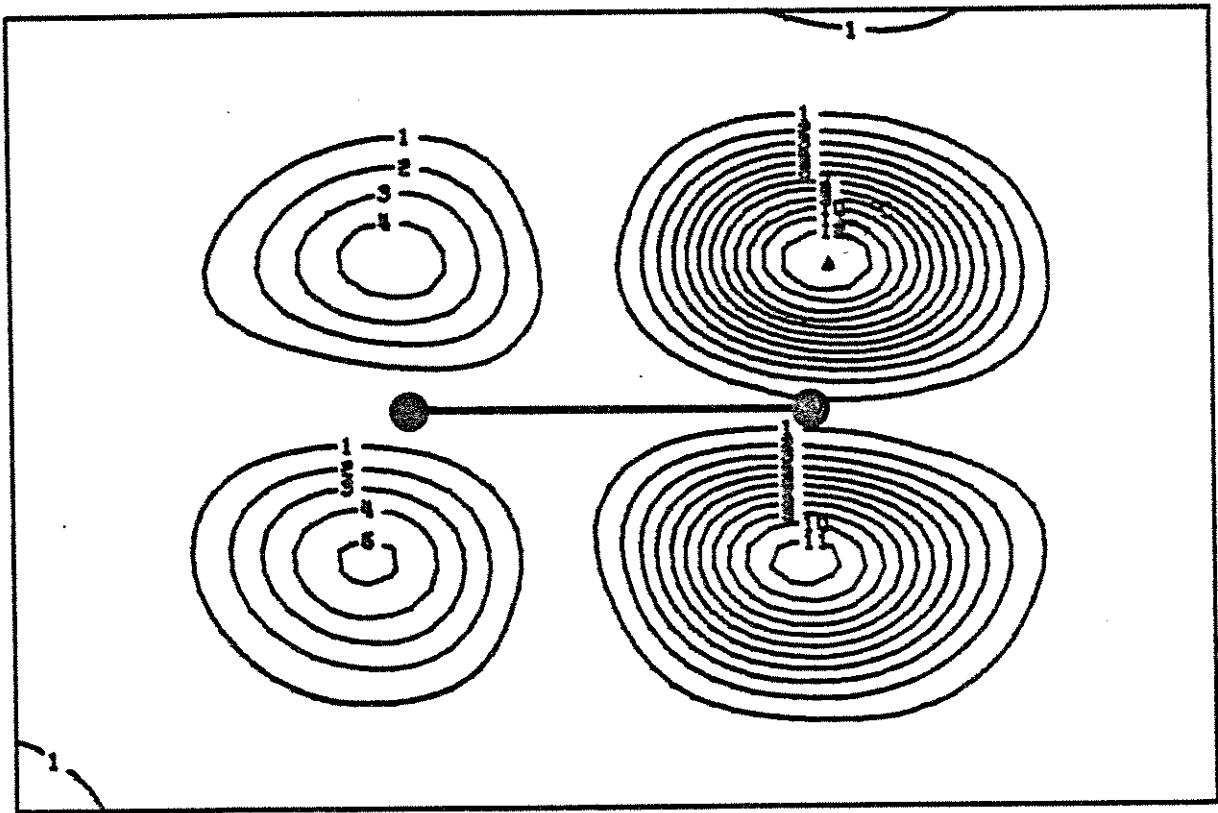
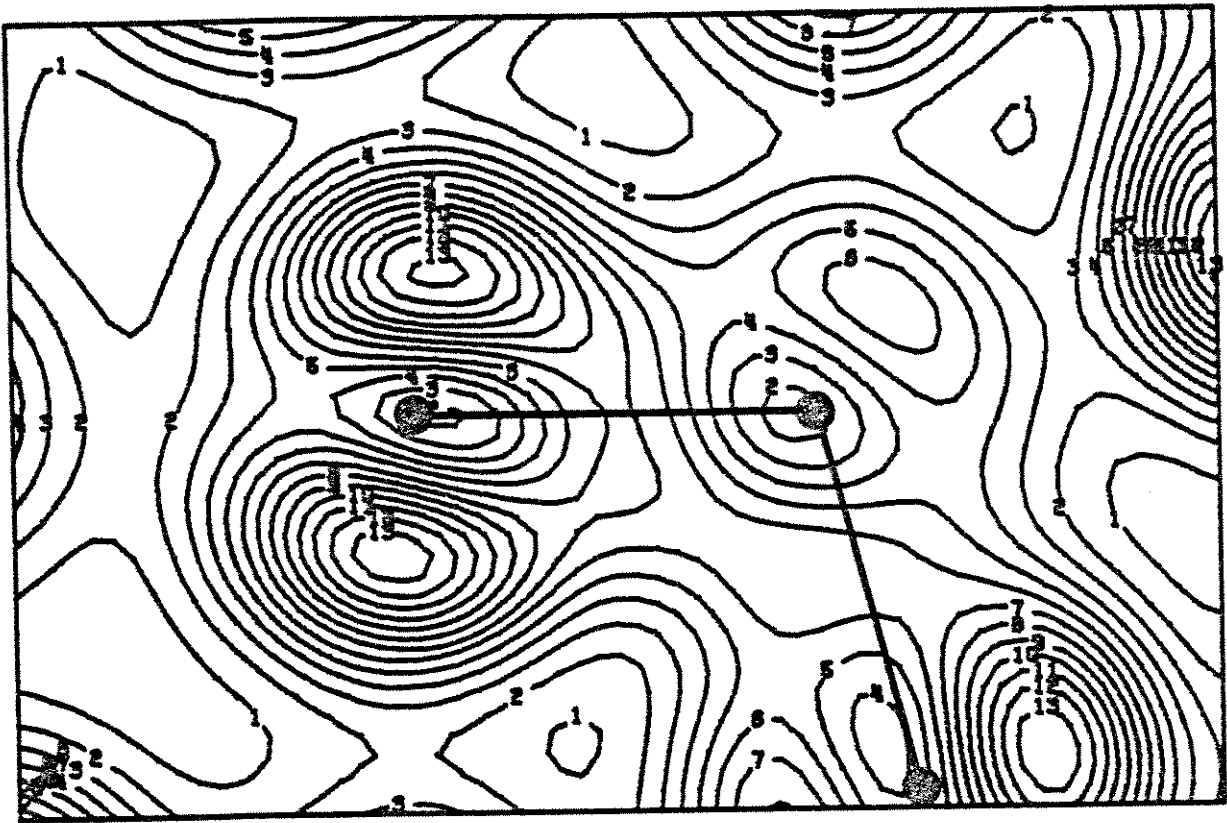
The super-lattice calculations by themselves are not a sensitive probe of the electronic gap states associated with defects, because of the inevitable interactions between defects in neighboring cells. However, the present calculations are consistent with, and lend support to, the previous work which has been detailed in Sec. II(b) and (c). The location of defect bands in the gap, and the self-energy shifts found at defect sites, are essentially unchanged. The bond contraction relaxation at the 1-fold site strengthens the lone pair  $\pi$ -interaction, but lowers the self-energy shift of the 1-fold atom by  $\sim 0.3$  eV. The result of this competition is that the defect band shifts very slightly deeper ( $\sim 0.06$  eV) into the gap.

Finally, it is straightforward to extract charge density information from the converged SCPSP calculation. This is done in Fig. 11 to illustrate the nature of the gap state associated with the onefold defect. It clearly confirms our earlier finding that this defect state is a  $\pi$ -antibonding combination of lone-pair orbitals, with most of the weight on the defect site.

Figure 11

SCPSP charge densities. (a) Lone pair band of crystalline Se, plotted in a plane containing three atoms along the chain, for reference. (The non-bonding p-orbital of the central atom is not visible because it is directed normal to the plane of the plot.) (b) Defect band for onefold defect of Fig. 9. Geometry is the same as in (a), except lower-right atom has been removed to create the vacancy.





(e) Summary and Conclusions

We have developed an approach to the study of structural defects in chalcogenides which makes use of flexible tight-binding techniques while remaining grounded in realistic self-consistent pseudopotential calculations. A variety of defect structures are studied. The neutral onefold coordinated defect gives rise to a deep hole trap state near midgap. This state is highly localized near the defect site, as it derives from a  $\pi$ -antibonding combination of nonbonding orbitals on the defect site and its neighbor along the chain. The neutral threefold defect produces a less localized but nonhydrogenic nondegenerate electron trap state below the conduction-band edge, which derives from a  $\pi$ -bonding combination of antibonding orbitals. The IVAP gives rise to both of these defect states, but they are closer in energy to the band edges.

The results emphasize the unique nature of defects in chalcogenides as opposed to other amorphous semiconductors. In particular, the availability of nonbonding orbitals on sites neighboring the defect atom allows the formation of an anomalous  $\pi$  bond at the undercoordinated defect and gives rise to the positive self-energy shift. We find a remarkable degree of analogy between the behavior of the onefold and threefold defects. The latter has an anomalous  $\pi$  interaction between antibonding orbitals which gives rise to an electron trap and a negative self-energy shift. This parallelism can occur only in the chalcogenides, where overcoordinated defects can form

freely because no s-p hybridization is required. We expect that defects in a-As, a-Si, and even a-As<sub>2</sub>Se<sub>3</sub> will be quite different in electronic structure than those in Se.

Finally, we have developed an accurate pseudopotential approach to the calculation of total energies in trigonal Se. We then extended this method to the calculation of defect energies in glassy Se by choosing super-cells to contain structural defects in relative isolation. To our knowledge, this is the first calculation of structural defect energies in a semiconductor using an approach of this kind. We find the 1-fold defect to be lowest in energy, and the 3-fold defect to be non-existent.



### CHAPTER III ARSENIC SELENIDE

#### (a) Introduction

Theory and experiment have not yet combined to give an unambiguous picture of the defects responsible for the properties of chalcogenide glasses. In large measure, this is due to the fact that the experiments are usually more conveniently done on As-Se and As-S glasses, whereas the theory tends to focus on the simpler case of glassy Se. Experimentally, some of the properties of heteropolar and homopolar glasses appear similar, and the distinction between the two is often blurred. However, there are important differences. For example, both the photoluminescence efficiency and photoinduced ESR intensity are much lower in Se than in  $\text{As}_2\text{Se}_3$ , and both experiments show quite different dopant dependencies in the two cases.<sup>49</sup> Moreover, it is now becoming clear that the properties of individual defects in  $\text{As}_2\text{Se}_3$  and Se are expected to be quite different. As we saw in Chapter II, the coordination defects in glassy Se have deep gap states whose origin can be traced to an anomalous  $\pi$ -bonding mechanism which does not carry over to the case of heteropolar glasses. Thus the corresponding deep gap states are not expected in  $\text{As}_2\text{Se}_3$ . On the other hand, the heteropolar glass allows for anomalous bond orbitals (like-atom bonds and nonbonding As orbitals) which will give rise to deep gap states having no counterpart in pure Se.<sup>50</sup>

Clearly the two kinds of systems must be treated quite differently.

In this paper, we will attempt to bridge some of these gaps by applying the ideas developed in the study of glassy Se to the case of  $\text{As}_2\text{Se}_3$ . Because the structures are more complicated and the defect possibilities more numerous in  $\text{As}_2\text{Se}_3$ , serious theoretical calculations are still in their infancy. Nevertheless, by using some simple models, and arguing by analogy with the case of Se where appropriate, we will be able to say quite a bit about the nature of intrinsic structural defects in the heteropolar glass. The discussion will be carried out using  $\text{As}_2\text{Se}_3$  as a model system, but many of the conclusions may be applicable to sulfide or telluride glasses as well.

In Sec. III(b), we discuss various methods for classifying defects, and attempt to identify those defects which are most likely to occur. Section III(c) provides a detailed discussion of the electronic structure of the various defects, with an emphasis on similarities and differences with respect to Se. In Sec. III(d) we present a speculative discussion of defect total energies, and discuss the structural equilibrium at  $T_g$  which determines the defect densities in the glass. Finally, we summarize our results in Sec. III(e).

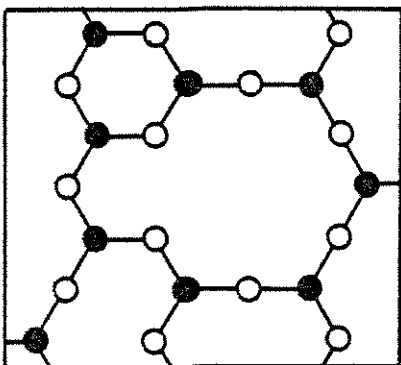
#### (b) Classification of Defects

In order to introduce the concept of defects, we must first define what is meant by a glass that has no defects. To visual-

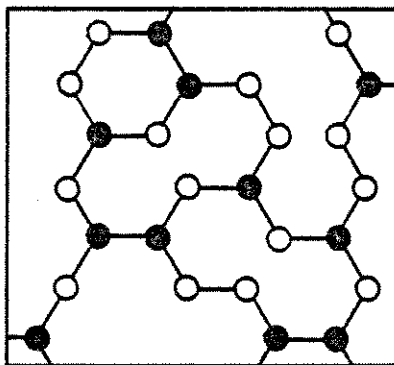
Figure 12

Schematic diagram of a random network with (a) no defects, (b) many like-atom bonds, (c) many malcoordinated atoms but no like-atom bonds, (d) many malcoordinated atoms but no like-coordinated-atom bonds. Solid circles represent As, open circles Se.

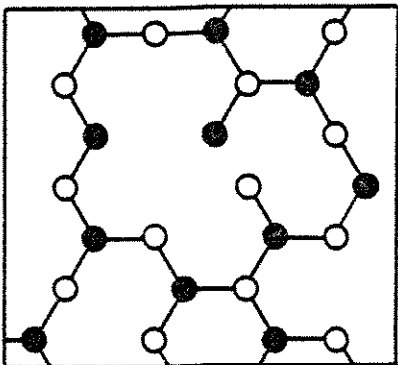
(a)



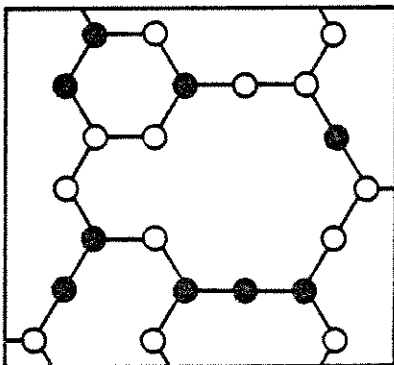
(b)



(c)



(d)





ize such a perfect glass, consider a continuous random network such as the one shown in Fig. 12(a). Every As atom has its preferred threefold coordination satisfied, every se atom has its preferred twofold coordination satisfied, and every bond is heteropolar. We will adopt the point of view that the perfect glass is the lowest-energy disordered structure, and that every deviation from the perfect-glass structure costs energy.

Such deviations take the form of various point defects in the glass. Figure 12(b) shows an amorphous network in which As and Se atoms retain their preferred coordination, although a large number of "wrong<sup>21</sup>" or like-atom bonds (LAB's) occur. On the other hand, Fig. 12(c) depicts a structure containing many malcoordinated<sup>16</sup> atoms (MCA's), but no LAB's. If the system tries to lower its energy by minimizing the number of LAB's and MCA's present,<sup>51</sup> then we expect the material to consist of a bulk network having the perfect-glass structure, interrupted occasionally by a defect containing one or more LAB's or MCA's.

A sampling of possible defect configurations is given in Fig. 13. The defects have been categorized according to the number of MCA's and LAB's they contain. Thus, our energy minimization principle implies that the defects in the upper left-hand region of the chart are the most likely ones. We have developed a notation which uniquely designates each defect; these are shown to the right of each configuration. The letters C and P correspond to chalcogen and pnictide, respectively, and the subscript denotes the coordination, fol-

Figure 13

Defect topologies organized according to number of malcoordinated atoms (marked by arrows) and number of like-atom bonds. Solid circles As, open circles Se. External bonds are understood to connect to a chalcogen atom of the bulk glass structure.

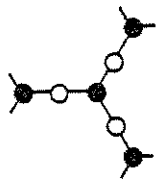
# Number of Mal-Coordinated Atoms

0

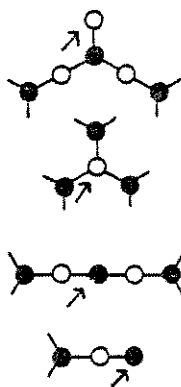
1

2 (IVAP)

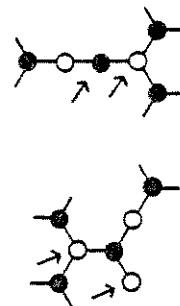
0



Bulk

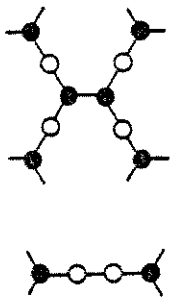


$C_1$   
 $C_3$   
 $P_2$   
 $P_1$

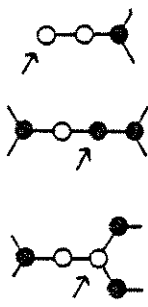


$P_2 C_3$   
 $C_1 P C_3$

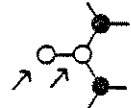
1



$P_3'$   
 $C_2'$



$C_1'$   
 $P_2'$   
 $C_3'$



$C_1' C_3'$

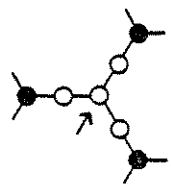
2



$P_2''$

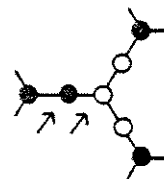
(SUBSTITUTION)

3



$C_3'''$

(SUBSTITUTION)



$P_2' C_3''$

(EXCHANGE)

NUMBER OF LINKS - POINT BONDING

lowing many previous authors. Each superscript prime indicates one LAB connected to that site, and multiple MCA's are represented using a molecular notation for the smallest cluster of sites which includes all the MCA's.

Obviously there are many more possibilities for defects than are shown in Fig. 13; we have attempted to include the simplest and most interesting ones. The most elementary defects are the single MCA's ( $C_1$ ,  $C_3$ ,  $P_2$ ,  $P_1$ ) and the single LAB's ( $P_3'$ ,  $C_2'$ ). ( $P_4$  has been left out because the s-p hybridization puts it in a class by itself.) The simplest close defect pairs ("intimate valence alternation pairs" or IVAP's)<sup>14</sup> are the nearest-neighbor pairs  $P_2C_3$  and  $C_1'C_3'$  and the second-neighbor pair  $C_1P_3C_3$ . If a Se is substituted for an As (or vice versa) a  $C_3'''$  (or  $P_2''$ ) center results; an interchange of neighboring Se and As atoms gives rise to a  $P_2'C_3''$  IVAP. Phillips's outrigger contains a  $C_2''$  (Ref. 19).

Note that the substitutions and interchanges have the property of conserving the structure of the bulk glass, as can be seen by comparing Figs. 12(a) and (d). In particular, they allow the system to avoid bonds between like-coordinated atoms. (That is, twofold coordinated sites are always bonded to threefold sites, and vice versa.) We shall use the term like-coordinated-atom bond (LCAB) for a bond connecting two sites with the same coordination number, in analogy to like-atom bonds (LAB's). If it is energetically favorable for the glass to avoid LCAB's, we say the LCAB principle is in

force. This possibility provides an alternative to the LAB principle, assumed earlier, which supposes that like-atom bonds are unfavorable. In the absence of MCA's, a LCAB is a LAB and vice versa, but different coordination defects will be expected depending on which principle is in force.

The LAB principle can be given a physical motivation, based on the fact that the partial ionic character of a heteropolar bond makes it stronger than the average of the two different homopolar bonds. The difference  $\Delta E$  in bond energies  $D$  can be estimated from the As-Se electronegativity difference  $\Delta x$  (Ref. 21):

$$\begin{aligned}\Delta E &= D(\text{As-Se}) - [D(\text{As-As}) + D(\text{Se-Se})] / 2 \\ &= (\Delta x)^2 = 0.2 \text{ eV}\end{aligned}\tag{1}$$

Thus the fraction of like-atom bonds frozen in at the glass transition temperature is expected to be on the order of  $\exp(-\Delta E/kT_g) \approx 5 \times 10^{-3}$ . Put another way, this argues that a defect with  $n$  LAB's is at least 100 times as likely as a similar defect with  $n+1$  LAB's. However, the foregoing assumes that the energy cost of a LAB is independent of its local environment. It probably is not; for example, one could argue that a  $C_3$  atom will be more electropositive than a normal  $C_2$ , and may therefore allow chalcogen neighbors at no extra cost in energy. Moreover, an argument can be made in favor of the LCAB principle if one assumes that the continuous random network consists

of a disordered arrangement of layers with the internal structure of those in the crystal. In that case, defects which preserve the layer structure will be favored. The LCAB-free defects (basically As-Se substitutions) are of this structure-preserving type.

The same argument applies much more certainly to structural defects in crystalline  $\text{As}_2\text{Se}_3$ . Alternative defects (e.g., vacancies, or interstitials covalently bonded to neighboring layers) introduce a minimum of two MCA's, while a simple substitution has only one MCA and is entirely strain free. In fact, the assumption that the LCAB principle dominates the crystal while the LAB principle dominates the glass may provide a natural explanation for the higher photoluminescence quantum efficiency and absence of fatigue in crystalline  $\text{As}_2\text{Se}_3$  (Ref. 6) if the dominant nonradiative center in the glass contains a LCAB. For definiteness, we will proceed on the assumption that the LAB principle dominates in the glass, but will pay special attention to the substitution and interchange defects as we go along.

Table II shows an alternative classification scheme which is also instructive. Defects have been categorized according to two properties: (i) The "natural charge state" (the charge on the defect when all the valence like states are filled with electrons and all the conduction like states are empty; that is, when no carriers are trapped), and (ii) the "chalcogen excess." The latter is defined as the number of Se atoms, over

Table II

Classification of defects according to chalcogen excess and natural charge state.

		NATURAL CHARGE STATE			
		+2	+1	0	-1
CHALCOGEN EXCESS	3/2				$C_1'$
	1			$C_2', C_1'C_3'$	
	1/2		$P_4, C_3'$	←————→	$C_1$
	0	$P_2'C_3''$	←————→	$C_1P_3C_3$	
	-1/2		$C_3, P_2$		
	-1	$P_1, P_2C_3$	←————→	$P_3'$	
	-3/2		$P_2'$		

and above the usual  $\text{As}_2\text{Se}_3$  ratio, introduced by the defect. For example, if an As-Se bond in the bulk is broken and an extra Se atom inserted, a  $\text{C}_2'$  (Se LAB) defect results; thus the chalcogen excess of this defect is  $Q_c(\text{C}_2') = +1$ . Similarly, a Se atom removed from the bulk creates two  $\text{P}_2$  defects, so that  $Q_c(\text{P}_2) = -1/2$ . The chalcogen excess of any defect can be obtained in this way. The glass can thus become Se rich (As rich) by incorporating defects with  $Q_c > 0$  ( $Q_c < 0$ ).

This stoichiometry dependence is not the only interesting feature of the chalcogen excess. The defect model has been able to explain Fermi-level pinning in the absence of paramagnetism by the interconversion of defects from one natural charge state to another.<sup>12,14</sup> For example, the reaction  $(\text{C}_1)^- \leftrightarrow (\text{C}_3')^+$  is accomplished by breaking (reforming) a bond. However, the reaction  $\text{C}_1^- \leftrightarrow \text{C}_3^+$  is impossible because breaking a bond adjoining the  $\text{C}_3$  results in a  $\text{P}_2$ , not a  $\text{C}_1$ . We can systematize these considerations by noting that  $Q_c(\text{C}_1) = +1/2$  while  $Q_c(\text{C}_3) = -1/2$ , so that the reaction  $\text{C}_1 \leftrightarrow \text{C}_3$  can only occur if a Se atom is physically removed from the vicinity of the defect. Thus, such a reaction is forbidden, no matter what local relaxations take place, because atoms are not free to migrate at temperatures below the glass transition  $T_g$ . The charge state of a defect may easily be changed by trapping or emitting carriers, but the chalcogen excess of a defect is permanently fixed. Therefore, Fermi-level pinning reactions must be accomplished by horizontal transitions in Table II.



This "selection rule" greatly restricts the number of allowed defect interconversion processes in heteropolar glasses.<sup>51</sup>

It may come as a surprise that the  $P_2$  and  $P_1$  defects are assigned positive natural charge states. This is a consequence of the fact that a nonbonding As p orbital is more correctly identified as a conduction-band state than a valence-band state, as will be discussed in Sec. III(c). If one then considers transitions between the relatively favorable defects in the upper left-hand region of Fig. 13, one finds very few simple candidates for the pinning mechanism. These are shown by arrows in Table II. The most prominent candidates are almost certainly  $C_1^- \leftrightarrow (C_3')^+$  and perhaps  $C_1^- \leftrightarrow P_4^+$ . The others involve bond switching, not simply bond breaking and forming; they require a compensating population of some other negatively charged defects to maintain charge neutrality, and they involve the relatively unfavorable  $P_2$  and  $P_1$  defects [see Sec. III(c)].

Above  $T_g$ , atoms become free to migrate, and an equilibrium is set up among the populations of defects of differing chalcogen excess. The establishment of this equilibrium, which determines the defect densities frozen in at  $T_g$ , will be discussed in Sec. III(d). We turn now to a discussion of the electronic states associated with the defects of interest.

### (c) Electronic States

As a starting point, we consider crystalline  $\text{As}_2\text{Se}_3$ . Because the short-range order, or local chemical configuration, is expected to be the most important factor in determining the density of states, we expect the perfect glass to have bands and gaps similar to those in the crystal.  $\text{As}_2\text{Se}_3$  forms into a layered orpiment structure with 20 atoms per unit cell. Figure 14 shows the theoretical density of states for this structure, as calculated by Bullett<sup>50</sup> using chemical pseudopotentials. It is in good agreement with x-ray photoelectron spectroscopy (XPS) measurements.<sup>52</sup> The Se non-bonding states near -1 eV overlap the  $\sigma$ -bonding-like states near -4 eV to form the principle valence band, and the  $\sigma^*$ -like conduction band is centered near 2 eV. There are also two lower-lying s bands.

In order to make a simplified model, we consider a first-neighbor tight-binding Hamiltonian with just three p orbitals per site. Moreover, each layer is distorted slightly to make all bond angles precisely  $90^\circ$ , so that the structure of Fig. 15(a) results. Since all bonds are now parallel to the x, y, or z axis, the systems of  $p_x$ ,  $p_y$ , and  $p_z$  orbitals decouple into three noninteracting subsystems which may be solved independently. The interaction diagram for any one such subsystem, in terms of  $\sigma$  and  $\pi$  bonds, is shown in Fig. 15(b). Next, the bulk perfect glass will be represented by a Bethe-lattice version of this interaction diagram, which is shown in Fig. 16(a).

Figure 14

Theoretical density of states for crystalline  $\text{As}_2\text{Se}_3$  as given by Ref. 50. The zero of energy is the valence band maximum.

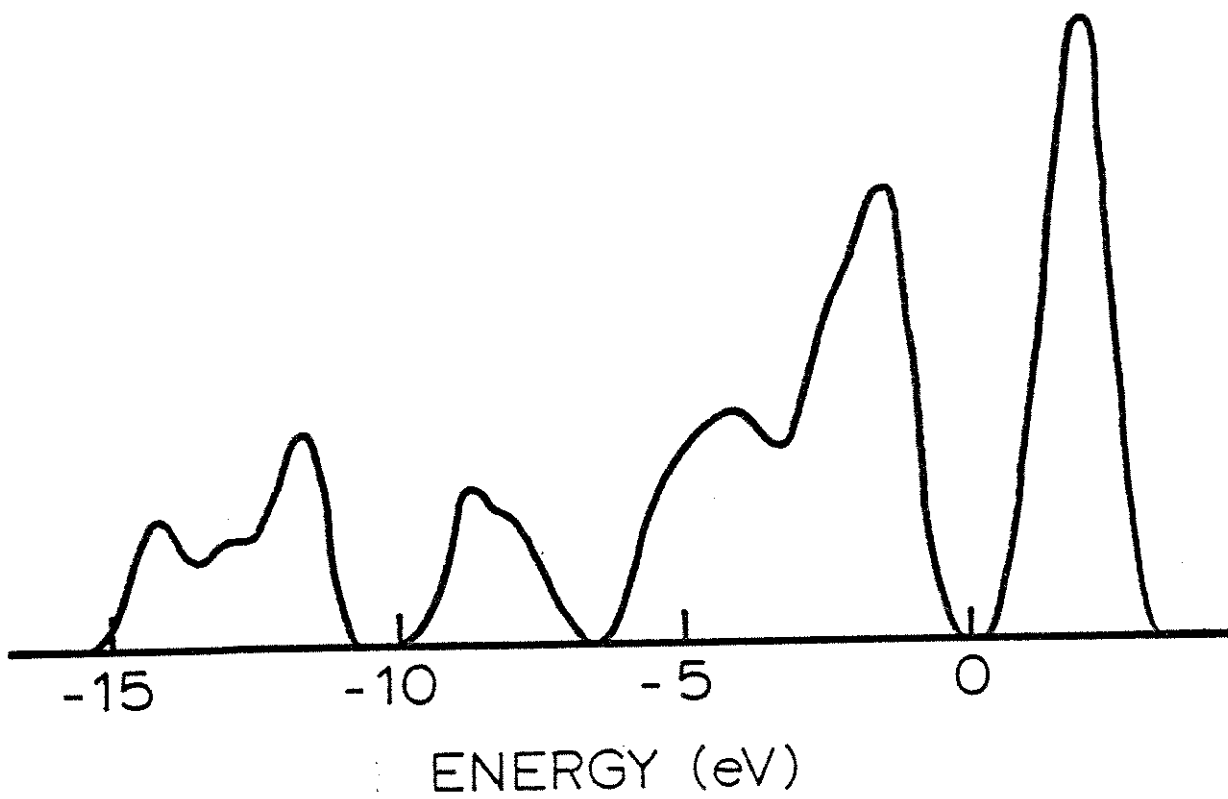
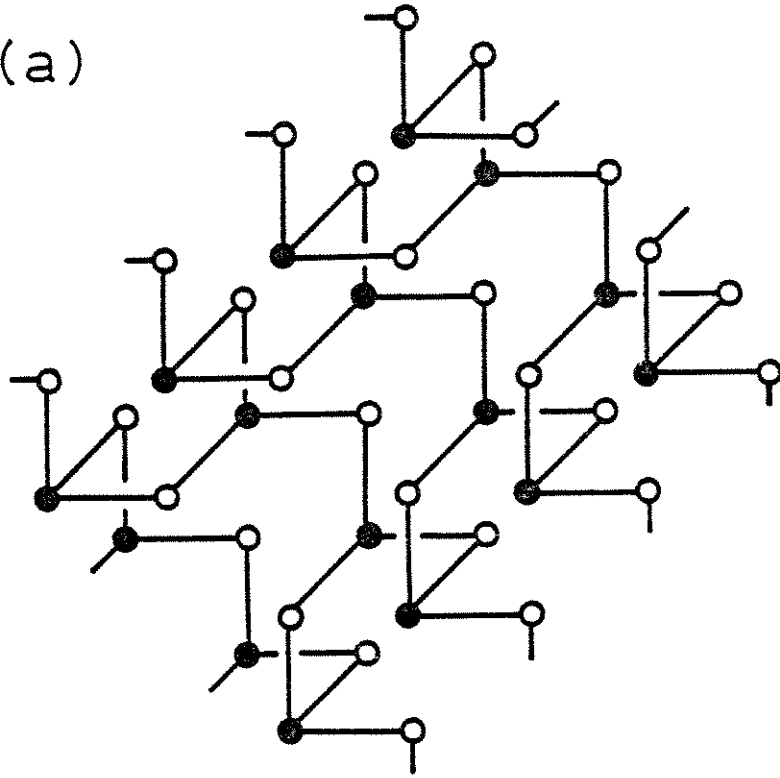


Figure 15

(a) Right-angle version of  $\text{As}_2\text{Se}_3$  layer. Solid circles are As, open circles Se. (b) Corresponding interaction diagram for  $p_x$  orbitals; double lines represent  $V_\sigma$ , single lines  $V_\pi$ .

(a)



(b)

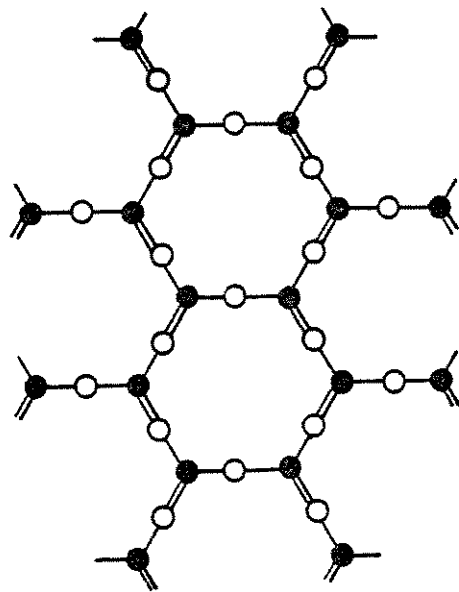
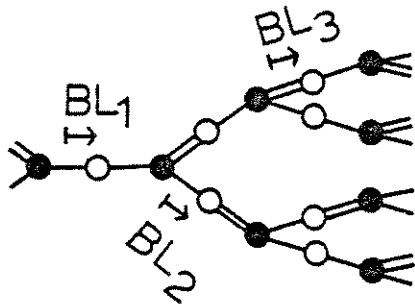


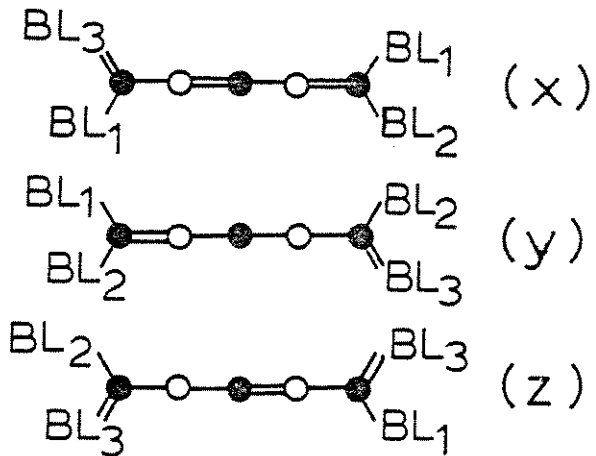
Figure 16

(a) Interaction diagram for Bethe-lattice structure. Solid circles  $A_s$ , open circles  $S_e$ ; double lines  $V_\sigma$ , single lines  $V_\pi$ . The semi-infinite trees  $BL_1$ ,  $BL_2$ , and  $BL_3$  are defined by the arrows. (b) Interaction diagram for the  $p_x$ ,  $p_y$ , and  $p_z$  orbitals of the  $P_2$  cluster-Bethe-lattice structure. Notations at  $A_s$  bonds indicate which semi-infinite tree is to be attached. (c) Same for  $C_2'$ . (d) Same for  $C_3$ .

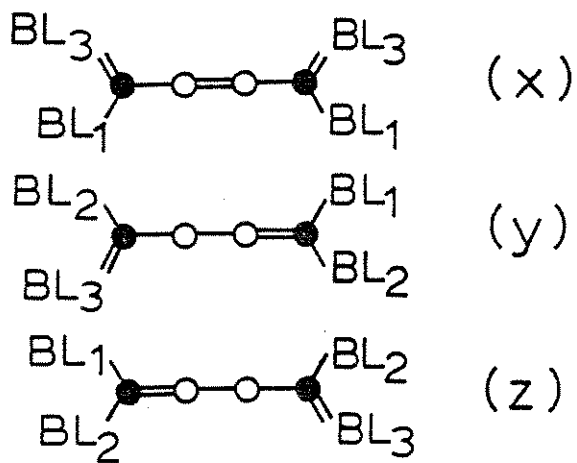
(a) B.L.



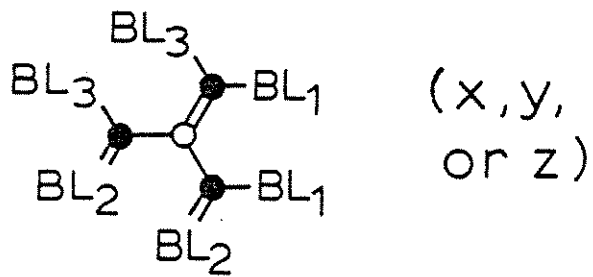
(b)  $P_2$



(c)  $C'_2$



(d)  $C_3$





Finally, we will use the cluster-Bethe-lattice method (CBLM)<sup>53</sup> to model defects. For example, a  $P_2$  defect results from breaking a bond in Fig. 15(a); the corresponding CBLM interaction diagrams are shown in Fig. 16(b). Similarly, the interaction diagram for over-coordinated defects on like-atom bonds can be modeled as shown in Figs. 16(c) and 16(d).

The resulting model is an elementary example of the CBLM, and is easily solved using Green's-function techniques. Before proceeding, however, it is necessary to comment upon the many simplifying assumptions which went into this model. We have effectively omitted interlayer interactions by choosing a nearest-neighbor Hamiltonian; we have chosen  $90^\circ$  bond angles and omitted s and d orbitals; and we have obscured the ring topology by introducing the Bethe lattice. We have been motivated in these choices by the fact that the interlayer interactions are weak;<sup>54</sup> that the s-p hybridization does not dominate the bonding (average As and Se bond angles are  $100^\circ$  and  $94^\circ$ , respectively, in the crystal)<sup>55</sup>; and that the rings in  $As_2Se_3$  are large and are presumably randomized in the glass anyway. Moreover, we are guided by the philosophy that we are looking for trends when defects are introduced into the glass; it is expected that perturbations due to s states or interlayer interactions will act equally, to a first approximation, upon the defect and bulk states.

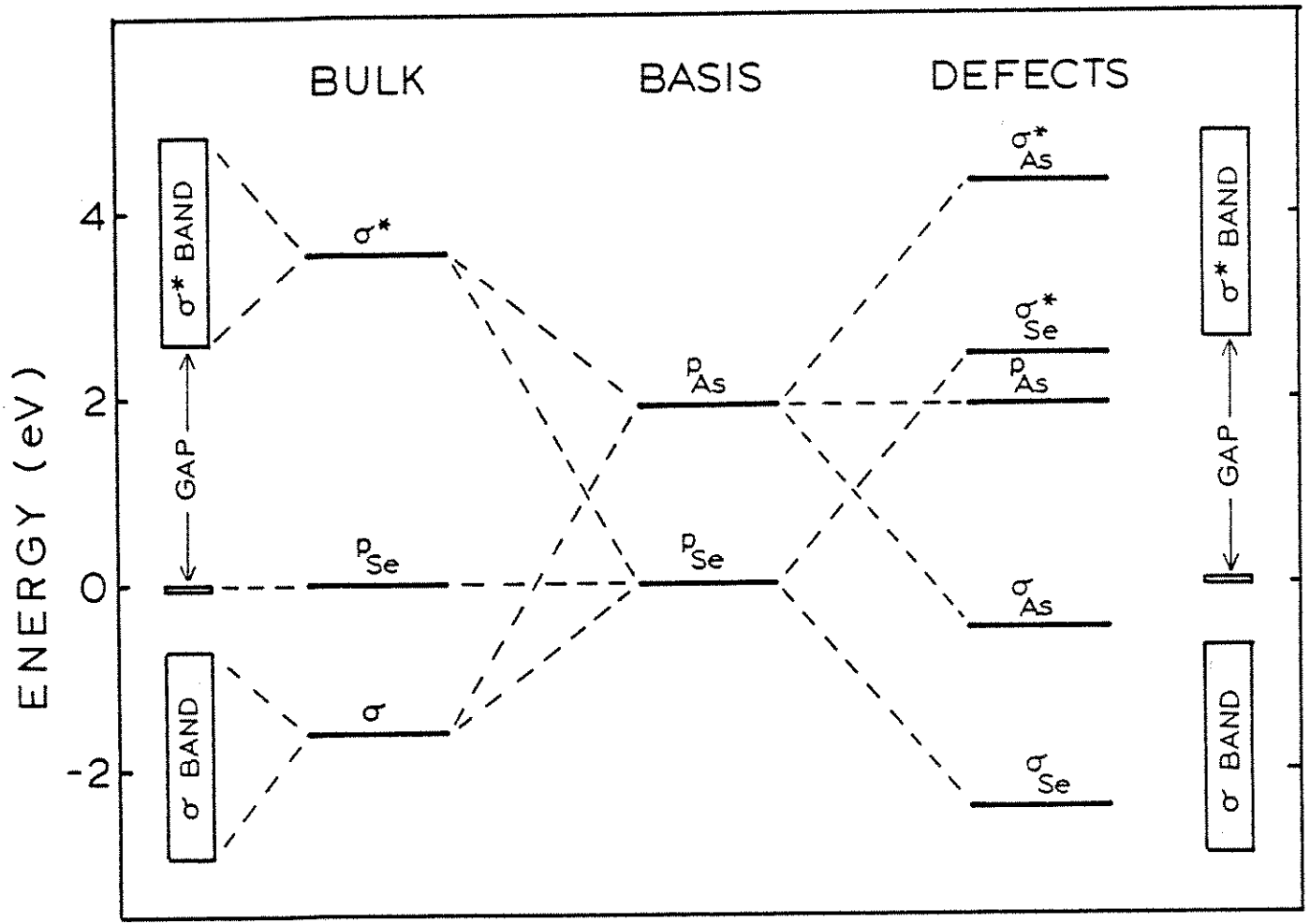
A more serious problem is the fact that the calculations are not self-consistent. For this reason, the location of gap

states in this model cannot be taken at face value. We will rely here upon the lessons of our earlier study on glassy Se in Chapter II, where we carried out much more thorough calculations. In addition to a simple tight-binding model for Se, exactly analogous to the one just described for  $\text{As}_2\text{Se}_3$ , we carried out more sophisticated tight-binding calculations on isolated defects and realistic self-consistent pseudopotential calculations on superlattice structures containing defects. It was found, you will recall, that self-consistency requires certain tight-binding parameters to be altered in the vicinity of defects. We will frequently argue by analogy to Se in order to predict, in general terms, the effects which would be introduced into our very simple model by self-consistency or other complications.

Let us turn now to a detailed consideration of our elementary model. It contains two kinds of basis orbitals: Se p orbitals, to which we assign energy level  $E(p_{\text{Se}}) = 0$  by convention, and As p levels at  $E(p_{\text{As}}) = \Delta$ . Since As is more electropositive than Se,  $\Delta$  is positive, and we choose  $\Delta = 1.89$  eV following Bullett.<sup>25</sup> These basis levels are shown in the center of Fig. 17. Consider now the bulk Bethe-lattice of Fig. 16(a); focusing on the dominant  $\sigma$  bonding, we note that it contains three kinds of bond orbitals. These are the heteropolar  $\sigma$  and  $\sigma^*$  bond-orbitals and the nonbonding Se p orbital, shown at left in Fig. 17. Taking  $V_\sigma = 2.42$  eV (Ref. 56) we find,

Figure 17

Energy levels relevant to a-As<sub>2</sub>Se<sub>3</sub>, as determined from simple model of text.  $\sigma$  interactions between basis p orbitals (center) give rise to bulk bond orbitals (left) and defect bond orbitals (right).  $\pi$  interactions between the former in turn give rise to Bethe-lattice bulk bands (extreme left, repeated extreme right for reference).



$$E(\sigma^*) = (\Delta/2) + [(\Delta/2)^2 + V_\sigma^2]^{1/2} = 3.54 \text{ eV}, \quad (2a)$$

$$E(p_{Se}) = 0, \quad (2b)$$

$$E(\sigma) = (\Delta/2) - [(\Delta/2)^2 + V_\sigma^2]^{1/2} = -1.65 \text{ eV}. \quad (2c)$$

When  $\pi$  interactions are included [we let  $V_\pi = 1.05$  eV (Ref. 56)] and the Bethe-lattice model is solved using standard Green's-function techniques,<sup>53</sup> the  $\sigma$  and  $\sigma^*$  levels broaden into the bands shown at extreme left of Fig. 17. The lone-pair band occurs at  $E = 0$  eV but remains unbroadened, a peculiar artifact of the bonding geometry of the model. It therefore avoids overlapping with the  $\sigma$ -bonding band, leaving a secondary gap where the real glass has none. The  $\sigma$  and lone-pair bands are valence bands, while the  $\sigma^*$  band forms the conduction band. In our model, the fundamental gap occurs between 0 eV and

$$\begin{aligned} E(\sigma^* \text{ band min.}) &= (\Delta/2) + [(\Delta/2)^2 + (V_\sigma + V_\pi)^2]^{1/2} \\ &= 2.61 \text{ eV} \end{aligned} \quad (2d)$$

Also shown at the right of Fig. 17 are the other five simple bond orbitals which can occur in glassy  $As_2Se_3$ , but only in the presence of defects. Note that several of these bond orbitals (the nonbonding  $p_{As}$  orbital and the homopolar  $\sigma^*_{Se}$  and  $\sigma_{As}$  orbitals) have energy levels within, or close to, the fundamental gap:

$$E(\sigma^*_{Se}) = V_{\sigma} = 2.42 \text{ eV}, \quad (2e)$$

$$E(p_{As}) = \Delta = 1.89 \text{ eV}, \quad (2f)$$

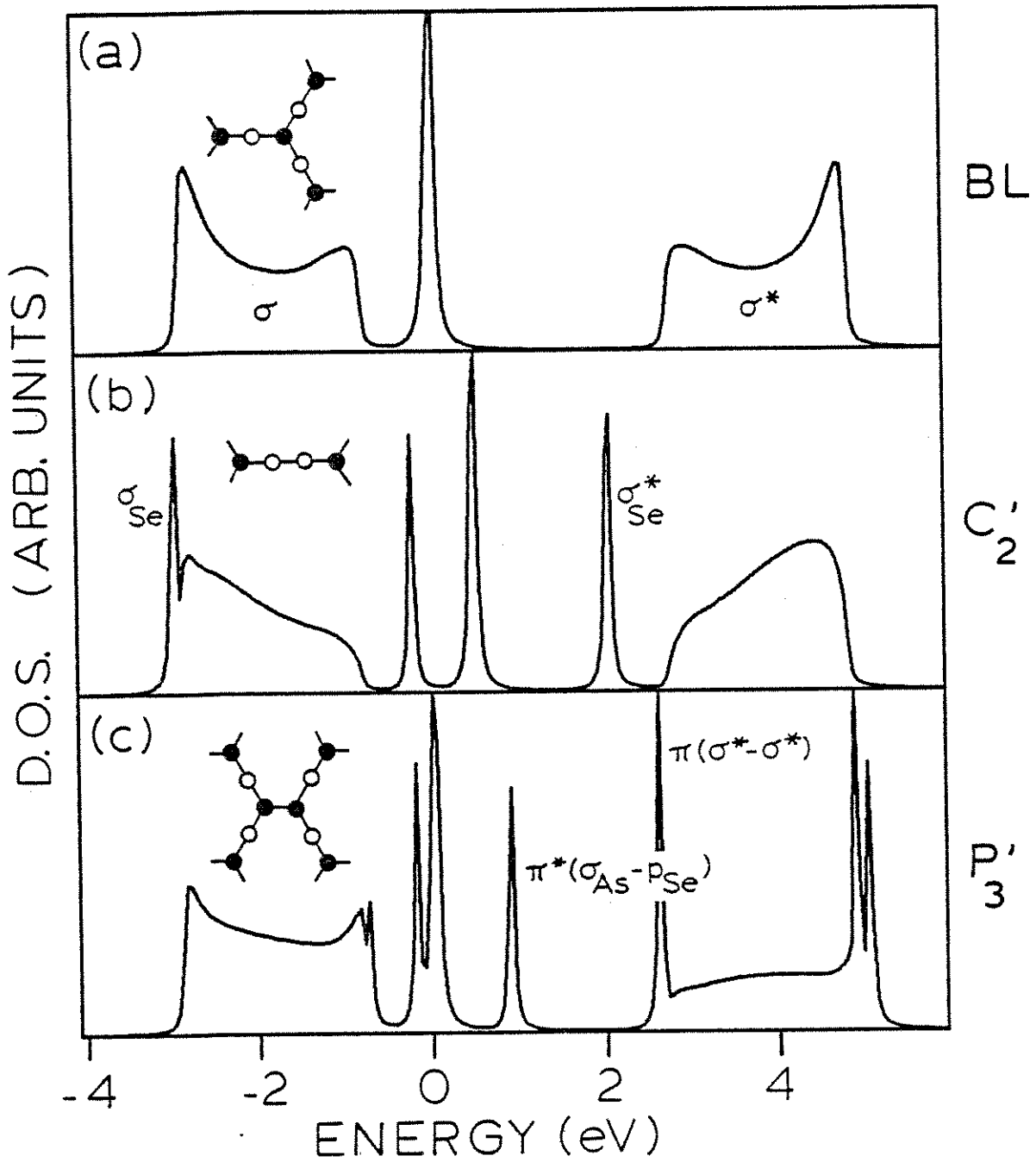
$$E(\sigma_{As}) = \Delta - V_{\sigma} = -0.53 \text{ eV}. \quad (2g)$$

As we shall see, it is these orbitals which give rise to the gap states of interest in  $As_2Se_3$ . It is the electronegativity difference which is responsible for raising the  $p_{As}$  level out of the lone-pair band, and shifting the  $\sigma_{As}$  and  $\sigma^*_{Se}$  levels out of the  $\sigma$  and  $\sigma^*$  bands.

Let us begin our survey of defects by considering the simple like-atom bonds. This is a class of defects which has no counterpart in glassy Se. In Fig. 18(a) and 18(b) we present the density of states (as obtained from the Green's function) for the bulk Bethe lattice and for the  $C_2'$  defect (Se LAB), respectively. The locations of the  $\sigma$ ,  $p_{Se}$ , and  $\sigma^*$  bulk bands in Fig. 18(a) correspond exactly with those shown at the extreme left of Fig. 17. The new features which at the  $C_2'$  defect can be understood by referring to the corresponding interaction diagrams in Fig. 16(c). Note that the system of x orbitals contains a new  $\sigma$ -bonded pair of  $p_{Se}$  orbitals. The resulting  $\sigma^*_{Se}$  level falls within the gap at 2.42 eV (see Fig. 17); when embedded in the bulk, it is shifted deeper into the gap by the  $\pi$  interactions to form the state at 2.02 eV, below the conduction-band edge in Fig. 18(b). (The defect is neutral

Figure 18

(a) Density of states for Bethe-lattice structure. (b) Local density of states, averaged over sites near the defect, for  $C_2'$  like-atom bond. (c) Same for  $P_3'$  like-atom bond. Certain features discussed in the text are labeled according to their dominant character. Schematic diagrams show the structure (solid circles As, open Se; external As bonds are understood to have Bethe-lattices attached). The fundamental gap extends from 0 to 2.61 eV. Note that a Lorentzian broadening of half-width 0.05 has been introduced by the Green's function calculation.





when this state is unoccupied.) Similarly, there is a  $\sigma_{\text{Se}}$  state appearing just at the bottom of the bonding band. (There is also some splitting of states out of the lone-pair band due to slight changes in the interactions near the defect, but this may be an artifact of the lack of lone-pair bandwidth.) We expect the situation to be similar for the states of the  $\text{C}_2''$  defect, which appears in the raft model of Phillips.<sup>19</sup>

The situation for the  $\text{P}_3'$  defect of Fig. 18(c) is similar, but now a pair of As p orbitals give rise to  $\sigma_{\text{As}}$  and  $\sigma_{\text{As}}^*$  bond orbitals. Notice from Fig. 17 that the  $\sigma_{\text{As}}$  level lies quite close in energy to the Se nonbonding orbitals. In fact the  $\sigma_{\text{As}}$  bond orbital interacts via a direct  $\pi$  interaction with two neighboring  $\text{p}_{\text{Se}}$  orbitals, giving rise to  $\pi$  and  $\pi^*$  complexes. The  $\pi^*$  complex would occur at

$$\begin{aligned} E_{\pi^*}(\sigma_{\text{As}}-\text{p}_{\text{Se}}) &= E(\sigma_{\text{As}})/2 + \{[E(\sigma_{\text{As}})/2]^2 + V_{\pi}^2\}^{1/2} \\ &= 0.82 \text{ eV} \end{aligned} \quad (2h)$$

if isolated. This is the origin of the gap state at 0.87 eV in Fig. 18(c). (The defect is neutral when this gap state is fully occupied.) Presumably this state will not be quite so far above the valence-band edge when a lone-pair band of nonzero width is considered, and it may be inhomogenously broadened by variations in the local environments of the defects. Finally, notice that a state appears to emerge just at or below the conduction-band edge as well. This occurs because the  $\sigma^*$

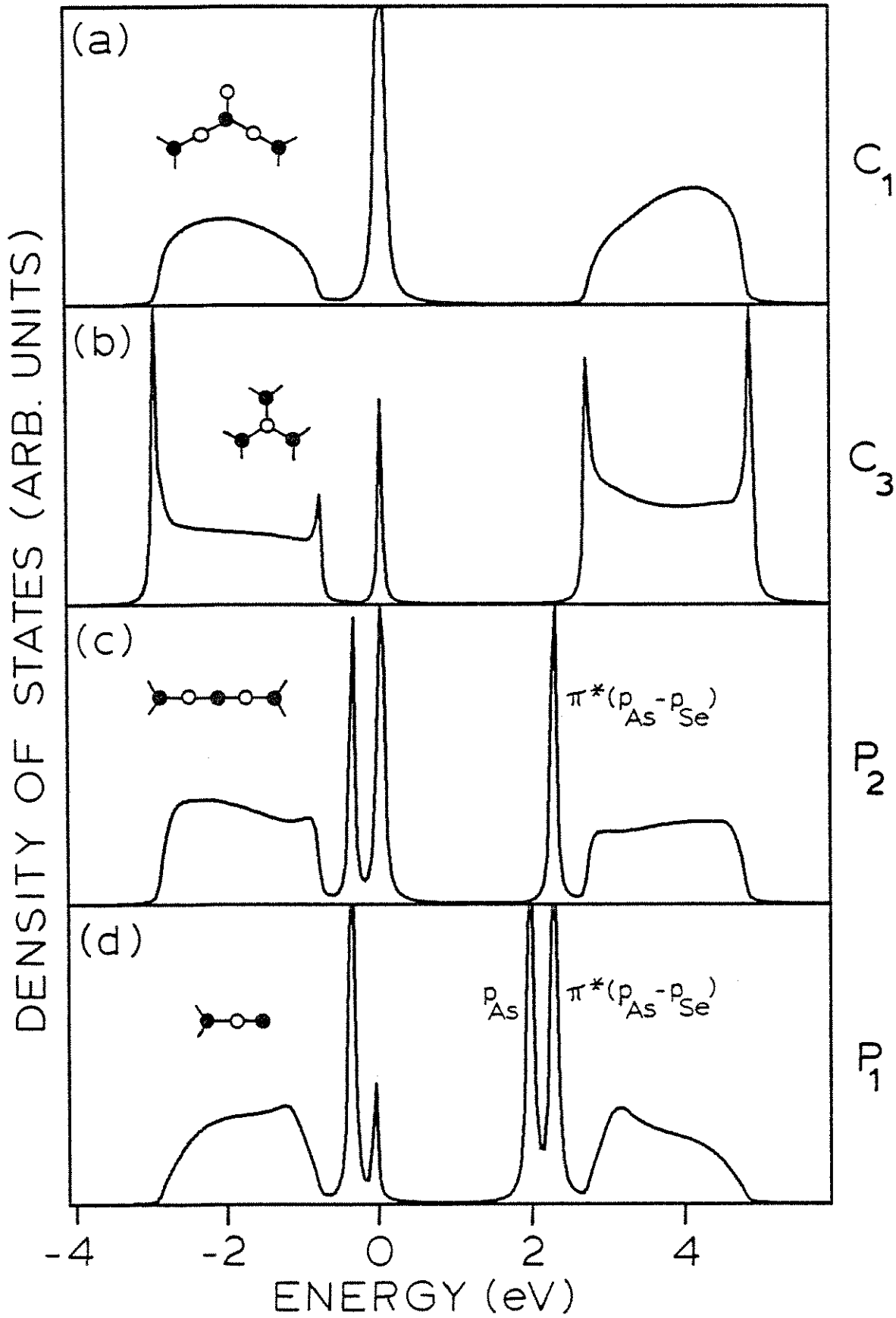
bandwidth results from  $\pi$  interactions which couple the Se side of one  $\sigma^*$  bond orbital to the As side of the next, while the  $\pi$  interaction along the As-As bond couples to the As side of both  $\sigma^*$  orbitals. The  $\sigma^*$  orbitals have more As than Se character (a consequence of the electronegativity difference), so the effective interaction at the As-As bond is atypically large, resulting in  $\pi(\sigma^*-\sigma^*)$  and  $\pi^*(\sigma^*-\sigma^*)$  levels just below and above the conduction-band edges.

Next, we turn to the simple coordination defects. Figures 19(a) and 19(b) show the density of states in the vicinity of an under- or overcoordinated Se atom, respectively. One finds no deep gap states, although the additional interactions among  $\sigma$  and  $\sigma^*$  orbitals in the neighborhood of the  $C_3$  defect give rise to strong resonances near the corresponding band edges. The lack of deep gap states is a consequence of the fact that no novel bond orbital or interaction has been introduced. This is unlike the case of  $C_1$  and  $C_3$  defects in amorphous Se, where a direct  $\pi$  interaction between nonbinding or  $\sigma^*$  orbitals gives rise to deep gap states. In  $As_2Se_3$ , direct  $\pi$  interactions between  $\sigma^*$  orbitals already occur in the bulk, and no direct  $\pi$  interactions between Se nonbonding orbitals (NBO's) are introduced by the  $C_1$  defect, because the neighboring site contains an As atom.

At this point we must point out that the single MCA defects are not neutral in their natural charge states, unlike the LAB's. The  $C_1$  and  $C_3$ , for example, have charges -1 and +1

Figure 19

Local density of states for simple malcoordinated atom defects. (a)  $C_1$ , (b)  $C_3$ , (c)  $P_2$ , and (d)  $P_1$ . Details of Fig. 18 apply.



respectively, if the valence bands are fully occupied and the conduction band empty. This means that while these defects will not support deep gap states, they will support hydrogenic acceptor and donor levels. Consider, for example, the neutral  $C_3$  defect, which has one electron in the conduction band. If that electron were localized to the defect site, each atom would be approximately neutral, and self-consistency would not be important. Instead, that electron tries to go into a conduction-band state, leaving behind a positive charge which in fact traps the electron in a hydrogenic orbit. The same self-consistency argument applies to a hole on the neutral  $C_1$  defect, which gives rise to a hydrogenic acceptor level.

Unlike the  $C_1$  and  $C_3$ , the undercoordinated  $P_2$  and  $P_1$  defects do introduce a new bond orbital, namely, an As nonbonding orbital. When isolated, this orbital lies at  $E = \Delta$ , below the conduction-band edge. In the  $P_2$  defect it interacts with a Se NBO via a direct  $\pi$  interaction, and moves further upward  $\sim 0.3$  eV, to form the state at 2.27 eV in Fig. 19(c). In the  $P_1$  there is a second NBO which remains near its unperturbed energy on the same As site, giving rise to a second defect state in Fig. 19(d).

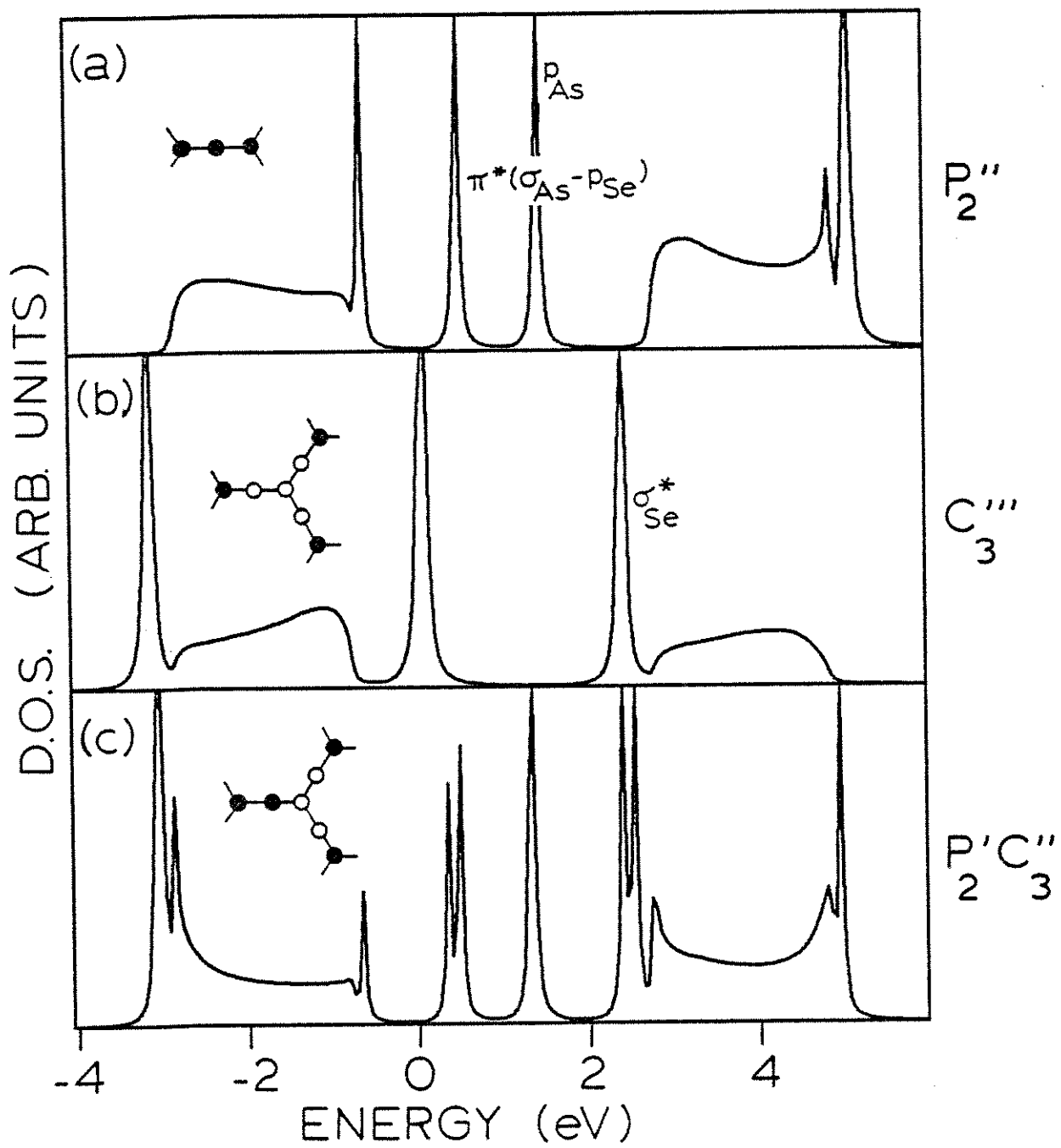
For a neutral  $P_2$  the gap level is half-occupied, and for a  $P_1$  the lower gap level is fully occupied while the higher one is empty. Self-consistency would not be important at these defects, because the gap states are well localized and the sites comprising the defect are therefore individually

neutral. Note that if the Fermi Level falls near midgap (as it is known to do experimentally), the  $P_2$  and  $P_1$  will carry charge +1 and +2, respectively. The usual picture of coordination defects in chalcogenides considers the undercoordinated defects to be naturally negatively charged, but the conventional wisdom fails here because the As NBO level falls closer to the conduction band than to the valence band. This unusual state of affairs is due to the electronegativity difference, which shifts the  $p_{As}$  level well above the  $p_{Se}$  valence levels (see Fig. 17). Physically, a  $P_2^-$  is unfavorable because the extra electron would be highly localized to an electropositive As site. Of course, in a lone-pair semiconductor, the unoccupied As p orbitals on the  $P_2^+$  will be unstable towards dative bonding with a neighboring  $p_{Se}$  (or even  $s_{As}$ ) lone pair.<sup>16</sup> Thus it appears that undercoordinated pnictide defects are unlikely to occur at all [see also Sec. III(d)]. However, if local bonding constraints occasionally forbid such dative bonding, we expect the resulting defects to be positively charged.

The last class of defects to be considered explicitly will be the substitutions and interchanges. As pointed out in Sec. III(b), these are likely to occur as defects in crystalline  $As_2Se_3$ , or perhaps as a consequence of the LCAB principle in the glass. If an As atom is substituted for a Se, the  $P_2''$  defect of Fig. 20(a) results. The LAB's give rise to the two-fold degenerate  $\sigma_{As}$  level near 0.45 eV, and the As NBO on the  $P_2$  site gives rise to the level near midgap. (It lies lower in

Figure 20

Local density of states for substitution and exchange defects. (a)  $P_2''$ , (b)  $C_3'''$ , and (c)  $P_2'C_3''$ . Details of Fig. 18 apply.





energy than in the  $P_2$  because of an interaction with  $\sigma^*$  orbitals.) The defect is neutral when the As NBO is half-occupied; this state is highly localized and self-consistency should not be important. The resulting picture is quite consistent with the work of Bullett,<sup>50</sup> who has carried out chemical pseudopotential calculations on the substitutional defects. (In the latter calculation, the introduction of the correct bond angles breaks the twofold degeneracy, giving rise to a pair of  $\sigma_{As}$ -like states above the valence-band edge.)

Figure 20(b) shows the density of states near the  $C_3'''$  defect, in which a Se atom has replaced an As. The gap state below the conduction-band edge is threefold degenerate and has most of its character on  $\sigma_{Se}^*$  orbitals at the defect. Just as for the  $C_3$  defect in pure Se, the threefold degeneracy will be broken when the bond angles are allowed to differ from  $90^\circ$ , and a nondegenerate state will fall somewhat lower in the gap. Self-consistency arguments suggest that this state will be lowered further still by the negative self-energy shift at the defect site which is necessary to restore charge neutrality. Once again we have good agreement with the previous results of Bullett,<sup>50</sup> who finds a triplet of gap states centered  $\sim 0.4$  eV below the conduction-band edge.

Finally, the result of interchanging nearest-neighbor As and Se atoms is shown in Fig. 20(c). The density of states resembles a superposition of the  $C_3'''$  and  $P_2''$ , except that

the degenerate orbitals have been split by the lowered symmetry. The defect is neutral when the As NBO state near mid-gap contains two electrons. It would be quite possible to go on considering more complicated defects in detail. As the last examples illustrate, however, once we have understood the nature of the gap states introduced by individual LAB's and MCA's, the essential features of the more complex composite defects can readily be predicted.

(d) Defect Creation Energies

We shall end this chapter with some brief speculations about defect creation energies, defined as the ground-state total energy of a given defect configuration minus that of the same set of atoms in the bulk. The electronic contribution to this difference is (see Appendix A):

$$\delta E_{el} = \sum_i \int_{-\infty}^{\epsilon_F} \epsilon [\eta_i(\epsilon) - \eta_i^0(\epsilon)] d\epsilon \quad (3)$$

Where  $\eta_i(\epsilon)$  and  $\eta_i^0(\epsilon)$  are the local density of states on atom  $i$  for the defect and bulk configurations, respectively. This is not the entire expression for the total energy, however, because the latter also includes the Coulomb repulsion between cores and other corrections. If these are modeled as a constant repulsive energy  $R$  per bond (see Appendix A), the resulting contribution is

$$\delta E_R = (R/2) \sum_i (c_i - c_i^0) \quad (4)$$

where  $c_i$  and  $c_i^0$  correspond to the coordination number of the  $i^{\text{th}}$  site in the defect and bulk, respectively. Finally, the defect creation energy is

$$E_{\text{tot}} = \delta E_{\text{el}} + \delta E_R + Q_e \mu_e - Q_c \mu_c \quad (5)$$

Here  $Q_e$  is the charge on the defect and  $Q_c$  is the chalcogen excess. The chemical potentials  $\mu_e$  (Fermi level) and  $\mu_c$  are introduced as a reminder that the difference in energies of two defects with different charge state, or different chalcogen excess, is not uniquely defined. However, the heat of reaction for any physically realizable defect interconversion process is independent of the chemical potentials, since numbers of electrons and atoms must be conserved.

Estimates of defect total energies are notoriously difficult; realistic calculations are available only for pure Se (Chapter II), and these demonstrate that simple models which are based solely upon discrete bond-orbital energies, and which neglect intercore repulsion, give decidedly inadequate results. Adler<sup>57</sup> has recently modified such a simple model to include a constant repulsive  $R$  per bond, in order to give zeroth-order defect energy estimates. However, such a model still does not include the effects of nonzero bandwidths or gap states and resonances; nor can it account for

electronegativity differences, charge transfers, or the partial ionic character of bonds in the heteropolar glasses.

As was pointed out in Sec. III(b), the latter effects are expected to give the simple LAB defects a creation energy on the order of 0.2 eV. The creation energy of MCA's, however, is undoubtedly larger ( $\geq 0.5$  eV for Se in Chapter II) and much harder to estimate. The electronic binding energy gained (lost) in forming (breaking) a bond will be at least partly compensated by the corresponding  $\delta E_p$ . It may be overcompensated, in which case the bond will prefer not to form; it is difficult to predict, a priori, whether this will be so. In the case of pure Se, we found that the bond prefers to be broken, i.e.,  $C_3^{\circ} \rightarrow C_1^{\circ}$  is exothermic. However, this is largely due to the anomalous  $\pi$  interaction which stabilizes the  $C_1$  defect; we have shown this does not occur at the  $C_1$  defect in  $As_2Se_3$ . In fact, none of the simple MCA's ( $C_1$ ,  $C_3$ ,  $P_2$ ,  $P_1$ ) shows analogous  $\pi$  bond stabilization, and it is likely that the corresponding creation energies are closer than in pure Se. The  $(C_1')^{\circ}$  defect will, on the other hand, exhibit this  $\pi$ -bond stabilization, and it may be lower in energy than the  $C_1^{\circ}$  despite the necessity of adding a LAB.

There is some reason to expect the chalcogen MCA's to be preferred over the pnictide ones. Adler<sup>57</sup> points out that the Coulomb U for creating the  $P_2^-$  (or  $P_2^+$ ) involves placing two electrons (or holes) on a highly localized orbital, and is therefore larger than for the  $C_1$  or  $C_3$ . Moreover, the neutral

$C_1$  or  $C_3$  has a hole or electron in a band edge state, rather than at the bond-orbital energy (band center); this lowers the energy with respect to the zeroth-order model. The corresponding stabilization does not take place for  $P_2$  or  $P_1$  defects, again because of the deep gap nature of As NBO levels.

It has been proposed<sup>16</sup> that the electronegativity difference will favor the  $P_4^+$  defect over the  $C_3^+$ , and the  $C_1^-$  over the  $P_2^-$ . However, this will only be the case if the extra electron or hole is constrained to reside on the defect site itself. As shown in Sec. III(c), the  $P_2^-$  is quite unfavorable for this reason, with the extra electron localized predominantly to a single As NBO. On the other hand, the argument fails for the  $C_3^+$ , in which the added hole is shared by  $\sigma^*$  orbitals which have more As than Se character. (For the same reason, the  $C_3^0$  cannot be ruled out as the source of the broad "As center" seen in optically induced ESR experiments.<sup>5</sup>) In the absence of compelling arguments to the contrary, we therefore follow Adler<sup>57</sup> in assuming that the large s-p promotion energy makes  $P_4$  defects relatively unfavorable, although this may be compensated in part by the increased strength of the more highly directed tetrahedral bonds.

Of course, defects must occur in pairs (or triplets, etc.) whose total charge and chalcogen excess are constrained to zero in pure  $As_2Se_3$ . If we assume that the defects are frozen in at the glass transition temperature  $T_g$ ,<sup>14</sup> then the density of each defect is  $n_0 \exp(-E_{tot}/kT_g)$ , with  $n_0$  being roughly the density

of atomic sites and  $E_{\text{tot}}$  given by Eq. (5). The chemical potentials  $\mu_e$  and  $\mu_c$  are then determined by the requirements of charge neutrality and the stoichiometric ratio of the glass. For example, the estimate of Eq. (1) suggests that the neutral  $C_2'$  and  $P_3'$  LAB's will be by far the most common defects. In order to have an  $\text{As}_2\text{Se}_3$  alloy we require the densities of the two defects to be equal:

$$n_o \exp[-E_{\text{tot}}(C_2')/kT_g] = n_o \exp[-E_{\text{tot}}(P_3')/kT_g] \quad (6)$$

Since  $\delta E_R = Q_e = 0$ ,  $Q_c = \pm 1$  for these defects, Eq. (5) gives

$$\mu_c = [\delta E_{e1}(C_2') - \delta E_{e1}(P_3')] / 2 \quad (7)$$

and

$$\begin{aligned} n(C_2') &= n(P_3') \\ &= n_o \exp(-[\delta E_{e1}(C_2') + \delta E_{e1}(P_3')] / 2kT_g) \end{aligned} \quad (8)$$

(Note that if the defects occur only in close pairs,<sup>21</sup> the Boltzmann prefactor would be the sum rather than the average of the  $C_2'$  and  $P_3'$  energies, and a considerably lower density of defects would occur, reflecting the loss of the entropy of mixing.<sup>16</sup> Therefore distant pairs will predominate unless the binding energy of a pair exceeds the average creation energy, which is unlikely for neutral defects.)

$\mu_c$  is plausibly on the order of a tenth of an eV or less, so that the last term of Eq. (5) is probably never very important for coordination defects. That is, the chalcogen excess of the coordination defects in  $As_2Se_3$  (which presumably pin  $\epsilon_F$ ) is not constrained to balance, since it can easily be compensated by LAB's elsewhere in the glass. For the same reason, the concentrations of the various coordination defects are not expected to be highly stoichiometry dependent.

The picture which emerges of the glass at  $T_g$ , then, is of a fairly high density (perhaps  $\sim 10^{20} \text{ cm}^{-3}$ ) of LAB defects. If the  $P_3'$  gives rise to a  $\sigma_{As}$ -like states above the valence-band edge as expected, the observed intrinsic hole mobility must be governed by extended state conduction or fast hopping among this manifold of states, as suggested by Halpern.<sup>21</sup> Scattered among the LAB's would be a smaller number ( $\sim 10^{18} \text{ cm}^{-3}$ ) of charged coordination defects. These would be dominated by the lowest-energy charge-compensated pair [possibly  $C_1^-$  and  $C_3^+$ , which would occur, for example, at a density of

$$n_o \exp\{-[\delta E_{e1}(C_1^-) + \delta E_{e1}(C_3^+)]/2kT_g\} \quad (9)$$

in analogy to Eq. (8)]. One or both members of this pair could then give rise to the neutral excitations responsible for the photoinduced phenomena, and to the negative-U pinning of the Fermi level via one of the horizontal reactions in Table II, e.g.,  $C_1^- \leftrightarrow (C_3^+)$ .

This picture is admittedly speculative, and the identification of those defects actually involved would require realistic total energy calculations or new experimental information. Moreover, the effects of interactions between charged defects may be important, and have not been discussed above. Nevertheless, given our present state of knowledge in this area, we believe this description to be a plausible starting point for developing a viable model for defects in the heteropolar chalcogenides.

#### (e) Summary and Conclusions

We have developed a scheme for classifying and labeling defects in terms of the constituent like-atom bonds (LAB's) and malcoordinated atoms (MCA's). The concept of like-coordinated-atom bonds (LCAB's) is also discussed, and the fact that LCAB's are suppressed for defects in crystalline  $\text{As}_2\text{Se}_3$  is offered as a possible explanation for experimental indications that certain nonradiative recombination centers occur only in the glass. We have shown that each defect is characterized by a chalcogen excess which characterizes its contribution to the deviation from an exact 2:3 stoichiometry, and we have pointed out that conservation of chalcogen excess imposes a constraint upon the possible defect interconversion processes which can pin the Fermi level below  $T_g$ .

Next, an elementary Bethe-lattice model is presented and solved for a selection of simple defects involving LAB's and



MCA's. While the results of such a simple model cannot be taken at face value, we know how they are modified in more realistic (e.g., self-consistent) calculations for Se, and we can use this knowledge to predict the gross features of a realistic calculation in  $\text{As}_2\text{Se}_3$ . Our primary result is that defects in  $\text{As}_2\text{Se}_3$  are quite different from those in Se, and arise for entirely different reasons. While in Se, gap states arise at onefold and threefold sites because of unique  $\pi$  bonding between bond orbitals on the defect and its neighbors, in  $\text{As}_2\text{Se}_3$  they are associated with unique bond orbitals (like-atom bond  $\sigma$  and  $\sigma^*$  orbitals and As NBO's). Thus, while  $C_1$  and  $C_3$  defects give rise to deep gap states in Se, they admit only hydrogenic levels in  $\text{As}_2\text{Se}_3$ . Deep states do exist, however, at LAB's and at undercoordinated pnictide sites. We believe that the latter defects ( $P_2$  and  $P_1$ ) will not be negatively charged, as previously thought, but rather positively charged due to the position of the As nonbonding p orbitals near the conduction-band edge in  $\text{As}_2\text{Se}_3$ .

Finally, we discuss the structural equilibrium of the glass at  $T_g$ , and introduce a new chemical potential  $\mu_c$  which reflects the relative abundance of chalcogen atoms with respect to pnictides. We argue that the creation energies of the simple LAB's will be low (and will fix  $\mu_c$ ), so that a large number of LAB's (perhaps  $\sim 10^{20} \text{ cm}^{-3}$ ) will permeate the glass. We then suggest that a smaller density (perhaps  $\sim 10^{18} \text{ cm}^{-3}$ ) of coordination defects, predominantly on chalcogen sites, is

responsible for the negative-U properties and the pinning of the Fermi level.

## CHAPTER IV POLYACETYLENE

### (a) Introduction

The spectacular increase in the conductivity of polyacetylene upon doping<sup>58</sup> has drawn much attention to this material.<sup>59</sup> Partly as a consequence, it is being realized that undoped  $(CH)_x$  is a fascinating model system in its own right. It is perhaps the simplest system having a strong Peierls distortion, and has been proposed to support domain wall (soliton) excitations between regions of opposite bond alternation.<sup>60,61</sup> Furthermore, it has been suggested that charged solitons may exist in the ground state of lightly doped  $(CH)_x$ .

Many questions about the undoped material remain controversial or unanswered. Optical absorption and photoconductivity measurements consistently fail to show a sharp edge at the optical absorption threshold characteristic of 1D (one-dimensional) density of states.<sup>62,63</sup> There is difficulty obtaining a fit between theoretical calculations and the details of the experimental photoemission spectra.<sup>64,65</sup> The explanation of these discrepancies may depend upon gaining an understanding of the structure; the x-ray data in fact suggest a distribution of interchain distances,<sup>66</sup> but little else is known. What is the microscopic fibril structure? What is the typical size of a polymeric unit in the  $(CH)_x$  film? Is

there significant bending, twisting, splaying, and crosslinking of the polymer chains?

While we cannot unambiguously answer such questions, we have investigated theoretically the consequences of various kinds of structural disorder in  $(\text{CH})_x$ , and find that considerable disorder is consistent with the experimental work done to date. In Sec. IV(b) we discuss topological disorder, including chain ends and crosslinks, and discuss solitons and the question of soliton binding to these sites. In Sec. IV(c) we restrict ourselves to infinite chains whose bond alternation remains intact, but consider various kinds of structural disorder which can nevertheless exist. These include regions of cis- $(\text{CH})_x$ , bending and twisting of chains, local interactions between chains, and bond-length disorder. Finally, in Sec. IV(d) we present a summary and conclusions, with some speculation about the experimental anomalies mentioned above.

#### (b) Topological Defects

If it could be synthesized, crystalline trans- $(\text{CH})_x$  would consist of infinite zig-zag chains made up of CH units. The structure is shown in Fig. 21 for a terminated chain, a defect which will be discussed shortly. The system is planar, with each carbon forming  $sp^2$  hybrids in order to bond with its neighbors. There are two filled bonding orbitals per CH unit, and a half filled  $\pi$  band. The theory of the Peierls transition predicts that the system will undergo an asymmetric distortion

Figure 21

Structure of a trans-polyacetylene chain terminating on a radical R.



in order to double the periodicity and open a gap at the Fermi energy,<sup>67</sup> thus explaining (at least qualitatively) the alternation between long (weak) and short (strong) bonds which is sketched in Fig. 21.

By symmetry, the  $p_z$  orbitals decouple rigorously from all other orbitals in the system and may be treated independently. This is very useful because it is the  $\pi$  orbitals which control much of the interesting physics of the system, including the optical gap and the possible existence of solitons. The interactions between the  $\pi$  orbitals are shown schematically in Fig. 22(a) for the structure of Fig. 21. A defect of this kind will be designated  $1F_w$  because the last  $\pi$  bond connecting to the one-fold site is a weak bond. (The radical R, which may be H,  $CH_3$ , etc., is assumed to have no states of odd-z parity in the neighborhood of the Fermi energy and is therefore ignored.) In Fig. 22(b) we show a chain containing a soliton (S) and terminating on a strong bond ( $1F_s$  defect); we will show presently that the reaction

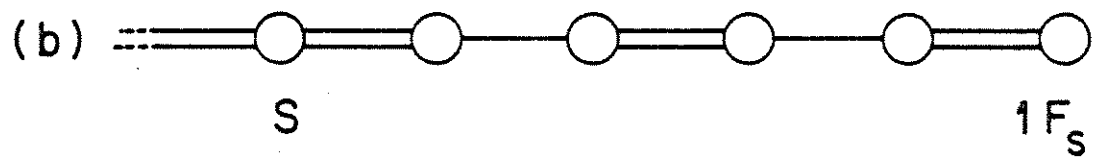
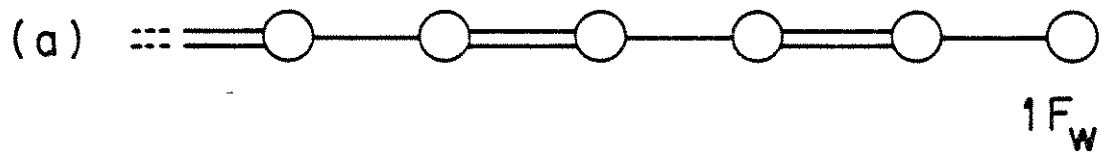


is exothermic. This spontaneous emission of a soliton by the  $1F_w$  surface configuration has the remarkable consequence that all odd-membered chains must contain a soliton somewhere in their interior.

Figure 22

(a) Schematic interaction diagram for  $\pi$  orbitals of a chain which terminates on a weak  $\pi$  bond ( $1F_w$  site). (b) Same for chain containing a soliton (S) and terminating on a strong  $\pi$  bond ( $1F_s$  site). The circles represent  $p_z$  orbitals, single lines represent weak bonds (interaction  $V_w$ ), and double lines represent strong bonds (interaction  $V_s$ ).





We now turn to a consideration of the electronic structure of such defects. Before presenting the results of the detailed calculations on these structures, we can learn a great deal about the nature of the electronic states, particularly the existence of the midgap states, by some general arguments.

All topological defects (chain ends, crosslinks, solitons) can be characterized by a bond-alternation parity (BAP) which is defined in the following way. For each semi-infinite chain leaving the defect, sever the chain on a weak bond; then count the atoms remaining in the finite central cluster. If this number is even (odd) the defect has even (odd) BAP. Now if one takes an elementary one-electron Hamiltonian whose zero of energy is at the center of the Peierls gap, it can be shown that the density of states is an even function of energy for all such defects, and that as a consequence, all defects of even (odd) BAP have an even (odd) number of localized midgap states. Thus, for example, the  $1F_w$  defect and the soliton both have odd BAP, and consequently both have a defect state at midgap. The  $1F_s$  defect has even BAP and has no defect states (only midband resonances). Note that BAP is a conserved quantity in reactions such as Eq. (10) above.

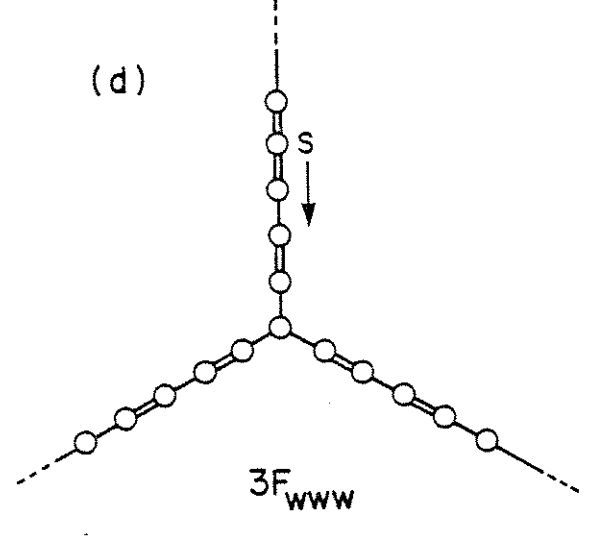
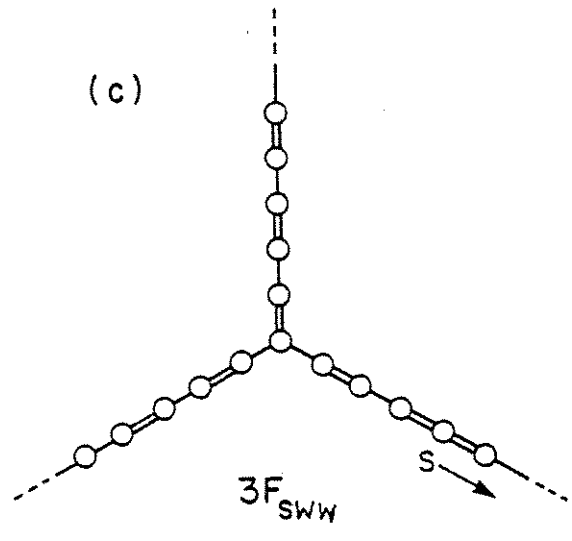
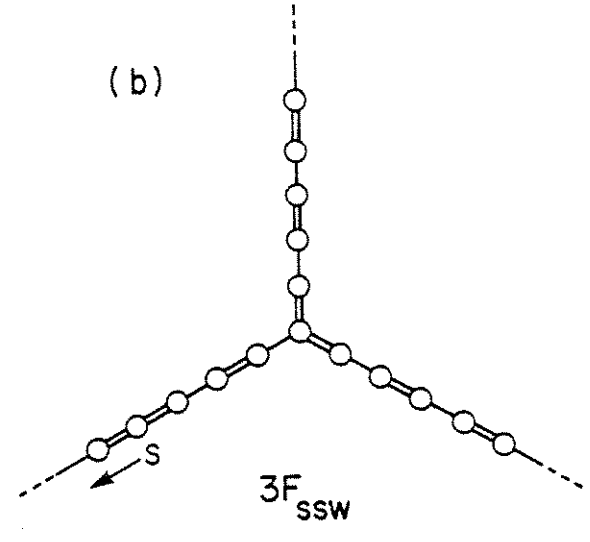
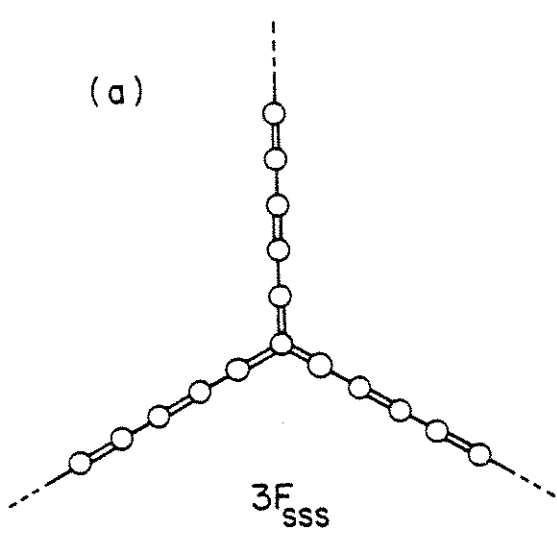
The existence of midgap states can be understood intuitively in the limit  $|V_s| \gg |V_w|$ . In this case one can think of clusters (usually pairs) of atoms connected by strong bonds as "molecules" which then interact slightly with one another via weak bonds. In Fig 2(a), for example, each strongly bonded

pair of sites gives rise to bonding and antibonding molecular orbitals at  $E=\pm V_s$ ; these are then broadened into bands whose width is on the order of  $V_w$ . The last  $\pi$  orbital of the chain has been left out, however, and gives rise to a midgap state at  $E=0$ . In the case of Fig. 22(b), the "triatomic molecule" at the soliton gives rise to an  $E=0$  midgap state, as well as states at  $E=\pm V_s/2$ . A glance at the planar crosslink defects of Fig. 23 indicates that the  $3F_{sss}$ ,  $3F_{ssw}$ ,  $3F_{sww}$ , and  $3F_{www}$  defects will have 2, 1, 0, and 1 midgap states, respectively, these being the number of  $E=0$  eigenvalues of the corresponding central molecule. (Of course the assumption  $|V_s| \gg |V_w|$  is unphysical, but the existence of midgap states is symmetry related and remains invariant as long as  $|V_s| > |V_w|$ .)

It is straightforward to use Green's function techniques to solve for the electronic structure of the defects shown in Figs. 22 and 23. However, this is somewhat pointless because we have no guarantee that any of these defects, in their present form, will be stable in the material. Firstly, we expect distortions (relaxations) of the system in the neighborhood of the defects; for example, it is well known from previous theoretical work that the soliton will not remain confined to a pair of strong bonds, as shown in Fig. 22(b), but will spread out over many sites, creating a region in which the degree of bond alternation is reduced.<sup>61</sup> Secondly, and more drastically, we will show that several of these defects are unstable against the emission of a soliton down one of the chains. Thus our

Figure 23

Planar crosslink configurations (a)  $3F_{sss}$ , (b)  $3F_{ssw}$ , (c)  $3F_{sww}$ , and (d)  $3F_{www}$ . Note that the centers interconvert via the reactions  $3F_{sss} \rightarrow S+3F_{ssw}$ ,  $3F_{ssw} \rightarrow S+3F_{sww}$ , and  $3F_{sww}+S \rightarrow 3F_{www}$ .



immediate task is to develop a means by which to calculate total energies of various defect configurations, and thereby find stable defects (and their energies) by searching for total energy minima in configuration space.

Following Su, Schrieffer, and Heeger,<sup>61</sup> we associate to each site  $n$  a coordinate  $u_n$  for the displacement of the  $n^{\text{th}}$  CH unit parallel to the chain. The Hamiltonian is then written

$$\begin{aligned}
 H = & - \sum_{n,s} v_{n+1,n} C_{n+1,s} C_{n,s} + \text{H.c.} \\
 & + \sum_n K(u_{n+1}-u_n)^2/2 - \sum_n \Gamma(u_{n+1}-u_n) \\
 & + \sum_n M u_n^2/2 .
 \end{aligned} \tag{11}$$

Here  $v_{n+1,n}$  is the transfer integral,  $C_{n,s}$  the annihilation operator for spin  $s$  on site  $n$ ,  $K$  the bond stretch spring constant, and  $M$  the CH mass. The Hamiltonian is identical to that of Su, Schrieffer, and Heeger except that the third term has been added to stabilize the chain against contraction due to  $\pi$ -bonding. This term has no effect as long as attention is restricted to trial solutions of the infinite chain which do not vary the lattice parameter, but it is important at defect sites because local contractions can and do occur.

In linear order

$$v_{n+1,n} = v_0 - \alpha(u_{n+1}-u_n) \tag{12}$$

so that the static Hamiltonian becomes

$$\begin{aligned} \tilde{H} = & - \sum_{n,s} [(1+x_n)C_{n+1,s}C_{n,s} + \text{H.c.}] \\ & + \sum_n \kappa x_n^2/2 + \sum_n \gamma x_n, \end{aligned} \quad (13)$$

where the renormalizations  $\tilde{H} = H/v_0$ ,  $x_n = (v_{n+1,n} - v_0)/v_0$ ,  $\kappa = K v_0/\alpha^2$ , and  $\gamma = \Gamma/\alpha$  have been applied to obtain a dimensionless Hamiltonian. For the undimerized chain one has  $x_n = x = \text{constant}$  and the energy per CH unit is

$$\tilde{E} = -(4/\pi)(1+x) + \kappa x^2/2 + \gamma x, \quad (14)$$

so that to stabilize the chain we choose  $\gamma = 4/\pi$ . The dimerized chain has  $x_n = (-1)^n x$  and

$$\tilde{E} = -f(x) + \kappa x^2/2 \quad (15)$$

where  $f$  is obtained by integrating numerically over the density of states. It has the form<sup>62</sup>

$$f(x) = x^2 \pi^{-1} [2 \ln(4/x) - 1] + O(x^4). \quad (16)$$

Taking  $\partial E/\partial x = 0$  for stability determines

$$\kappa = f'(x)/x. \quad (17)$$

Notice that the couplings  $\kappa$  and  $\gamma$  of Eq. (13) are determined solely by the dimerization parameter  $x$ . Thus, within the model,  $x$  alone determines the length of solitons (in units of the lattice constant), the relaxations which occur at defects, the relative total energies of various relaxed defects, etc. For completeness, however, we now derive estimates for the real spring constant  $K$  and coupling  $\alpha$ , obtaining results somewhat different from those reported previously.

We have chosen a  $\pi$  bandwidth of 12 eV ( $v_0 = 3$  eV) and Peierls gap of 1.4 eV, which leads to  $x = 0.117$  and  $\kappa = 3.255$ . To proceed further, we make use of the experimentally observed<sup>68</sup> Ramon-active mode at  $1470 \text{ cm}^{-1}$  which corresponds to a zone-center optical phonon in which CH units move primarily parallel to the chain axis, with neighboring CH units  $180^\circ$  out of phase.<sup>69</sup> To simplify the analysis we consider the undimerized geometry, and find the screened spring constant  $K_T$  is given by  $K_T = (1/4)Mw^2 = 25.9 \text{ eV-Å}^{-2}$  for  $M=13$  amu. This is related to the unscreened spring constant  $K$  of Eq. (11) by

$$K_T = K + K_{\text{scr}} = (\alpha^2/v_0) [\kappa - f'(x)]$$

where  $K_{\text{scr}}$  corresponds to the electronic screening from the first term in the Hamiltonian of Eq. (11). For  $x = 0.117$  we find  $k = f'(x)/x = 3.255$  and  $f'(x) = 2.028$ , so that  $\alpha^2/v_0 = 21.1 \text{ eV-Å}^{-2}$  and  $K = 68.6 \text{ Å}^{-2}$ . This differs from the estimates of Su,



Schreiffer, and Heeger, who give  $\alpha^2/v_0 = 6.8 \text{ eV-Å}^{-2}$  and  $K/2 = 10.5 \text{ eV-Å}^{-2}$  (Ref. 61).

Taking  $v_0 = 3.0 \text{ eV}$ , we find  $\alpha = 8.0 \text{ eV-Å}^{-1}$ , in agreement with Mele and Rice<sup>70</sup> who give  $\alpha = 6.9 \text{ eV-Å}^{-1}$ . Then from Eq. (12) we obtain  $0.022 \text{ Å}$  for the horizontal displacement of the CH units, and  $0.038 \text{ Å}$  for the bond-length change. Su et al.<sup>61</sup> give  $\alpha = 4.1 \text{ eV-Å}^{-1}$  and a bond-length change of  $0.073 \text{ Å}$ . We believe our smaller estimate of the bond-length alternation to be in better agreement with bond-length estimates from other work.<sup>71</sup>

For each defect configuration, Green's-function techniques are used to determine the local density of states  $N_n(E)$  at each site  $n$ . The formalism used is a straightforward specialization of the cluster Bethe-lattice method<sup>72</sup> to the case of twofold coordination (see Appendix A). The contribution of each site to the defect creation energy is, from Eq. (13),

$$\delta\tilde{E}_n = \int_{-\infty}^{\epsilon_F} \epsilon N_n(\epsilon) d\epsilon + \sum_{n'} (\kappa x_{n'}^2 + \gamma x_{n'}) / 2 - \tilde{E}_0 \quad (18)$$

where the sum is over sites which are nearest neighbors to  $n$ , and  $\tilde{E}_0$  is the total energy per site of the uniformly dimerized chain. The sum  $\sum_n \delta\tilde{E}_n$  converges rapidly and serves to uniquely define the defect total energy. Note that the Hamiltonian of Eq. (13) does not include the changes in  $\sigma$  bonding as one goes to onefold or threefold defects. Therefore, it will only give total energies which are valid for comparing defects of the

same coordination. For each defect configuration, the electronic part of Eq. (13) is solved using Green's-function techniques and the lattice energies are simply summed. As a test case we have calculated the ground-state soliton with trial function  $u_n = (v_0 x / 2\alpha) w_n$  given by

$$w_n = (-1)^n \tanh(n/l) \quad (19a)$$

to have length  $l = 9$  and  $E = v_0 \tilde{E} = 0.442$  eV, in agreement with Su Schrieffer, and Heeger.<sup>61</sup>

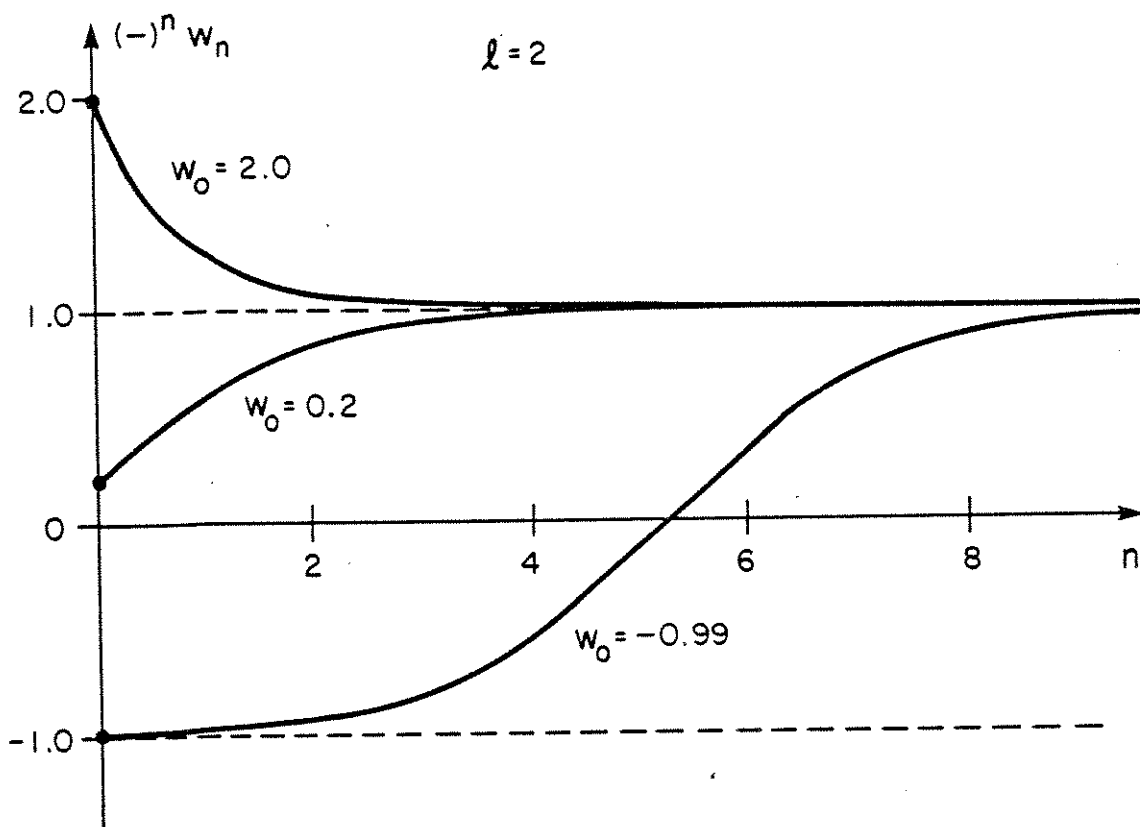
It now becomes necessary to choose trial functions for the atomic positions in the neighborhood of defects. One wants a trial solution which (i) heals to the normal chain with dimerization  $x$  far from the defect site, (ii) allows for regions of decreased (or increased) dimerization near defects, and (iii) allows the defect site to relax in toward (or out away from) any chain connected to it. We choose

$$w_n = (-1)^n \frac{w_0 + \tanh(n/l)}{1 + w_0 \tanh(n/l)} \quad (19b)$$

where the labeling  $n=0,1,2,\dots$  proceeds from the defect down the chain. The position of the defect site  $w_0$  and decay length  $l$  are the two free parameters. Several examples are shown in Fig. 24. [For  $|w_0| < 1$  these are truncated solitons since Eq. (19) can be rewritten as  $w_n = (-1)^n \tanh(n/l + \tanh^{-1} w_0)$ .] We have

Figure 24

Trial solutions for the order parameter  $(-)^n w_n$  as a function of  $n$ , following Eq. (19). In all cases  $l=2$ ; three choices of  $w_0$  are shown.



experimented with other trial solutions, but the essential features are obtained with Eq. (19).

In Figs. 25 and 26 we show the calculated total energies of the  $1F_s$  and  $1F_w$  defects as a function of  $w_0$  and  $l$ . The  $1F_s$  defect has an energy minimum at  $l=5$  and  $w_0=1.75$  with  $E=0.808$  eV. The result  $w_0 > 1$  indicates that the final site relaxes in toward the rest of the chain compared to the uniformly dimerized case, and there is a region of increased bond alternation near the defect. The  $1F_w$  defect, however, shows a remarkable behavior: no stable configuration exists, but the system wants to relax towards  $w_0 = -1$ . As can be seen from Fig. 26, this corresponds to the emission of a soliton, leaving behind an unrelaxed  $1F_s$  defect. Furthermore, by counting bonds it becomes clear that any finite chain with an odd number of CH units must contain either a  $1F_w$  chain or a soliton; we have shown that the latter is energetically favorable, and thus all odd chains must contain a soliton.

For the three-fold defects of Fig. 23 we used trial solutions with a common  $l$  for all three chains and different  $w_0$  for different kinds of outgoing chains. We found that the decays  $3F_{sss} \rightarrow 3F_{ssw} + S$  and  $3F_{ssw} \rightarrow 3F_{sww} + S$  are exothermic, and that the only stable defects are  $3F_{sww}$  ( $l = 4$ ,  $w_0^1 = w_0^2 = 1.3$ ,  $w_0^3 = 0.35$ ,  $E = -0.743$  eV) and  $3F_{www}$  ( $l = 8$ ,  $w_0^1 = w_0^2 = w_0^3 = 0.5$ ,  $E = -0.447$  eV). As in the case of the one-fold defects, the configuration with no midgap state is, not surprisingly, lowest in energy. However, note that  $E(3F_{www}) - E(3F_{sww}) < E(\text{soliton})$  so

Figure 25

Total energy in eV of  $1F_s$  defect as a function of  $w_0$  and  $l$ .  
The + marks the energy minimum.

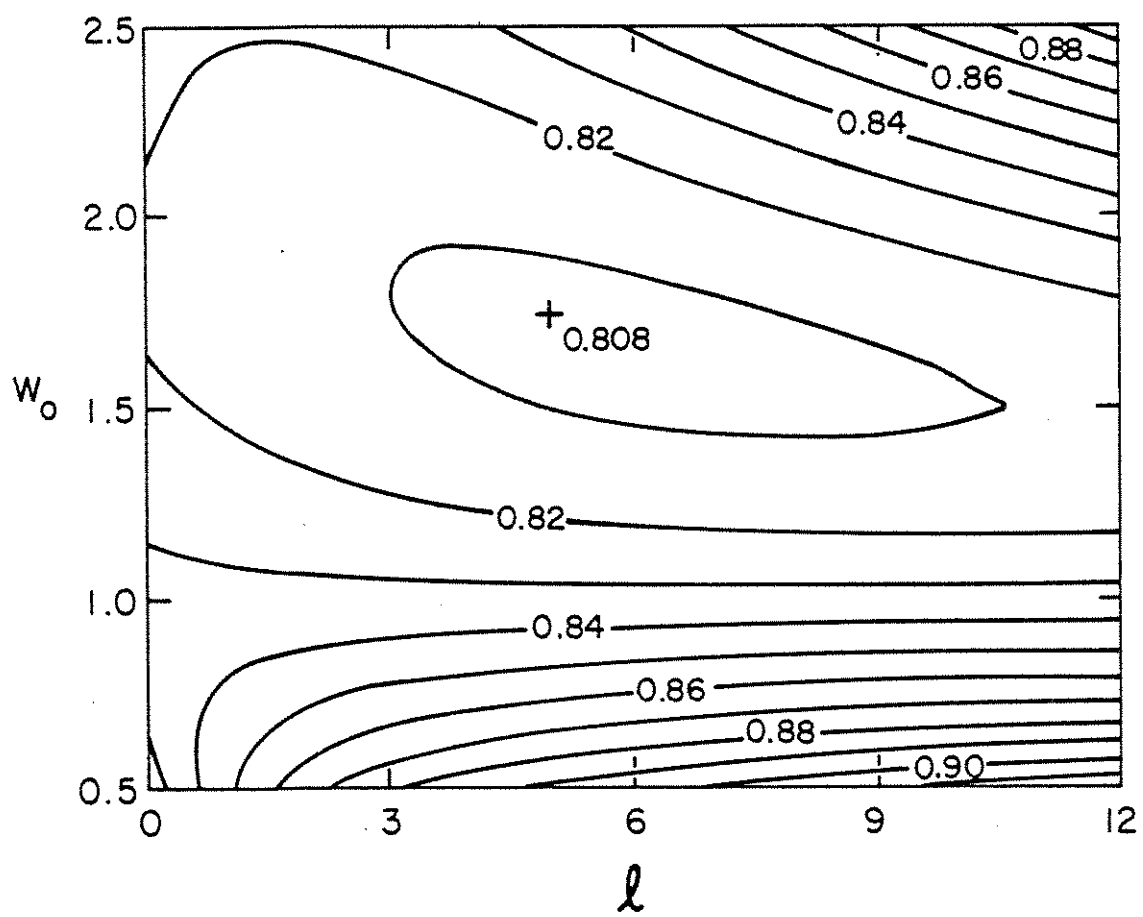
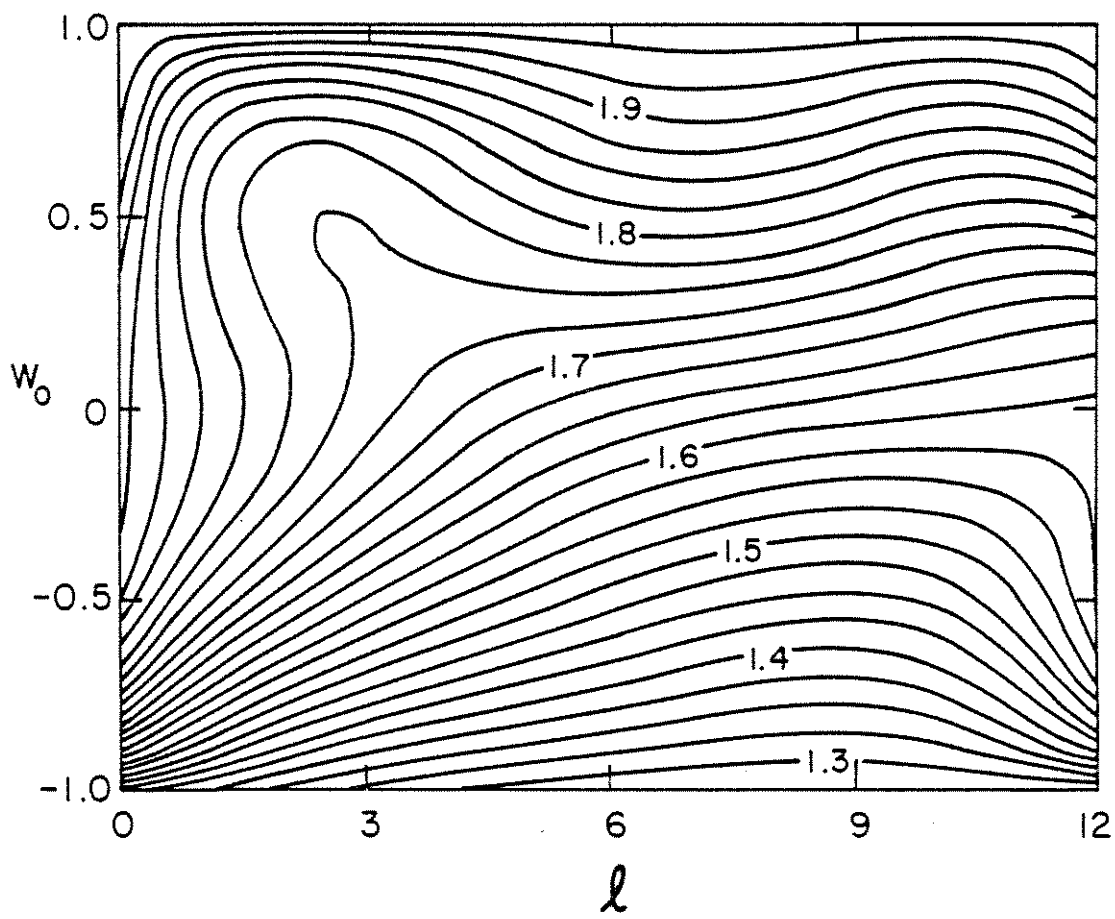


Figure 26

Total energy in eV of  $1F_w$  defect as a function of  $w_0$  and  $l$ .



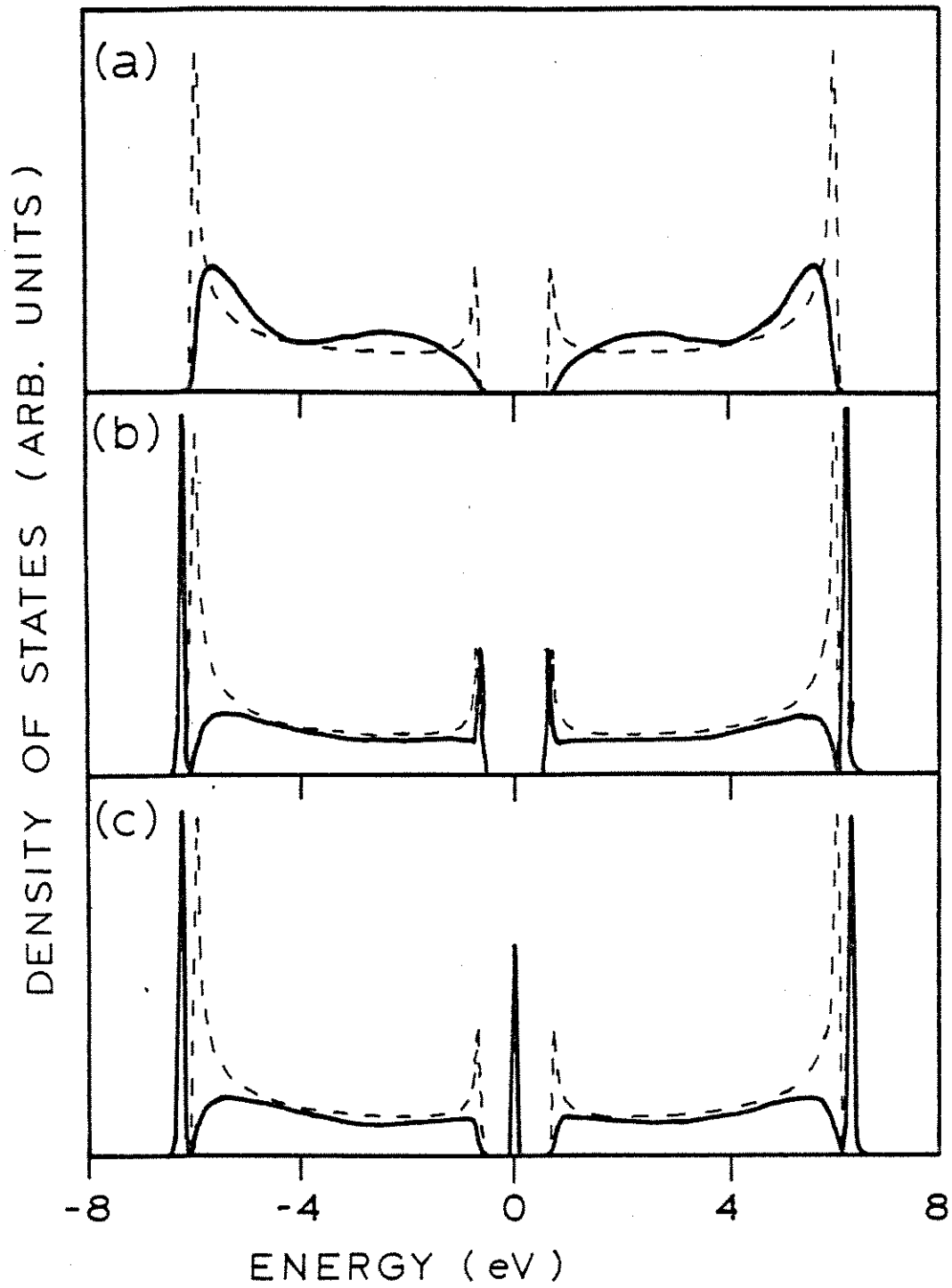


that the  $3F_{www}$  is stable and will not decay into  $3F_{sww}$  by soliton emission. This contrasts with the situation for one-fold defects, where we found that only the  $1F_s$  was stable. Since the  $3F_{www}$  defect has a midgap state, we have the remarkable result that there is a stable threefold defect which is either paramagnetic (if half occupied) or charged (if fully occupied or empty). In the latter case, it is possible that such a defect would show up as a broad IR-active feature at phonon frequencies. Of course, if free solitons are available, the reaction  $3F_{www} + S \rightarrow 3F_{sww}$  will purge such defects from the material.

In summary, the stable bond-coordination defects identified to date are the  $1F_s$ ,  $3F_{sww}$ , and  $3F_{www}$  defects. Figure 27 shows the local density of states averaged over a cluster of sites near the defect for each of these three species. Figure 27(a) shows that the  $1F_s$  defect causes a considerable shift of weight from the sharp 1D Peierls edges deeper into the bands. This is suggestive in terms of the experimentally observed absence of a sharp optical edge, but the effect has been exaggerated in Fig. 27(a) by averaging the local density of states over a small cluster (four atoms) and consequently a very high density of broken chains would be needed to explain the experimental result. Notice in Fig. 27(b) that the  $3F_{sww}$  defect gives rise to shallow trap states 0.06 eV from the Peierls edges, but no deep gap states. Finally for the  $3F_{www}$  defect, the formation of the midgap state subtracts a great deal of

Figure 27

Calculated densities of states near stable defects (solid curves) and bulk densities of states (dashed curves) for comparison. (a)  $1F_s$  defect, (b)  $3F_{sww}$  defect, and (c)  $3F_{www}$  defect.



weight from the Peierls edges [the effect is underestimated by Fig. 27(c) due to averaging the density of states over a small cluster], but any reasonable density of such defects in the pristine material should give rise to noticeable midgap optical absorption (which has not been reported to date). Similar conclusions about the relative stability of the  $1F_s$  and  $1F_w$  have recently been discussed by Su.<sup>73</sup>

A variety of other topological defects could occur, such as nonplanar crosslinks or benzene-ring chain terminations, but the list of such possibilities is too extensive to pursue here. We prefer to turn our attention at this point to a class of nontopological defects in which we consider only infinite chains with no reversals of bond alternation. There remains considerable variety to be explored in this realm.

### (c) Structural Disorder

Even if every polymeric unit in the polyacetylene film were infinite and contained no solitons, there could still be various kinds of disorder present, including regions of  $\text{cis}-(\text{CH})_x$ , bending and twisting of chains, local interactions between touching chains, and variations in amplitude of bond alternation. This disorder could affect the density of states at the Peierls edges or deeper in the valence bands, with consequences for the interpretation of optical and photoemission experiments. We shall consider each of these forms of structural disorder in turn.

In order to do so, we have developed a tight-binding model by fitting to a first-principles calculation on trans-(CH)<sub>x</sub>. The calculation of Karpfen and Petkov<sup>56</sup> was chosen for fitting because its bandwidths appear to agree with photoemission better than other calculations available.<sup>64,74</sup> The resulting tight-binding Hamiltonian is detailed in Table III. Briefly, the basis consists of three sp<sup>2</sup> hybrids, the p<sub>z</sub> and the H orbitals in each CH unit. Interactions between all hybrids on nearest neighbors are included.

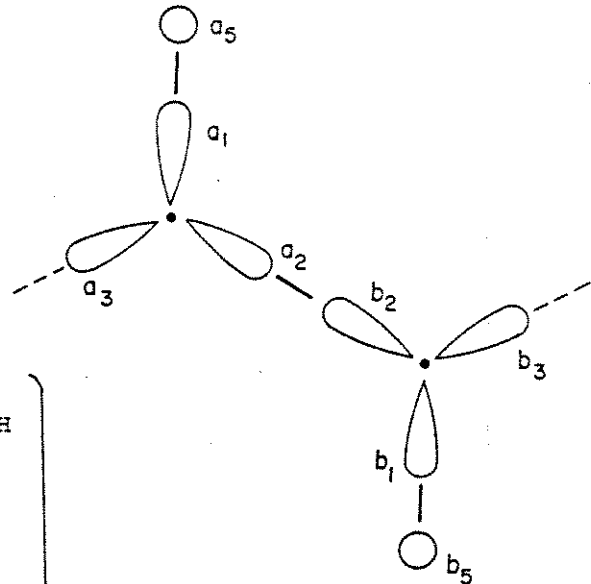
Figure 28 shows the resulting density of states for trans, in cis-transoid, and trans-cisoid polyacetylene. The p<sub>z</sub> orbitals continue to decouple rigorously, so that the region of the Peierls gap is unaffected (assuming the degree of bond alternation is unchanged). However, there are clearly considerable changes deeper in the σ valence bands. If many chains in the film retain regions of cis-isomerization, the resulting density of states would resemble a weighted average of Figs. 28(a)-(c). However, it was found that such averaging would not materially improve the fit to photoemission by the trans-isomer alone.

In Fig. 29 we show what is meant by the terms bend and twist. Both kinds of disorders involve dihedral angle variations only; the sense of the dihedral angle is constant for twist and alternates for bend. ("Splay," or coplanar bending, involves bond angle rather than dihedral angle variations and was therefore dismissed as a less likely possibility.) We have

Table III

Tight-binding Hamiltonian for  $(\text{CH})_x$ . Insert shows basis orbitals:  $a_1, a_2, a_3$  are carbon  $sp^2$  hybrids,  $a_4$  (not shown) is  $p_z$ ,  $a_5$  is a hydrogen s orbital.  $H_0$  gives the form of the on-site part of the Hamiltonian,  $V$  gives the nearest-neighbor coupling,  $\theta$  is the dihedral angle. Parameters are given in the bottom panel.

$$H_0 = \begin{pmatrix} -\epsilon & -\epsilon & -\epsilon & 0 & -V_{\text{CH}} \\ -\epsilon & -\epsilon & -\epsilon & 0 & 0 \\ -\epsilon & -\epsilon & -\epsilon & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -V_{\text{CH}} & 0 & 0 & 0 & E_{\text{H}} \end{pmatrix}$$



$$V = \begin{pmatrix} (V_{\pi} \cos \theta)/2 & -V_{\sigma} & -(V_{\pi} \cos \theta)/2 & (V_{\pi} \sin \theta)/\sqrt{2} & 0 \\ -V_{\sigma} & -V_{\sigma} & -V_{\sigma} & 0 & 0 \\ -(V_{\pi} \cos \theta)/2 & -V_{\sigma} & (V_{\pi} \cos \theta)/2 & -(V_{\pi} \sin \theta)/\sqrt{2} & 0 \\ (V_{\pi} \cos \theta)/\sqrt{2} & 0 & -(V_{\pi} \sin \theta)/\sqrt{2} & -V_{\pi} \cos \theta & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

On site	Strong bond	Weak bond
$\epsilon=0.65$ eV	$V_{\sigma}=18.87$ eV	$15.91$ eV
$E_{\text{H}}=8.30$ eV	$V_{\pi}=4.45$ eV	$3.75$ eV
$V_{\text{CH}}=15.65$ eV	$V_{\sigma}=3.26$ eV	$2.74$ eV

Figure 28

Tight binding densities of states for crystalline isomers of polyacetylene. (a) Trans-(CH)<sub>x</sub>, (b) cis-transoid (CH)<sub>x</sub>, (c) trans-cisoid (CH)<sub>x</sub>. Insets show the geometries of each isomer.



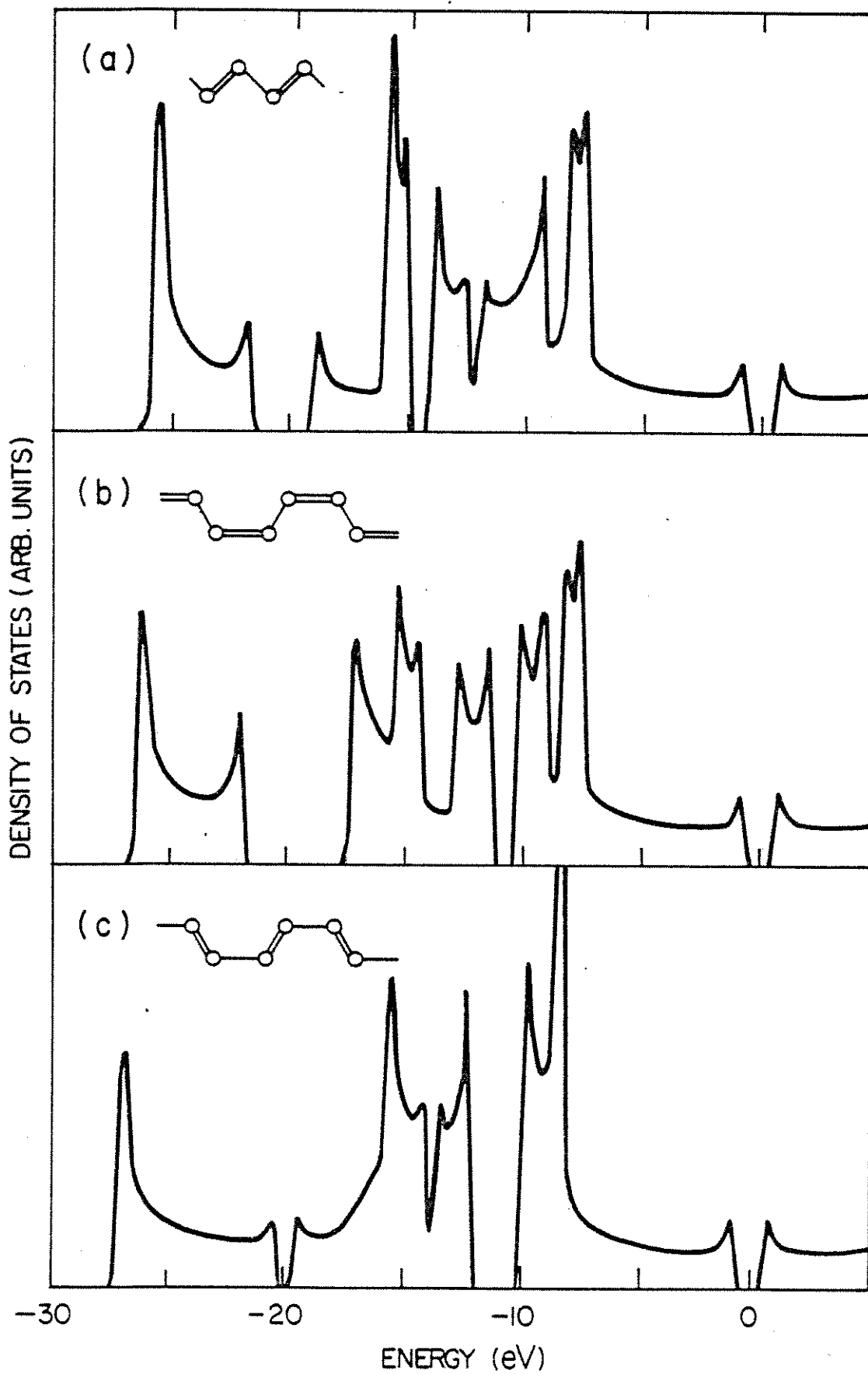
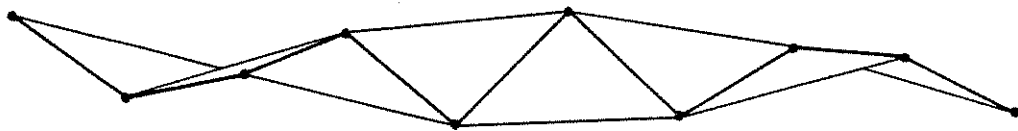


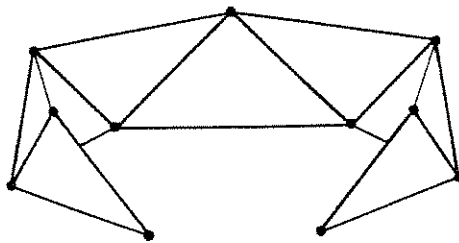
Figure 29

Structures resulting from dihedral angle variations may be constructed by folding a strip of paper. Each vertex represents a carbon atom and each heavy line represents a carbon-carbon nearest-neighbor C-C bond. A flat strip of paper corresponds to undimerized trans-(CH)<sub>x</sub>. (a) A strongly twisted chain formed by uniform dihedral angle variations of 45°. (b) A strongly bent chain formed by alternating dihedral angles of ±45°.

(a) TWIST



(b) BEND

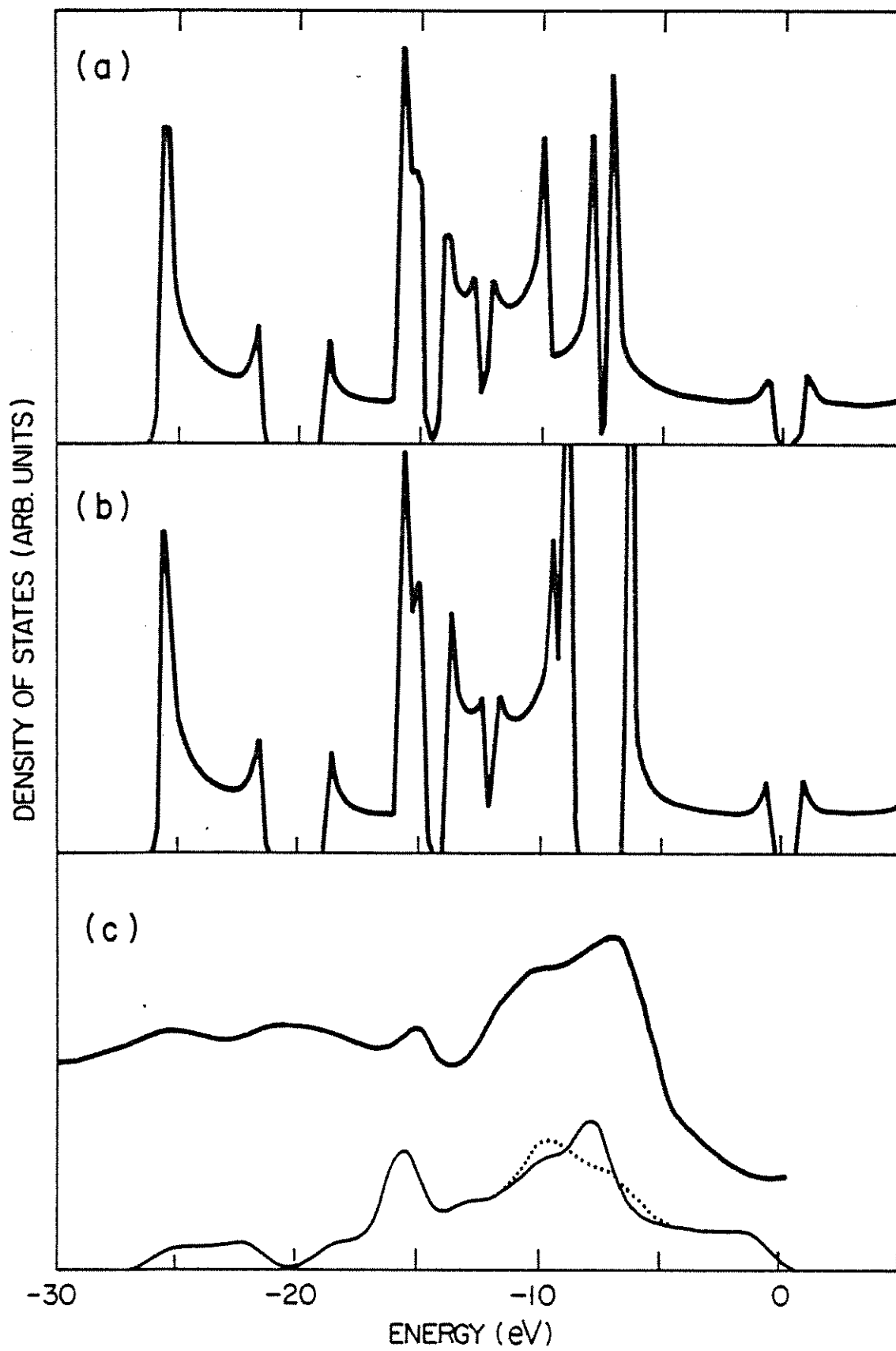


calculated the density of states for structural models with constant bend or twist. These give rise to the circular or helical patterns of Fig. 29 and have the property of being periodic in interaction space. Therefore the simplest approach is just to calculate the 1D band structure as a function of the "wave vector"  $\phi$  which specifies the relative phase of the wave function on neighboring  $C_2H_2$  units. The results are shown in Figs. 10(a) and 10(b). While in principle the bonding and antibonding hybrids could interact strongly with the  $p_z$  orbitals, this does not occur for twist at  $\phi=0$  or for bend at  $\phi=\pi$ , because the interactions along the strong and weak bonds are out of phase and almost cancel. For bend the interaction near  $\phi=0$  is strong and opens a gap near -8 eV where the  $\pi$  band previously crossed a bonding band. For twist near  $\phi=\pi$  it manifests itself weakly as an asymmetry in the Peierls edges because there are no bonding or antibonding states in that region of the spectrum.

It is quite easy to imagine virtually every chain in the material having some modest amount of bend. We have calculated the density of states for different dihedral angles (in  $5^\circ$  increments). Figure 30(a) shows the results of superposing these densities of states according to a Gaussian weighted average with an rms dihedral angle of  $10^\circ$ . Also shown are the density of states without bend and the experimental photoemission of Duke et al.<sup>64</sup>. Inclusion of some bend allows a modest improvement in the fit to experiment.

Figure 30

(a) Tight-binding densities of states for trans-(CH)<sub>x</sub> with constant twist of 15°. (b) Same for bend of 15°. (c) Light solid line: tight-binding p-state projection of density of states for trans-(CH)<sub>x</sub> without twist or bend broadened by a 1-eV-wide Gaussian. Dotted line: same for rms bend of 10°. Heavy solid line: experimental photoemission of Ref. 64.

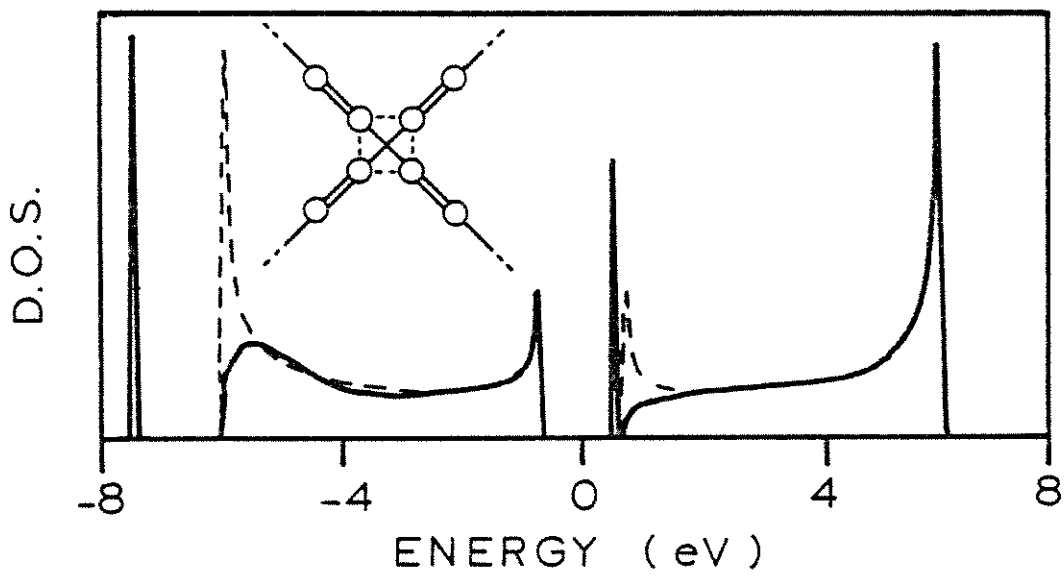


In a search for interactions which might show a strong influence on the Peierls region of the spectrum, we have considered the case of local interchain interactions and of bond-strength disorder. The interaction configuration and the resulting density of states for one model of local interchain interactions between crossing chains is shown in Fig. 31, with an exaggerated value of 1.5 eV for the interchain interaction  $w$ . The  $p_z$  orbitals once again decouple. Because of the presence of oddfold rings in interaction space, the density of states is not constrained to be even in energy. We find that an electron trap state has been pulled out of the conduction band. A  $90^\circ$  rotation of both chains about their axes induces a similar effect at the valence-band edge. Thus a distribution of interchain interaction along a pair of chains would be expected to induce a slight broadening of the band edges, suppressing the one-dimensional  $1/\sqrt{E}$  singularities at each edge. Such an effect has been described by Grant and Batra<sup>74</sup> who have performed calculations on an idealized three-dimensional model of  $\text{cis}-(\text{CH})_x$ . Since there is arguably a distribution of interchain separations in these quasicrystalline samples, this broadening is probably more appropriately described as the band tails of an otherwise unperturbed 1D spectrum than the  $\sqrt{E}$  3D threshold obtained in Ref. 74. The former interpretation requires that the interchain conductivity occur via hopping between localized tail states.

Figure 31

Density of states in the vicinity of a localized interaction between a pair of crossing  $(CH)_x$  chains. Insert gives a schematic interaction diagram; single lines denote  $v_s$ , dotted lines  $v_w$ .





Finally, we consider the possibility that bond lengths vary randomly away from uniform bond alternation. Bond-length disorder should be considered less probable than stochastic variations in bond angles or interchain separations. However, anticipating that local variations in the packing density of contiguous chains may introduce significant fluctuations in the total crystal potential experienced by the valence electrons, we will consider bond-length fluctuations as well. According to our estimate the bond-length change responsible for the gap in crystalline  $(\text{CH})_x$  is only  $\pm 2.7\%$ , so it is quite easy to imagine that static bond-length disorder of only 1-2% could begin to wash out the gap.

We have tested this idea by constructing a chain segment of 5200 CH units with site positions given by  $x_n = (-1)^n(1+s_n)$ , where  $s_n$  is chosen independently according to a Gaussian distribution with  $\sigma=0.5$  and zero mean. (For  $\sigma=0$  this would be the perfectly dimerized chain.) Then the bond lengths also follow a Gaussian distribution, with  $\sigma' = (\sigma/\sqrt{2}) \times 2.7\%$  or 1% of the bond length. We then connected uniformly dimerized chains to either end of this segment, and averaged the calculated local density of states over the central 5000 atoms of the segment. Since this is a one-dimensional model, all states are Anderson localized, so that the total density of states is really a sum of delta functions corresponding to each localized state. Therefore, to obtain a smooth average it is necessary to study a segment which is long enough to have many delta functions in

every energy interval. Even with 5000 atoms per segment, it was necessary to average the results over four such segments.

The results are shown in Fig. 32. With only a 1% rms bond-length change on each C-C bond, we find a very substantial broadening of the 1D band edges. This effect is perhaps better expressed following the arguments presented in Ref. 67 and 75 which treat the electronic spectrum near the Peierls gap in one-dimensional systems in which the lattice distortion becomes uncorrelated over a length  $\xi$ . Although in these treatments  $\xi$  is envisioned as resulting from thermal disorder, we readily generalize  $\xi$  to include static disorder. One obtains a lifetime broadening,  $\Gamma$ , of a "Bloch" state

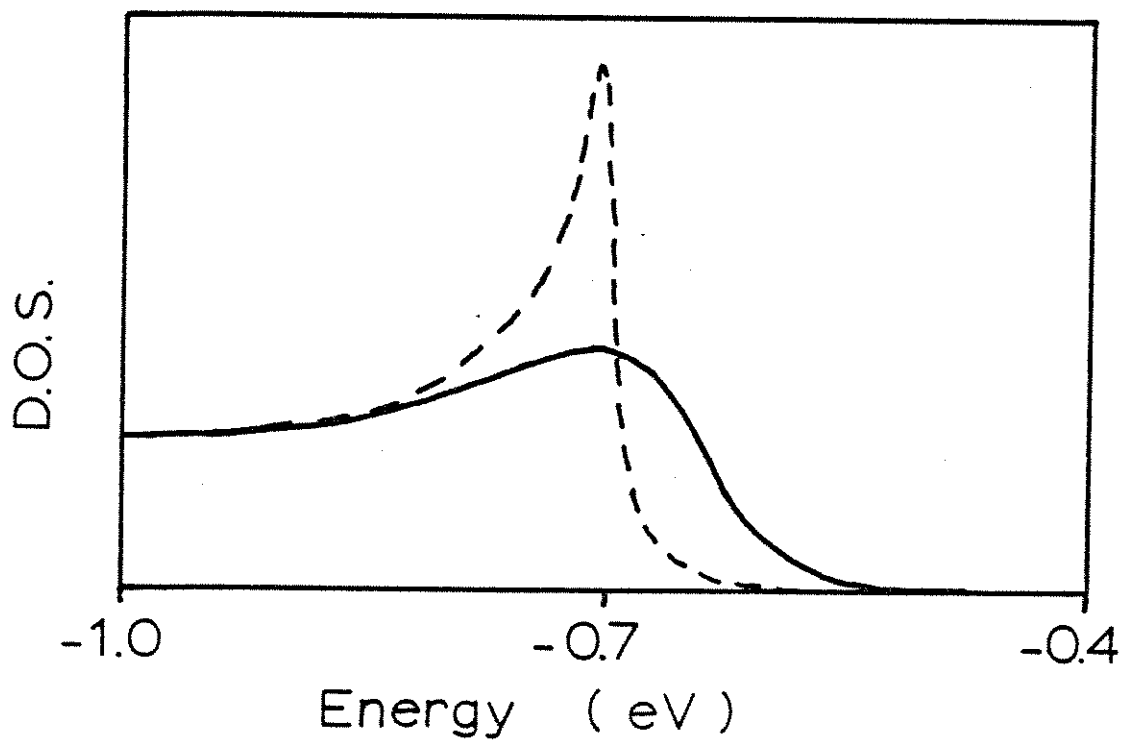
$$\Gamma = \hbar v_F \xi^{-1}, \quad (20)$$

where  $v_F$  is the Fermi velocity in the metallic (undistorted) state.  $\Gamma$  has the effect of "smearing" the electronic spectrum replacing the Peierls gap with a "pseudogap." We estimate  $\hbar v_F = 8 \text{ eV-}\overset{\circ}{\text{A}}$  in  $(\text{CH})_x$ ; consequently a mean coherence length of  $40 \overset{\circ}{\text{A}}$  on each chain would be required to explain the observed broadening of the absorption edge, which is not an unreasonably short estimate.

We believe the results described in Figs. 31 and 32 to be the most likely explanation for the absence of a sharp edge in the experiments. Because there is no k-selection rule for a system of localized tail states, this would also explain the

Figure 32

Solid curve: average density of states near the valence-band edge for a  $(\text{CH})_x$  chain with stochastic variations of 1% rms in bond lengths. Dashed curve: crystalline  $(\text{CH})_x$  for comparison.



observed momentum independence of the absorption edge.<sup>76</sup> Because we have only carried out 1D calculations, we have no way of estimating the position of the mobility edge. However, the fact that the disorder in  $(\text{CH})_x$  need only be comparable to the bond alternation (which is already small), together with the quasi-1D nature of the system, suggests that this material may be a model system for severe Anderson localization, with an unusually large tail of localized states. Finally, we point out that the degree of broadening of the Peierls edges may be highly sample dependent since the amount of disorder is likely to depend upon preparation conditions. We should note that while the present results suggest mechanisms which broaden the absorption edge, a complete description of the optical spectrum below 2 eV should include a treatment of the excitonic final state.<sup>64</sup> In particular, photoconductivity measurements<sup>63</sup> may be taken as evidence that optical absorption is excitonic within  $\sim 0.1$  eV of threshold.

#### (d) Summary and Conclusions

We have studied many different kinds of disorder which can occur in polyacetylene films. The results indicate that moderate structural disorder, including twisting, bending, or local touching of chains, could be virtually universal without having a major impact on the electronic properties of the film. There appears to be a minor improvement in the fit to photoemission when some bend is included. Admixtures of

cis-isomerized  $(\text{CH})_x$  chains at levels approaching 10% cannot be ruled out by photoemission. Note that Raman<sup>73,75</sup> scattering has been shown to be a more sensitive probe of the cis-trans ratio on  $(\text{CH})_x$ .

By studying the stability of various possible topological defects to soliton emission and relaxation, we have identified one stable chain termination and two species of stable planar crosslinks. One of the latter has a midgap state, which is not seen in the IR absorption. The other two defects could occur at large densities (on the order of a percent) without having any dramatic effect on the density of states.

Finally, since the bond alternation is already weak, stochastic variations of only 1% in bond length can strongly broaden the 1D Peierls edges. This effect and interactions between randomly packed chains are thought to be the most likely explanations for the observed absence of a sharp edge in some samples.





## APPENDIX A BETHE LATTICE FORMALISM

In Chapters II-IV, we have repeatedly solved systems in which a tight binding Hamiltonian is applied to a cluster-Bethe-lattice structure. In Chapters II and IV, one has twofold coordination, so that a Bethe-lattice simply corresponds to a semi-infinite chain, whereas in Chapter III we must deal with a true Cayley tree.

The cluster-Bethe-lattice method (CBLM) solution using Green's function techniques has been detailed previously.<sup>77-78</sup> However, in this Appendix we report two new extensions of this method. Firstly, we have generalized the solution to allow for nonorthogonal basis orbitals. To our knowledge, such a generalization has not been previously reported. Because there are many systems for which the inclusion of nearest-neighbor overlaps is natural and useful, we outline the method here. The discussion will specialize to systems of twofold coordination, but the generalization to higher coordination is straightforward. Secondly, we have used the CBLM density of states to calculate estimates of defect total energies.

We begin by defining the Green's-function operator

$$\hat{G}(E) = (E - \hat{H})^{-1} = \sum_{\alpha} \frac{|\psi_{\alpha}\rangle\langle\psi_{\alpha}|}{E - E_{\alpha}} \quad (A1)$$

$\hat{H}$  is the Hamiltonian, the  $|\psi\rangle$  are energy eigenstates with eigenvalues  $E_\alpha$ , and the energy  $E=\varepsilon+i\delta$  is taken to have an infinitesimal imaginary part. Then

$$-(1/\pi) \text{Im tr } \hat{G} = \sum_{\alpha} \delta(\varepsilon-E_{\alpha}) = n(\varepsilon) \quad (\text{A2})$$

gives the density of states. Now choose basis vectors  $|i\rangle$ , linearly independent but not necessarily orthonormal, and construct the duals  $|i^*\rangle$  such that

$$\langle i^*|j\rangle = \delta_{ij}, \quad (\text{A3})$$

and therefore

$$\sum_i |i^*\rangle\langle i| = 1. \quad (\text{A4})$$

Define the generalized Green's-function matrix elements as

$$G_{ij} = \langle i^*|\hat{G}|j\rangle. \quad (\text{A5})$$

Using Eqs. (A1) and (A4), one easily shows

$$-(1/\pi) \text{Im tr } G = n(\varepsilon), \quad (\text{A6})$$

so that the natural definition of the local density of states is

$$- (1/\pi) \text{Im tr } G_{ii} = n_i(\varepsilon) . \quad (\text{A7})$$

Taking matrix elements, Eq. (A1) can be recast into the form

$$\sum_j \langle i | E - \hat{H} | j \rangle \langle j^* | \hat{G} | k^* \rangle = \delta_{ik} . \quad (\text{A8})$$

We define

$$\begin{aligned} S_{ij} &= \langle i | j \rangle , \\ H_{ij} &= \langle i | \hat{H} | j \rangle , \\ \tilde{G}_{ij} &= \langle i | \hat{G} | j \rangle , \end{aligned} \quad (\text{A9})$$

where the basis is assumed to have been chosen such that the  $S_{ij}$  and  $H_{ij}$  are real. Equation (A8) now becomes

$$\sum_j (E S_{ij} - H_{ij}) \tilde{G}_{jk} = \delta_{ik} . \quad (\text{A10})$$

$G$  is related to  $\tilde{G}$  by

$$G_{ij} = \sum_k \tilde{G}_{ik} S_{kj} . \quad (\text{A11})$$

We will proceed by solving first for  $\tilde{G}$ , then  $G$ .

In typical applications we label orbitals by a site label "i" and an internal label "a" which runs over the  $N$  atomic orbitals on each site. In this notation, each matrix element

( $H_{ij}$ ,  $S_{ij}$ , etc.) is itself an  $N \times N$  matrix in the internal space. We adopt this convention and furthermore restrict ourselves to first-neighbor interactions and overlaps for the purpose of this discussion. Thus for bulk Se, which corresponds to a 1-d chain of sites, the inputs are the matrices

$$\begin{aligned}
 S_0 &\equiv S_{ii} , \\
 S &\equiv S_{i,i+1} , \\
 E_0 &\equiv ES_{ii} - H_{ii} , \\
 V &\equiv H_{i,i+1} - ES_{i,i+1} .
 \end{aligned}
 \tag{A12}$$

Note that  $S_0$  and  $E_0$  are symmetric and that  $S_{i+1,i} = S^T$ , etc.

When attached to a defect, the chain segment under consideration will extend uniformly to infinity in one direction ( $i \rightarrow +\infty$ ) and will be attached to a defect at  $i=0$ . If we define the fields  $x_i$  and  $\bar{x}_i$  by

$$\begin{aligned}
 G_{i+1,i-n} &\equiv x_i V^T G_{i,i-n} \\
 G_{i-1,i+n} &\equiv \bar{x}_i V G_{i,i+n}
 \end{aligned}
 \tag{A13}$$

for  $n > 0$ , Eq. (A10) gives, e.g.,

$$\begin{aligned}
 E_0 \tilde{G}_{i,i-n} - V \tilde{G}_{i+1,i-n} - V^T \tilde{G}_{i-1,i-n} &= 0 , \\
 (E_0 x_{i-1} - V x_i V^T x_{i-1} - 1) V^T \tilde{G}_{i-1,i-n} &= 0 ,
 \end{aligned}$$

leading to

$$\begin{aligned} \chi_{i-1} &= (E_0 - V\chi_i V^T)^{-1}, \\ \bar{\chi}_{i+1} &= (E_0 - V^T \bar{\chi}_i V)^{-1}. \end{aligned} \quad (A14)$$

By means of these recursion relations, it can be shown that  $\chi_i$  and  $\bar{\chi}_i$  are symmetric matrices and are independent of  $n$ , as implied in Eq. (A13). Furthermore, because the chain is semi-infinite in one direction, we have  $\chi_i = \chi = \text{constant}$ ; this does not hold true for the  $\bar{\chi}_i$ . The latter must be calculated individually, starting at the defect and moving down the chain.

In order to calculate the local density of states on each site, we need

$$G_{ii} = \tilde{G}_{ii} S_0 + \tilde{G}_{i,i+1} S^T + \tilde{G}_{i,i-1} S. \quad (A15)$$

From Eq. (A10),

$$E_0 \tilde{G}_{ii} - V \tilde{G}_{i+1,i} - V^T \tilde{G}_{i-1,i} = 1,$$

or

$$\tilde{G}_{ii} = (E_0 - V\chi V^T - V^T \bar{\chi}_i V)^{-1}. \quad (A16)$$

Also,

$$\begin{aligned}\tilde{G}_{i,i+1} &= (\tilde{G}_{i+1,i})^T = (xV^T\tilde{G}_{ii})^T = \tilde{G}_{ii}Vx, \\ \tilde{G}_{i,i-1} &= (\tilde{G}_{i-1,i})^T = (\bar{x}_iV\tilde{G}_{ii})^T = \tilde{G}_{ii}V^T\bar{x}_i,\end{aligned}$$

so that finally Eq. (A15) becomes

$$G_{ii} = \tilde{G}_{ii} (S_0 + VxV^T + V^T\bar{x}_iS). \quad (A17)$$

The generalization to the calculation of  $G_{oo}$  at the defect site itself is straightforward.

Once we have obtained the density of states in the vicinity of the defect from the realistic tight-binding calculation, it is relatively straightforward to find the band-structure energy  $E_{BS}$ , which is just the sum of filled one-electron levels. However, the total energy is not simply  $E_{BS}$ , but contains the repulsive Coulomb interactions between ion cores and other corrections.

In order to write down the total energy, we adopt the Hartree-Fock-Slater point of view and consider the system to be composed of valence electrons and frozen ion cores of charge  $Z=+6$  for Se. Then the total energy is<sup>79</sup>

$$\begin{aligned}E_T &= \sum_i \int dr \psi_i^*(r) [-(\hbar^2/2m)\nabla^2 + V_{ion}(r)] \psi_i(r) \\ &+ (1/2) \sum_i \int dr V_H(r) |\psi_i(r)|^2 \\ &+ (3/4) \sum_i \int dr V_{xc}(r) |\psi_i(r)|^2\end{aligned}$$

$$+ (1/2) \sum_{nm} \frac{e^2 Z^2}{|R_n - R_m|} = E_{BS} + U \quad (A18)$$

where  $V_H$  and  $V_{x\alpha}$  are the Hartree and Slater exchange potentials. The corrections to  $E_{BS}$  have been lumped together in the term  $U$ . Since

$$E_{BS} = \sum_i \varepsilon_i = \sum_i \int dr \psi_i^*(r) \times \\ [-(\hbar^2/2m)\nabla^2 + V_{ion}(r) + V_H(r) + V_{x\alpha}(r)] \psi_i(r) \quad (A19)$$

we can write

$$U = \sum_i \int dr [-(1/2)V_H(r) - (1/4)V_{x\alpha}(r)] |\psi_i(r)|^2 \\ + (1/2) \sum_{nm} \frac{e^2 Z^2}{|R_n - R_m|}. \quad (A20)$$

The term  $U$  thus contains corrections due to overcounting of the Hartree and exchange-correlation energies [the Coulomb repulsion between each pair of electrons was counted twice in Eq. (A19)] and the repulsive ion-ion interaction. There is little hope of evaluating this term directly within the tight-binding model, so we turn instead to a rough empirical approach.

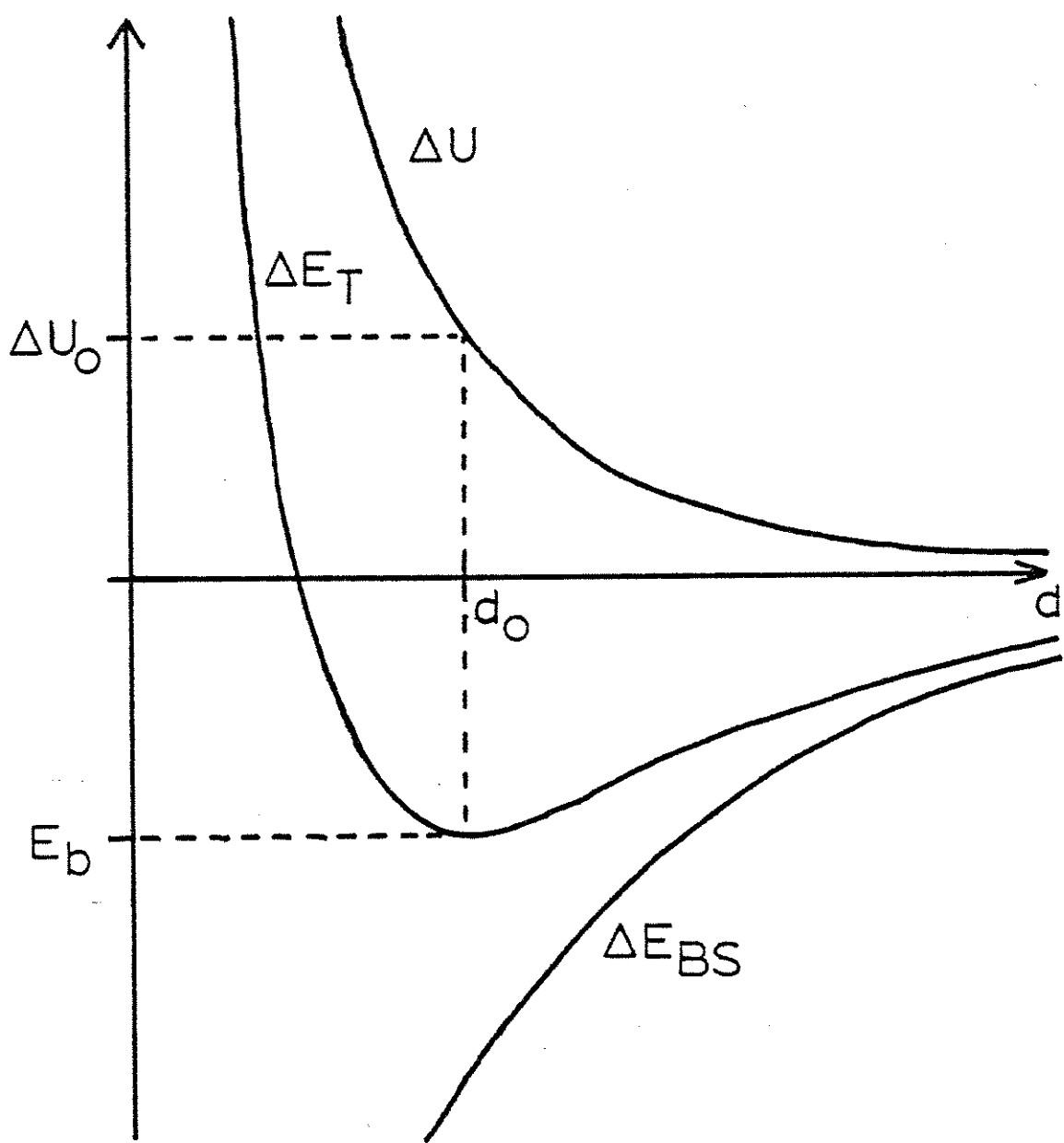
In Fig. 33 we sketch the dependence of  $E_T$ ,  $E_{BS}$ , and  $U$  as a function of bond length for a uniform expansion or contraction of the trigonal crystal. Here we have taken the atom as the

Figure 33

Sketch of contributions  $\Delta E_{BS}$  and  $\Delta U$  to the total energy per bond  $\Delta E_T$ , as a function of bond length  $d$ .



ENERGY (arb. units)



zero of energy. This corresponds to taking the bond length  $d \rightarrow \infty$ . The cohesive energy of the crystal  $E_D$  is given by the minimum of  $E_T$  which occurs at the equilibrium bond length  $d_0$ . We can expand  $\Delta U$  as a function of  $d$  about  $d_0$  (Ref. 80):

$$\begin{aligned} \Delta U &= \Delta U_0 + \alpha(d-d_0) \\ &+ \beta(d-d_0)^2/2 + \text{higher terms.} \end{aligned} \quad (\text{A21})$$

Since  $\Delta E_{BS}$  can be calculated as a function of  $d$ , it is possible to obtain the constants  $\Delta U_0$ ,  $\alpha$ , and  $\beta$  if the binding energy and bulk modulus are known experimentally.

Since it is very difficult to calculate  $U$  for an arbitrary structural configuration, it is necessary to introduce a simplifying assumption at this point. Following Chadi<sup>80</sup> we assume

$$U = \sum_l \Delta U(d_l), \quad (\text{A22})$$

where  $l$  labels first-neighbor covalent bonds in the material and  $d_l$  is the corresponding bond length. This approach has been used previously to estimate the relaxed bond length at semiconductor surfaces<sup>80</sup> where only  $\alpha$  and  $\beta$  need be known. In the present context, however, we assume bonds are either normal ( $d=d_0$ ) or completely broken. Thus we need only find a way to estimate the equilibrium value  $\Delta U_0$ . Having done so, we calculate the total energy of any defect structure by integrating over the density of states to obtain  $E_{BS}$  and correcting by  $\pm \Delta U_0$ .

for each broken or extra bond in the structure. (From now on,  $E_T$  and  $E_{BS}$  will always be referred to the energy of an equal number of atoms in the normally coordinated bulk as the zero of energy.) We will describe the calculation of  $E_{BS}$  in detail shortly, but let us first estimate  $\Delta U_0$  using the case of selenium (Chapter II) as an example.

The most straightforward way to determine a value for  $\Delta U_0$  would be to imagine the dissociation of trigonal Se into separate atoms. The  $E_T$  is the experimental cohesive energy per atom  $E_b = 2.35$  eV,<sup>81</sup> and  $U = E_T - E_{BS} = -\Delta U_0$ . The determination of  $E_{BS}$  is complicated, however, by the fact that in our Hamiltonian, the atomic-orbital energy levels on each atom are shifted as a function of the coordination of that atom. The self-energy shifts for threefold, twofold, and onefold sites are  $\Delta_3 = -1.25$  eV,  $\Delta_2 = 0$  eV, and  $\Delta_1 = 1.25$  eV. If there were no self-energy shift on the free atom, we would find  $E_{BS} = 4.47$  eV. A simple extrapolation gives  $\Delta_0 = 2.50$  eV,  $E_{BS} = 19.47$  eV, and  $\Delta U_0 = 17.12$  eV. However, we have little confidence in such an extrapolation, and we prefer instead to find an experimental reference which does not involve free atoms.

This has been accomplished by considering the dissociation of the trigonal crystal into free  $Se_2$  molecules instead of free atoms. To be precise, we proceed conceptually in the following stages: (i) dissociate the crystal into free atoms; (ii) combine free atoms to form  $Se_2$  dimers; and (iii) adjust the bond length of the dimer to be  $d_0$ . The corresponding energy gains of

the system, per two atoms at  $0^\circ\text{K}$ , are (i)  $2E_b = 4.70$  eV (Ref. 81); (ii)  $-72.96$  kcal/mole  $= -3.18$  eV (Ref. 82); and (iii)  $0.46$  eV from vibrational data.<sup>83</sup> Thus for the entire process  $E_T = 1.98$  eV. The calculation of the corresponding  $E_{BS}$  is straightforward. Applying the Hamiltonian of Table I to the dimer, we calculate  $E_{BS} = 17.83$  eV and, therefore,  $U = E_T - E_{BS} = -15.85$  eV. Since this corresponds to breaking one bond, we have finally  $\Delta U_0 = 15.85$  eV. It is this value we will use for all subsequent calculations of defect total energies.

Our goal now is to determine  $E_{BS}$  and  $U$  for the onefold defect, threefold defect, and IVAP of Fig. 4. Because two onefold (threefold) defects are required to break (form) a single bond, we have  $U = \pm \Delta U_0 / 2$  and  $0$  for these three cases, respectively. We turn now to a detailed discussion of the evaluation of  $E_{BS}$ .

Recall that  $E_{BS}$  is now defined with respect to an equal number of atoms in the normally coordinated bulk, as a zero of energy. Thus if  $N_n(\epsilon)$  is the local density of states on the  $n^{\text{th}}$  site and  $N^0(\epsilon)$  is the local density of states at any site in the bulk, we can write

$$E_{BS} = \sum_n \int_{-\infty}^{\epsilon_F} \epsilon \delta N_n(\epsilon) d\epsilon, \quad (\text{A23})$$

where  $\epsilon_F$  is the Fermi level and  $\delta N_n(\epsilon) = N_n(\epsilon) - N^0(\epsilon)$ . The energy integral is carried out numerically, and the sum over sites is typically well converged (within  $-0.01$  eV) with the inclusion

of the first 10 sites extending down each of the semi-infinite chains attached to the defect.

There is one fine point which arises in the evaluation of the energy integral. In the Green's-function formalism, the density of states is

$$N_n(\varepsilon) = - (1/\pi) \sum_{\alpha} \frac{|\langle n|\alpha\rangle|^2}{\varepsilon - \varepsilon_{\alpha} + i\delta}, \quad (\text{A24})$$

where  $\alpha$  are eigenstates of the system and  $\delta$  is a small imaginary part which must be included in the calculation. Thus each eigenstate contributes to the total density of states as

$$(1/\pi) \frac{\delta}{(\varepsilon - \varepsilon_{\alpha})^2 + \delta^2}$$

This is a Lorentzian which goes to a delta function in the limit  $\delta \rightarrow 0$ . For calculational purposes,  $\delta$  is chosen to be comparable to the energy mesh spacing and is typically 0.04 eV. The fact that a nonzero  $\delta$  must be employed in the calculation has two unfortunate consequences. Firstly, the density of states does not quite drop to zero in the gap regions (see, for example, Figs. 3-5), so that it can be hard to decide where to put the Fermi level in the gap. Secondly, and more seriously, the energy integral over the tail of the Lorentzian goes as

$$\int_{-\infty}^{-x} (\varepsilon/\pi) \frac{\delta}{(\varepsilon - \varepsilon_{\alpha})^2 + \delta^2} d\varepsilon \approx (\delta/\pi) \int_{-\infty}^{-x} \varepsilon^{-1} d\varepsilon,$$

and is therefore logarithmically divergent.

We have corrected these problems by deconvoluting the Lorentzian out of the density of states, and subsequently broadening by a Gaussian of the same width. Using standard fast-Fourier-transform routines, we numerically transform the density of states, divide and multiply by the transforms of the Lorentzian and Gaussian, respectively, and then reverse transform back to energy space. Because the Gaussian falls off much faster in the tail regions, the divergences disappear and the density of states in the gap typically falls to a value several orders of magnitude lower than it did previously. This appears to be a simple yet general technique for obtaining an accurate energy integral.

The above program has been carried out to calculate the total energies of the onefold, threefold, and IVAP defects of Fig. 4. The results are shown in Table IV. The last three columns represent total energy estimates of increasing sophistication.  $E_{\text{SBO}}$  refers to "simple-bond-orbital" estimates<sup>14</sup> which are based on discrete bond-orbital electron levels and which neglect inter-ion repulsion entirely. The quantity  $E'_{\text{BS}} = E_{\text{BS}} - Z\Delta_{1 \text{ or } 3}$  is essentially our calculated  $E_{\text{T}}$  except that only that portion of  $U$  necessary to cancel the strong effects of the self-energy shifts in  $E_{\text{BS}}$  is included. Thus  $E_{\text{T}}$  effectively includes the ion-ion repulsion, while  $E'_{\text{BS}}$  does not. Note the striking reversal in the relative onefold and threefold energies in the last three columns. The physical

Table IV

Tight-binding total energy calculation for defects. First two columns are contributions to  $E_{BS}$  or  $E_T$ ; last three are total energy estimates of increasing sophistication. See text.

	$E_{BS}$ (eV)	U (eV)	$E_{SBO}$ (eV)	$E_{BS}'$ (eV)	$E_T$ (eV)
1-fold	9.09	-7.92	3 - 4	1.59	1.17
3-fold	-6.36	7.92	1	1.14	1.56
IVAP	1.22	0.0	?	1.22	1.22

significance of these results is discussed in Chapter II of the text.



## APPENDIX B SELF-CONSISTENT PSEUDOPOTENTIAL METHOD

In this appendix, we describe in detail some features of the self-consistent pseudopotential (SCPSP) method used in Secs. II(c) and (d). The plan is as follows. First, we will discuss the choice of local-density functional, paying special attention to the tail region and the spin-multiplet splitting in free atoms. Next, we will review the relative merits of local and nonlocal pseudopotentials. Then, we describe certain features of the crystalline calculations, including the implementation of the Lowdin perturbation theory and other convergence aids. Finally, we discuss the calculation of total energies and demonstrate that it is possible to obtain accurate equilibrium geometries and cohesive energies in crystalline Si, Ge, and Se.

We begin with local-density theory. In a series of seminal papers<sup>84-86</sup>, Hohenberg, Kohn and Sham considered a system of interacting electrons in an external potential, and proved that the many-body ground state wave function is a (universal) functional of the electronic charge density  $\rho$  alone. It follows that the ground state energy  $E$  of the system is also a functional of  $\rho$ . They then defined the exchange-correlation energy functional  $E_{xc}[\rho]$  by

$$E[\rho] \equiv \int [V_{\text{ext}}(r) + \epsilon_H(r)] dr + T[\rho] + E_{xc}[\rho] \quad (\text{B1})$$

where  $V_{\text{ext}}$  is the external potential,  $T[\rho]$  is the kinetic energy of a non-interacting electron system, and  $\epsilon_H$  is the Hartree energy

$$\epsilon_H(r) = (1/2) \int dr' \rho(r') / |r-r'| \quad (\text{B2})$$

(Atomic units, au and Hartree, are used in this Appendix except where otherwise specified.)

The only assumption that enters local density theory is that  $E_{\text{xc}}[\rho]$  can be approximated as a local function of  $\rho$ :

$$E_{\text{xc}}[\rho] = \int dr \rho(r) \epsilon_{\text{xc}}(\rho(r)) \quad (\text{B3})$$

Then, since  $E[\rho]$  has its variational minimum at the true  $\rho$ , one can take the functional derivative  $\delta/\delta\rho$  of Eq. (B1) to show that

$$0 = \int \delta\rho(r) [V_{\text{ext}}(r) + \mu_H(r) + \mu_{\text{xc}}(r) + \delta T/\delta\rho(r)] dr \quad (\text{B4})$$

where  $\mu_H(r) = 2\epsilon_H(r)$  and

$$\mu_{\text{xc}}(r) = \delta[\rho\epsilon_{\text{xc}}(\rho)]/\delta\rho. \quad (\text{B5})$$

To obtain the  $\rho$  which solves Eq. (B4), one need only solve for a non-interacting system with potential  $V_{\text{ext}} + \mu_H + \mu_{\text{xc}}$ ; this is now correctly done using a one-electron solution of the Schroedinger equation:

$$[-\nabla^2/2 + V_{\text{ext}}(r) + \mu_H(r) + \mu_{\text{XC}}(r)] \psi_i(r) = \epsilon_i \psi_i(r) \quad (\text{B6})$$

This must be solved self-consistently using

$$\rho(r) = \sum_i |\psi_i(r)|^2 \quad (\text{B7})$$

Finally, the ground state energy is given by

$$E = \sum_i \epsilon_i + \int \rho(r) [-(1/2)\mu_H(r) + \epsilon_{\text{XC}}(r) - \mu_{\text{XC}}(r)] dr \quad (\text{B8})$$

In order that  $\epsilon_{\text{XC}}(\rho)$  have the correct value in the limit of slow density variations, it is natural to set  $\epsilon_{\text{XC}}(\rho)$  equal to the exchange-correlation energy density in a uniform electron gas of density  $\rho$ . The exchange contribution is defined as the portion that would be present in a non-interacting uniform electron gas. It is easily given from Hartree-Fock as

$$\epsilon_X(\rho) = -3 (3\rho/8\pi)^{1/3} \text{Ryd} \quad (\text{B9})$$

with the corresponding potential from Eq. (B5) being

$$\mu_X(\rho) = -4 (3\rho/8\pi)^{1/3} \text{Ryd} . \quad (\text{B10})$$

The correlation contribution,  $\epsilon_C \equiv \epsilon_{\text{XC}} - \epsilon_X$ , contains the physics of many-body interactions.

Some years earlier, Slater had proposed a somewhat ad-hoc mean-field calculational scheme (known as the "Slater  $X_\alpha$  method")<sup>87</sup> which led to equations identical in form to (B6) and (B7). The great success of this method was in fact the primary stimulus for subsequent advances in local density theory. His approach lacked the rigor of the Hohenberg-Kohn theorem and led to a somewhat different coefficient in front of  $\rho^{1/3}$  in Eq. (B10), namely

$$\mu_{X_\alpha}(\rho) = -6\alpha (3\rho/8\pi)^{1/3} \text{ Ryd} \quad (\text{B11})$$

with  $\alpha = 1$ . Since  $\alpha = 2/3$  corresponds to pure exchange in the Hohenberg-Kohn-Sham picture, Slater subsequently suggested using Eq. (B11) and

$$\varepsilon_{X_\alpha}(\rho) = -(9/2)\alpha (3\rho/8\pi)^{1/3} \text{ Ryd} \quad (\text{B12})$$

with  $\alpha$  a parameter to be chosen between  $2/3$  and  $1$ , the idea being that  $\alpha > 2/3$  effectively includes some correlation in an approximate way. He suggested fitting  $\alpha$  to Hartree-Fock total energy calculations on isolated atoms (a procedure which generally yields  $\alpha_{\text{HF}} = 0.70-0.72$ ). However, this is hard to justify on the basis of first principles, since Hartree-Fock does not include correlation at all.

The next task of local density theory, then, was to obtain an accurate correlation function  $\varepsilon_c(\rho)$  from calculations of the

ground state energy of the electron gas, in order to supplement Eq. (B9). Unfortunately, the energy of the electron gas has proven to be extremely difficult to calculate for typical semiconductor or metallic electronic densities ( $r_s \approx 1-2$ , where  $r_s$  is conventionally defined by  $\rho^{-1} \equiv 4\pi r_s^3/3$ ). The earliest suggested formula is due to Wigner<sup>88</sup>, who proposed a simple interpolation between low and high-density limits; as subsequently corrected by Pines<sup>89</sup>, this leads to the still-popular "Wigner interpolation formula":

$$\varepsilon_c(r_s) = -0.88/(r_s+7.8) \text{ Ryd.} \quad (\text{B13})$$

As more accurate many-body and Monte-Carlo calculations of the electron-gas energy have become available, a large number of new formulas for the correlation energy have been spawned (for an extensive recent review, see Ref. 90). Many of these are generalized for spin-polarized ( $\rho^\uparrow \neq \rho^\downarrow$ ) systems as well. A common parametrization, due to von Barth and Hedin,<sup>91</sup> is

$$\varepsilon_{xc}(r_s, \xi) = \varepsilon_{xc}(r_s, 0) + [\varepsilon_{xc}(r_s, 1) - \varepsilon_{xc}(r_s, 0)]g(\xi)$$

where

$$\xi = (\rho^\uparrow - \rho^\downarrow) / (\rho^\uparrow + \rho^\downarrow)$$

$$g(\xi) = [(1+\xi)^{4/3} + (1-\xi)^{4/3} - 2] / [2^{4/3} - 2]$$

$$\epsilon_{xc}(r_s, 0) = \epsilon_x(r_s, 0) - c_0 f(r_s/r_0)$$

$$\epsilon_{xc}(r_s, 1) = 2^{4/3} \epsilon_x(r_s, 0) - c_1 f(r_s/r_1)$$

$$f(z) = (1+z^3) \ln(1+1/z) + z/2 - z^2 - 1/3 \quad (B14)$$

[Eq. (B5) is applied to obtain  $\mu_{xc}^{\uparrow}$  and  $\mu_{xc}^{\downarrow}$ .] The commonly-used formulas of von-Barth-Hedin,<sup>91</sup> Hedin-Lundquist,<sup>92</sup> Gunnarsson-Lundquist,<sup>93</sup> and Janak-Moruzzi-Williams<sup>94</sup> are of this form, each with different parameters  $c_0$ ,  $c_1$ ,  $r_0$ , and  $r_1$ .

The Wigner and subsequent formulas for  $\epsilon_{xc}$  may be interpreted as given by Eq. (B12) with a density-dependent  $\alpha$ . For a typical valence density in Se,  $r_s=2$ , one finds  $\alpha$  in the range 0.80-0.85 for all of them. Since it is pretty clear that the errors involved in making a local density approximation in the first place are of this magnitude or greater, it appears not to matter much which one is used. For example, the best calculations to date on the ground state properties and structural excitations of crystalline silicon<sup>44</sup> were done using Wigner exchange-correlation. In fact, leaving correlation out altogether ( $\alpha=2/3$ ) appears to give quite adequate results for binding energies and bond lengths in a number of applications [Refs. 41-42 and Sec. II(d)]. While  $\alpha=1$  often gives better estimates of excitation energies (semiconductor band gaps are

consistently too small for smaller  $\alpha$ ), it is clearly incorrect for ground state total energy estimates.

In order to calculate the cohesive energy of a solid (the energy to dissociate the crystal into free atoms), it is necessary to calculate the reference energy of an isolated atom. This is done in the spherical approximation, i.e. using a Herman-Skillman-type computer program.<sup>36</sup> However, two new complications arise. The first is that the ground state spin-multiplet is spin-polarized for open-shell atoms. For example, in Se the ground state is a  $^3P$  configuration, with 3 out of 4 valence p electrons having spin up. Von Barth<sup>95</sup> has shown how to calculate the spin-multiplet splittings correctly. In general this is a complicated problem, but it turns out that the energy of the most spin-polarized state (the ground state, by Hund's rule) is always easily calculated via a normal local-spin-density calculation. One simply continues to do a spherically-averaged calculation with unequal occupations for spin up and spin down electrons, using, e.g., Eq. (B14) to calculate  $\epsilon_{XC}$  and  $\mu_{XC}$ . We have calculated this spin-polarization energy correction using Gunnarsson-Lundquist exchange-correlation, and obtain, for example, 0.76 eV/atom in Se.

A second complication arises in the fact that the local density mean field potential has the wrong form in the tail region of the atom, decaying exponentially instead of having the correct  $1/r$  algebraic decay. Slater suggested simply

altering the potential in this region to the correct form, but it appears impossible to define a corresponding  $\epsilon_{xc}$  which preserves the variational relationship between  $\epsilon_{xc}$  and  $\mu_{xc}$  expressed in Eq. (B5). Moreover, even in the hydrogen atom, the "worst-case" system for a mean field theory, Gunnarsson-Lundquist exchange-correlation gives a binding energy of 0.984 Ryd, in error by only  $\sim 0.2\text{eV}$ . Errors in heavier atoms are probably not significantly greater. At this point, we prefer to live with the wrong form of the tail, taking the point of view that it is an intrinsic and unavoidable consequence of the local density approximation.

Once one has chosen a local-density exchange-correlation function, one can in principle do a calculation on a real crystalline solid by simply expanding in an appropriate basis set of Bloch functions, e.g. plane waves. However, to make the calculation tractable, one needs to appeal to methods which focus on valence electrons only. We have chosen to use a pseudopotential approach.

For self-consistent calculations, one needs to choose a pseudopotential to represent the ion core (charge  $Z=+6$  in Se). One general approach is to use an empirical pseudopotential, i.e. one which has been parametrized to fit experiment, usually optical properties. A well-known example is the Appelbaum-Hamann pseudopotential for silicon.<sup>96</sup> However, pseudopotentials fit to optical properties may give good excitation energies, but are not necessarily optimized to give



good ground state total energies. In fact, experience has shown that in general they do not.<sup>39,40</sup> Since this will be one of our primary applications, we prefer to adopt an alternate approach in which the pseudopotential is obtained from first principles in some way, usually by extracting information from an all-electron calculation on the free atom.

To see how this works, consider the case of the Se atom. The ionic pseudopotential, which represents the  $Z=+6$  ion, should have the property that, when self-consistently screened with six valence electrons, its 1s and 2p eigenvalues match the 4s and 4p eigenvalues of the all-electron atom. Furthermore, the corresponding wave functions should be the same outside the core region. Such a pseudopotential was proposed by Starkloff and Joannopoulos.<sup>45</sup> They adopted the analytic form

$$V_{\text{ion}}(r) = - f(r; \lambda, r_c) Z/r \quad (\text{B15})$$

where

$$f(r; \lambda, r_c) = [1 - e^{-\lambda r}] / [1 + e^{-\lambda(r - r_c)}] \quad (\text{B16})$$

is a steplike pseudizing function with core radius  $r_c$  and a reciprocal length  $\lambda$  characterizing the smoothness of the step. In principle,  $Z/r$  in Eq. (B15) should be replaced by the "base potential" obtained from the all-electron calculation by subtracting the screening due to valence electrons from the total

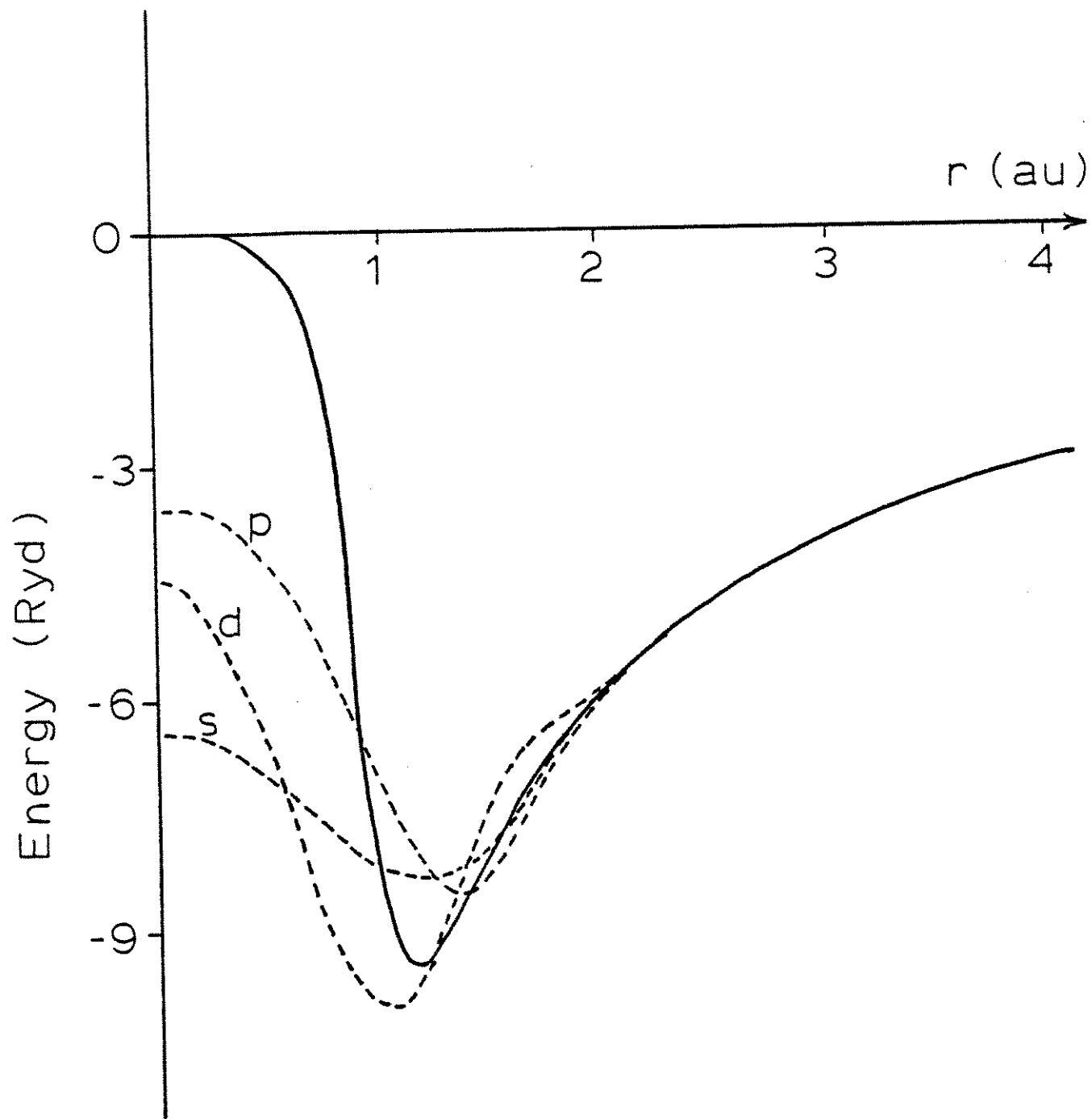
mean field potential. However, since this potential is already rapidly approaching  $Z/r$  at  $r_c$ , Eq. (B15) is generally perfectly adequate.

The idea now is to choose the two free parameters,  $\lambda$  and  $r_c$ , to exactly fit the two eigenvalues,  $\epsilon_s$  and  $\epsilon_p$ . This can be done for Se; with Wigner exchange-correlation, we obtain  $\lambda = 10.196 \text{ au}^{-1}$  and  $r_c = 0.92278 \text{ au}$ . The corresponding ionic pseudopotential is shown in Fig. 34. The self-consistently screened pseudopotential has been constructed so that it converges rapidly, outside the core, to the all-electron mean field potential; this, together with the equality of the eigenvalues, also insures that the pseudo wave functions converge rapidly to the all-electron wave functions outside the core. (In principle, this equality is only insured up to a multiplicative constant, but we find in practice that this constant is generally within ~1% of unity. To a good approximation, then, these are "norm-conserving"<sup>97</sup> pseudopotentials.)

The Starkloff-Joannopoulos pseudopotential has been used with success in a number of bandstructure<sup>35,99</sup> and total energy [Refs. 40, 98, and Sec. II(b)] calculations. However, it has certain drawbacks. Firstly, it is not always true that  $\lambda$  and  $r_c$  can be adjusted to fit  $\epsilon_s$  and  $\epsilon_p$  exactly. In Si and Ge, for example, the best fit is given with  $\lambda = \infty$  and  $r_c$  chosen for a least-squares fit of  $\epsilon_s$  and  $\epsilon_p$ . The method is therefore of rather limited utility in these cases; for example, we find

Figure 34

Ionic pseudopotentials for Se. Solid line: local potential of Ref. 45 with  $\lambda = 10.196 \text{ au}^{-1}$  and  $r_c = 0.92278 \text{ au}$ . Dashed lines: non-local potential of Ref. 97 with  $r_c(s) = 1.18 \text{ au}$ ,  $r_c(p) = 1.24 \text{ au}$ ,  $r_c(d) = 1.18 \text{ au}$ .



that trends in bond lengths and cohesive energies in going from crystalline Si to Ge are not even qualitatively correct. Secondly, even in Se we require a fairly sharp cut-off, which contributes high Fourier components to the pseudopotential. This in turn requires that we use a larger basis set (higher Fourier components) to expand the wave functions. Since computer time goes up roughly as the third power of basis size, or the ninth power of the cutoff wavevector, the abruptness of the cutoff may sharply increase the expense necessary to obtain a well-converged result.

These problems can be overcome by using a "non-local" pseudopotential, i.e. one in which we specify a different  $V_l(r)$  for each angular momentum component  $l$ . Among the non-local pseudopotential schemes suggested recently,<sup>97,100</sup> we prefer the method of Hamann, Schluter and Chiang<sup>97</sup> because it appears to give much smoother potentials. These authors start with the all-electron valence  $s$ ,  $p$ , and  $d$  wave functions and pseudize them so that they are nodeless, smooth, and identical to the all-electron wave functions outside the core. Then, using the known all-electron eigenvalue  $\epsilon_l$ , they invert the Schroedinger equation to obtain  $V_l(r)$  for each  $l$ . The resulting non-local pseudopotential (again using Wigner exchange-correlation) is shown along side the above-mentioned local potential in Fig. 34. It is clearly much smoother.

The implementation of non-local pseudopotentials in a crystal calculation is fairly straightforward.<sup>101</sup> It substan-

tially increases the time necessary to set up the Hamiltonian matrix, a procedure which usually dominates the computer time for these calculations. Luckily, some time can be saved by storing the non-local part of the Hamiltonian for each k-point, since only the local part of the potential (the screening) changes from iteration to iteration in the self-consistency process.

We turn now to a discussion of certain features of the self-consistent crystal calculations. These fall into two categories: (i) ways to improve the accuracy and efficiency of the Lowdin perturbation method, and (ii) tricks for minimizing the number of iterations necessary to obtain convergence.

We begin with the former. The Lowdin perturbation scheme<sup>37</sup> is a way of reducing the size of the Hamiltonian which must be diagonalized by including some higher Fourier basis vectors in first order perturbation theory. Let Roman letters label plane waves in set A with energy  $0 < \epsilon_n < E_A$ , and Greek letters label plane waves in set B with  $E_A < \epsilon_\alpha < E_B$ . Then we construct a renormalized Hamiltonian for set A which includes set B in perturbation theory:

$$U_{mn} = H_{mn} + \sum_{\alpha} \frac{H_{m\alpha} H_{\alpha n}}{\epsilon_n - \epsilon_{\alpha}} \quad (\text{B17})$$

This is sometimes called "folding in" of the higher-energy components. The new Hamiltonian  $U_{mn}$  has more accurate eigenvalues than  $H_{mn}$ .

Once  $U_{mn}$  has been diagonalized, the eigenfunctions must be constructed in order to obtain the charge densities for the next self-consistent iteration. On the one hand, it is possible to simply use the eigenstates of  $U_{mn}$  (that is, the wave functions would have no Fourier components in set B). On the other hand, it is better to use the perturbed wave functions. If  $\psi^0 = \sum_n c_n \phi_n$  is the eigenvector,  $U\psi^0 = \epsilon\psi^0$ , then the perturbed  $\psi$  is

$$\psi = \psi^0 + \sum_{\alpha} \left[ \sum_n \frac{U_{\alpha n}}{\epsilon - \epsilon_{\alpha}} c_n \right] \phi_{\alpha} \quad (\text{B18})$$

While the latter procedure (sometimes called "folding out" the wave function) is indeed preferable, it substantially increases the computer time required to compute and Fourier transform the charge densities. In practice, we find that if  $E_A$  and  $E_B$  are chosen large enough to be close to convergence (i.e. the cohesive energy is within  $\sim 1$  eV/atom of the converged value), the folding-out procedure provides very little improvement in the total energy ( $\sim 0.05$  eV/atom), although the charge densities are indeed substantially better. On the other hand, if  $E_A$  is low enough, the folding-out does substantially improve the total energy.

Before leaving the topic of Lowdin perturbation theory, we wish to make a final comment. In the denominator of Eq. (B17), it is necessary to replace the energy  $\epsilon$  by some representative or average value  $\bar{\epsilon}$ . For bandstructure or optical calculations,

it is often customary to set  $\bar{\epsilon}$  at or near the Fermi level, to give a good description of the gap. However, for ground state total energy calculations it is important to set  $\bar{\epsilon}$  at the center of gravity of the filled valence levels (well below  $\epsilon_F$ ). Otherwise, the energy denominator is systematically too small, so that the resulting eigenvalues and total energies are systematically too deep. We find that the correct choice of  $\bar{\epsilon}$  substantially improves the effectiveness of the Lowdin method in reducing the cost of total energy calculations.

We now comment upon two tricks for minimizing the number of iterations necessary to obtain convergence. This is important for two reasons: (a) when calculating total energies, the various contributions to  $E_{TOT}$  in Eq. (B8) are not well converged until a very high degree of convergence is reached, well after the eigenvalues are converged; and (b) for large unit cells with low symmetry, charge sloshing modes give rise to instabilities which must be carefully damped to obtain convergence at all.

The first of these problems is greatly reduced by making certain that  $E_{TOT}$  is calculated at the end of each iteration according to the variationally correct version of Eq. (B8). For example, one should replace  $(-1/2)\mu_H$  by  $\epsilon_H$  from the present iteration minus  $\mu_H$  from the previous one. This insures the variational property of  $E_{TOT}$ , i.e. that deviations in  $\rho$  only give rise to second-order deviations in  $E_{TOT}$ . Then  $E_{TOT}$  will be converged well before its individual contributions.



The second problem can be understood as follows. Let  $V^{(n)}$  be the mean field potential obtained on the  $n^{\text{th}}$  iteration,  $\rho^{(n)}$  the charge density obtained by solving the Schrodinger equation for  $V^{(n)}$ , and  $U^{(n)}$  be the new potential constructed by screening with  $\rho^{(n)}$ . If  $V^{(n)}$  were the converged potential  $V^0$ , we would have  $U^{(n)} = V^{(n)} = V^0$ . Now suppose we add a perturbation

$$V^{(n)}(\mathbf{r}) = V^0(\mathbf{r}) + \delta V \cos(\vec{G} \cdot \vec{r}) \quad (\text{B19})$$

Then we expect, by the definition of the dielectric constant  $\epsilon(\mathbf{q})$ , that

$$U^{(n)}(\mathbf{r}) = V^0(\mathbf{r}) + [1 - \epsilon(\mathbf{G})] \delta V \cos(\vec{G} \cdot \vec{r}) \quad (\text{B20})$$

Here the perturbing potential  $\delta V \cos(\vec{G} \cdot \vec{r})$  is the sum of the external perturbation  $V^{(n)} - U^{(n)}$  and the induced perturbation  $U^{(n)} - V^0$ . Clearly for  $\epsilon > 2$ , the perturbation has an oscillatory divergent behavior. Since  $\epsilon(\mathbf{q} \rightarrow 0) \approx 10-20$  in many semiconductors, this is a severe problem for small reciprocal lattice vectors. We have solved the problem by employing a wavevector-dependent damping parameter  $\alpha(\mathbf{q})$  in the construction of the new potential:

$$V^{(n+1)}(\mathbf{G}) = \alpha(\mathbf{G}) U^{(n)}(\mathbf{G}) + [1 - \alpha(\mathbf{G})] V^{(n)}(\mathbf{G}) \quad (\text{B21})$$

$(n-1)^{\text{th}}$   
↑

$V^{(n+1)} = V^{(n)}$  no change

$\epsilon(\mathbf{q}) = \epsilon(\mathbf{G})$

$\delta V / \epsilon = \frac{\delta V}{\epsilon}$

or  $\alpha(\mathbf{G}) = \frac{1}{\epsilon(\mathbf{G})}$

$\alpha(\mathbf{G}) = \frac{1}{\epsilon(\mathbf{G})}$

$\epsilon = 1 - \frac{1}{\epsilon}$

From Eq. (B20), the best  $\alpha$  is  $\alpha(q) = \epsilon^{-1}(q)$ . We use a Fermi-Thomas dielectric function with a small- $q$  cutoff:

$$\epsilon(q) = 1 + k_{\text{FT}}^2 / (q^2 + q_0^2) \quad (\text{B22})$$

By treating  $k_{\text{FT}}$  and  $q_0$  as adjustable parameters, it is usually possible to find values which give fast convergence.

Finally, we turn to our method of evaluating the total energy for crystalline solids. The formalism for doing this in momentum-space was first worked out by Ihm, Zunger, and Cohen<sup>102</sup>. While the treatment provided by these authors gives the right answer, it appears to be somewhat lacking treatment of the divergent Coulomb sums at  $q=0$ . For example, they refer to a "formal" expansion of the potential or charge density in the vicinity of  $q=0$ ; the meaning of such an expansion is unclear since these functions have delta-functions at  $q=0$  and are zero for finite  $q \neq G$ .

To avoid these problems, we introduce a new derivation which makes use of a powerful and physically appealing trick. We write down the total energy for a crystal in a universe whose Coulomb interaction has the Yukawa form

$$f(r) = e^{-\lambda r} / r, \quad (\text{B23})$$

and then show that  $E_{\text{TOT}} \rightarrow \text{constant}$  as  $\lambda \rightarrow 0$ . The true  $E_{\text{TOT}}$  is then just this limiting value.

To see how this works, define the pseudo core charge density to be

$$\rho_{ps}(r) = - (1/4\pi) \nabla^2 V_{ion}^{ps}(r) \quad (B24)$$

The fictitious potentials due to electrons and ions respectively are

$$\mu_H(r) = \int d^3r' \rho(r') f(r-r') \quad (B25)$$

$$V_{ps}(r) = \int d^3r' \rho_{ps}(r') f(r-r') \quad (B26)$$

We want to calculate the total energy per atom to assemble electrons and frozen cores from infinity into a crystal. Using Eq. (B8) this is

$$E_{TOT} = (1/N) \sum_i \epsilon_i + \int_{\Omega} d^3r \rho(r) [-(1/2)\mu_H + \epsilon_{xc} - \mu_{xc}](r) \\ + (1/2) \int d^3r \rho_{ps}(r) [\sum_{R \neq 0} V_{ps}(r-R)] \quad (B27)$$

To convert to momentum space, note that Eq. (B25) for example becomes

$$\mu_H(G) = \rho(G) f(G) \quad (B28)$$

with  $f(q) = 4\pi/\Omega(q^2 + \lambda^2)$ . Note that  $\mu_H(G=0) = 4\pi Z/\Omega\lambda^2$  is well-defined, and is cancelled by an identical term in the ion potential. Thus we are required to calculate the eigenvalues  $\varepsilon_i$  using a mean field potential  $V$  chosen so that  $V(G=0) = \mu_{xc}(G=0)$ . The total energy per atom becomes

$$\begin{aligned}
 E_{TOT} = & (1/N) \sum_i \varepsilon_i - (\Omega/2) \sum_{G \neq 0} \rho(G) \mu_H(G) \\
 & + \Omega \sum_G \rho(G) [\varepsilon_{xc} - \mu_{xc}](G) \\
 & + (\Omega/2) \sum_{G \neq 0} \rho_{ps}(G) V_{ps}(G) \\
 & - (1/2) \int d^3r \rho_{ps}(r) V_{ps}(r) \tag{B29}
 \end{aligned}$$

We have deleted  $G=0$  in the second and fourth terms since they give rise to cancelling contributions of  $\pm 2\pi Z^2/\Omega\lambda^2$ . Now there are no divergences for  $\lambda \rightarrow 0$  in Eq. (B29), so this is a rigorously correct expression for the total energy.

However, the utility of (B29) is somewhat limited because  $\rho_{ps}(G)$  and  $V_{ps}(G)$  are often poorly converged until very high  $G$ . We can solve this problem by making a transformation to cores consisting of Gaussian-shaped charge packets  $\rho_g$  of half-width  $\sigma$  and containing charge  $+Z$ . We choose  $\sigma$  small enough so that the charge clouds  $\rho_{ps}$  and  $\rho_g$  do not overlap for first neighbors, but large enough so that  $\rho_g(q)$  converges quickly for large  $q$ . Once again we resort to our Yukawa potential; the fictitious poten-

tial  $V_g(r)$  due to  $\rho_g(r)$  is defined in analogy to Eq. (B26).

Then

$$\begin{aligned}
 E_{\text{TOT}} = & (1/N) \sum_i \varepsilon_i - (\Omega/2) \sum_{G \neq 0} \rho(G) \mu_H(G) \\
 & + \Omega \sum_G \rho(G) [\varepsilon_{xc} - \mu_{xc}](G) \\
 & + (\Omega/2) \sum_{G \neq 0} \rho_g(G) V_g(G) \\
 & - (1/2) \int d^3r \rho_g(r) V_g(r) \\
 & + \alpha_g Z
 \end{aligned}
 \tag{B30}$$

To evaluate  $\alpha_g Z$ , note that

$$\begin{aligned}
 \alpha_g Z = & (1/2) \int d^3r \int d^3r' f(r-r') \\
 & \times \sum_{R \neq 0} [\rho_{ps}(r) \rho_{ps}(r'-R) - \rho_g(r) \rho_g(r'-R)] \\
 = & \int d^3r' \Delta V(r') [\sum_{R \neq 0} \bar{\rho}(r'-R)]
 \end{aligned}
 \tag{B31}$$

where  $\Delta V \equiv V_{ps} - V_g$  and  $\bar{\rho} \equiv (\rho_{ps} + \rho_g)/2$ .

If Gauss's law were valid, then we would have  $\Delta V(r') = 0$  for  $r'$  outside the charge cloud  $\Delta \rho \equiv \rho_{ps} - \rho_g$ . Since we assumed non-overlapping charge clouds,  $\alpha_g$  would vanish. However, Gauss's law is not valid for a Yukawa potential. The potential

at a distance  $r'$  from the center of a sphere of radius  $r$  containing unit charge is

$$(1 + r^2 \lambda^2 / 6) e^{-\lambda r'} / r' \quad (\text{B32})$$

for

$$r \ll r' \quad ; \quad r \ll \lambda^{-1} . \quad (\text{B33})$$

Using (B32), one obtains

$$\Delta V(r') = \int d^3 r \Delta \rho(r) (r^2 \lambda^2 / 6 r') e^{-\lambda r'} \quad (\text{B34})$$

Thus

$$\begin{aligned} \alpha_g Z &= [(2\pi/3) \int d^3 r r^2 \Delta \rho(r) \\ &\quad \times \int d^3 r' h(r') [\sum_{R \neq 0} \bar{\rho}(r' - R)] \end{aligned} \quad (\text{B35})$$

where we have defined  $h(r') \equiv \lambda^2 e^{-\lambda r'} / 4\pi r'$ . Note that this envelope function  $h$  integrates to unity for all  $\lambda$ . Moreover, it extends smoothly over a larger and larger region of  $r'$ -space as  $\lambda \rightarrow 0$ . Thus the  $r'$  integral just gives  $Z/\Omega$ , the average charge per atomic volume, in this limit. Note that Eq. (B33) is valid for all but a vanishing portion of the integral as  $\lambda \rightarrow 0$ . Finally,  $\alpha_g$  is now well-defined as  $\lambda \rightarrow 0$ .

$$\begin{aligned}\alpha_g &= (2\pi/3\Omega) \int d^3r r^2 \Delta\rho(r) \\ &= \int d^3r [V_{ps}(r) - V_g(r)] / \Omega .\end{aligned}\tag{B36}$$

The second line was obtained by integrating by parts and using Gauss's theorem.

To make the connection with Ihm, Zunger and Cohen<sup>102</sup>, we simply take the limit of a small Gaussian charge packet,  $\sigma \rightarrow 0$ . The third and fourth terms of (B30) become the Ewald sum " $\chi_{\text{Ewald}}$ " and  $\alpha_g$  becomes " $\alpha_1$ ." Since one usually calculates  $\chi_{\text{Ewald}}$  by replacing the delta-function potentials by Gaussians anyway, we prefer to use Eqs. (B30) and (B36) directly to calculate the total energy. (Of course, certain integrals and Fourier transforms involving Gaussians are performed analytically.) These equations were in fact implemented in our computer program.

We conclude this Appendix by giving some sample results for total energy calculations on crystalline Si, Ge, and Se. Here, we use Wigner exchange-correlation, and the Hamann-Schluter-Chiang non-local pseudopotentials. As a consistency check on our method, we begin by reproducing the previous calculation of Yin and Cohen<sup>44</sup> on crystalline Si using a basis set cut-off of  $E_A = E_B = 11.5$  Ryd (i.e. no Lowdin perturbation set). Choosing their theoretical equilibrium bond length, and using the same set of 10 special k-points, we reproduce their calculation of the cohesive energy (we find

4.68 eV/atom compared to their 4.67 eV/atom). Clearly this validates the fact that both programs are working correctly.

Table V shows that the application of Lowdin perturbation theory, following the prescription given earlier in this Appendix, gives equally accurate results while allowing the cost of the calculation to be cut in half (compare cases 2-3). Moreover, it shows that the calculation has not yet converged at  $E_A = E_B = 11.5$  Ryd; we find that the solid gains about 0.2 eV in additional energy as one goes to full convergence (see cases 3-5).

We have also reproduced the calculation of Yin and Cohen on germanium<sup>103</sup>, which was done in exact analogy to their calculation of Si (using the same cutoffs etc.). Moreover, we have extended the calculations to Se, where one must minimize with respect to three degrees of freedom in the unit cell instead of one. Our total energy was minimized for a first neighbor bond length of 2.367 Å, a bond angle of 102.60°, and a second neighbor (interchain) bond length of 3.104 Å. These differ from experiment<sup>32,46</sup> by -1.0%, +0.2%, and -9.3% respectively. These can be compared with the Yin and Cohen errors for the bond lengths of Si and Ge (+0.5% and +0.1% respectively). Clearly the first neighbor bond length and bond angle are obtained with comparable accuracy. The error in the interchain distance is not surprising, since this represents an interaction having largely van der Waals character. In fact, the previous set of calculations [Sec. (II(d))] using local pseudopotentials gave



Table V

Tests for convergence of the cohesive energy with respect to basis size for Si. Calculations were done at 1 k-point, scaled to 10 k-points, and then corrected by a constant for zero-point motion and spin polarization. Wave functions were folded out whenever Lowdin perturbation sets were used. Computer time per iteration per k-point is shown in the last column.

Case	$E_A$ (Ryd)	$E_B$ (Ryd)	$E_{coh}$ (eV)	Time (sec)
1	7.5	7.5	3.95	21
2	7.5	11.5	4.70	30
3	11.5	11.5	4.68	64
4	11.5	15.0	4.83	85
5	11.5	25.0	4.89	160

the neighboring chains to be weakly unbound; in that case, we fixed the second neighbor distance at the experimental value. On the whole, we consider the theoretical value of the inter-chain distance to be quite acceptable.

As can be seen from Table VI, the errors in the calculated cohesive energies are much larger than the errors in the equilibrium geometry. Furthermore, they are systematically too large (when fully converged), and the deviations appear to increase as one goes to heavier elements or elements containing more valence electrons. For Se, some of the error may be due to excessive interchain binding, as would be indicated by the fact that the equilibrium second neighbor distance is too short. However, even for Germanium the error is almost half an eV per atom.

Of course, in almost all calculations of interest, one is interested in the relative energies of different structural configurations. As indicated by the accurate equilibrium structures predicted above, these are very accurately reproduced. Therefore, we suggest that the method as it now stands, with the time-saving techniques discussed in this Appendix, can be applied to the study of many interesting physical problems.

Table VI

Calculated cohesive energies for Si, Ge and Se compared to experiment. "Previous" calculations are from Refs. 44 and 103 (using  $E_A = E_B = 11.5$  Ryd). "Present" calculations were performed by going to  $E_A = 11.5$  Ryd,  $E_B = 25.0$  Ryd for Si and Ge;  $E_A = 8.91$  Ryd,  $E_B = 20.15$  Ryd for Se.

	Experiment	Previous theory	Present theory
Si	4.63 eV	4.67 eV	4.89 eV
Ge	3.85 eV	4.02 eV	4.31 eV
Se	2.25 eV	----	3.13 eV



#### ACKNOWLEDGMENTS

Firstly, I would like to recognize two organizations for their support. These are: the National Science Foundation, which awarded me a Graduate Fellowship and provided partial research funding via NSF Grant No. DMR76-80895; and the Xerox Webster Research Center, which hosted me for a month in 1979 as a Visiting Scientist to initiate the work on polyacetylene.

Among the many individuals to whom I am indebted, let me single out just a few. First, thanks to Thomas Starkloff, who initiated me into the nitty-gritty of pseudopotential calculations. Second, special thanks to Dr. Eugene Mele, who was my collaborator on the polyacetylene work and who arranged my visit to Xerox. And third, I am deeply indebted to Professor Marc Kastner. I have benefited greatly by the opportunity of interacting with his experimental group, and his support and criticism have been invaluable.

Most of all, I wish to express boundless gratitude to my thesis advisor, Professor John Joannopoulos, for his great enthusiasm, helpfulness, supportiveness, and sense of humor.



REFERENCES

- <sup>1</sup>N. F. Mott and E. A. Davis, Electronic Processes in Non-Crystalline Materials (Clarendon, Oxford, England, 1971).
- <sup>2</sup>H. Fritzsche, Amorphous and Liquid Semiconductors, edited by J. Tauc (Plenum, New York, 1973), p. 221.
- <sup>3</sup>J. M. Marshall and E. A. Owen, *Philos. Mag.* 31, 1341 (1975).
- <sup>4</sup>S. C. Agarwal, *Phys. Rev. B* 7, 685 (1973).
- <sup>5</sup>S. G. Bishop, U. Strom, and P. C. Taylor, *Phys. Rev. B* 15,
- <sup>6</sup>R. A. Street, *Adv. Phys.* 25, 397 (1976).
- <sup>7</sup>P. W. Anderson, *Phys. Rev. Lett.* 34, 953 (1975).
- <sup>8</sup>P. W. Anderson, *J. Phys. (Paris)* 10, suppl. C4 (1976).
- <sup>9</sup>D. C. Licciardello, D. L. Stein, and F. D. M. Haldane, *Philos. Mag.* (in press).
- <sup>10</sup>D. C. Licciardello, *Comments Solid State Phys.* 9, 217 (1980).
- <sup>11</sup>R. Kaplow, T. A. Rowe, and B. L. Averbach, *Phys. Rev. B* 3, 168 (1968).
- <sup>12</sup>N. F. Mott, E. A. Davis, and R. A. Street, *Philos. Mag.* 32, 961 (1975).
- <sup>13</sup>R. A. Street and N. F. Mott, *Phys. Lett.* 35, 1293 (1975).
- <sup>14</sup>M. Kastner, D. Adler, and H. Fritzsche, *Phys. Rev. Lett.* 37, 1504 (1976).
- <sup>15</sup>M. Kastner, *J. Non-Cryst. Solids* 31, 223 (1978).
- <sup>16</sup>M. Kastner and H. Fritzsche, *Philos. Mag. B* 37, 199 (1978).

- <sup>17</sup>S. G. Bishop, U. Strom, and P. C. Taylor, The Physics of Selenium and Tellurium, edited by E. Gerlach and P. Grosse (Springer, New York, 1979), p. 193.
- <sup>18</sup>D. Emin, *J. Non-Cryst. Solids* 35-36, 969 (1980), and references therein.
- <sup>19</sup>J. C. Phillips, A. C. Beevers, and S. E. B. Gould, *Phys. Rev.* B21, 5724 (1980)
- <sup>20</sup>E. N. Economou, K. L. Ngai and T. L. Reinecke, Linear and Non-Linear Electron Transport in Solids (Plenum, New York, 1976), p.595.
- <sup>21</sup>V. Halpern, *Philos. Mag.* 34, 331 (1976).
- <sup>22</sup>K. L. Ngai and P. C. Taylor, *Philos. Mag. B* 37, 175 (1978).
- <sup>23</sup>R. Street, Proceedings of the 14th International Conference on the Physics of Semiconductors (Edinburgh), ed. B. Wilson, (The Institute of Physics, 1978), p. 1291.
- <sup>24</sup>N. Mott, Proceedings of the International Topical Conference on the Physics of SiO<sub>2</sub> and its Interfaces, ed. S. T. Pantelides (Pergamon Press, N.Y., 1978), p.1.
- <sup>25</sup>G. S. Higashi and M. Kastner, *J. Phys. C* 12, L821 (1979).
- <sup>26</sup>J. Orenstein, and M. Kastner (unpublished).
- <sup>27</sup>M. Abkowitz and R. C. Enck, *J. Non-Cryst. Solids* 35-36, 831 (1980).
- <sup>28</sup>M. Abkowitz, *Bull. Am. Phys. Soc.* 25, 231 (1980).
- <sup>29</sup>J. D. Joannopoulos, The Physics of Selenium and Tellurium, edited by E. Gerlach and P. Grosse (Springer, New York, 1979), p. 2.



- <sup>30</sup>B. Moreth, Phys. Rev. Lett. 42, 264 (1979).
- <sup>31</sup>D. R. McCann and L. Cartz, J. Appl. Phys. 43, 4473 (1972).
- <sup>32</sup>W. Lingelbach, J. Stuke, G. Weiser, and J. Treusch, Phys. Rev. B 5, 243 (1972).
- <sup>33</sup>W. B. Pollard and J. D. Joannopoulos, Phys. Rev. B 19, 4217 (1979). 2278 (1977).
- <sup>34</sup>M. A. Abkowitz, The Physics of Selenium and Tellurium, edited by E. Gerlach and P. Grosse (Springer, New York, 1979), p. 210.
- <sup>35</sup>J. D. Joannopoulos, Th. Starkloff, and M. Kastner, Phys. Rev. Lett. 38, 660 (1977).
- <sup>36</sup>F. Herman and S. Skillman, Atomic Structure Calculations (Prentice-Hall, Englewood Cliffs, 1963).
- <sup>37</sup>P. Lowdin, J. Chem. Phys. 19, 1396 (1951).
- <sup>38</sup>D. J. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).
- <sup>39</sup>H. Wendel and R. Martin, Phys. Rev. Lett. 40, 950 (1978).
- <sup>40</sup>J. Harris and R. O. Jones, Phys. Rev. Lett. 41, 191 (1978).
- <sup>41</sup>J. Ihm and M. L. Cohen, Solid State Comm. 29, 711 (1979).
- <sup>42</sup>E. O. Kane, Phys. Rev. B 21, 4600 (1980).
- <sup>43</sup>A. Zunger, Phys. Rev. B 21, 4785 (1980).
- <sup>44</sup>M. T. Yin and M. L. Cohen, Phys. Rev. Lett. 45, 1004 (1980).
- <sup>45</sup>Th. Starkloff and J. D. Joannopoulos, Phys. Rev. B 16 5212 (1977).
- <sup>46</sup>D. R. McCann and L. Cartz, J. Appl. Phys. 43, 4473 (1972).
- <sup>47</sup>C. Kittel, Introduction to Solid State Physics, (John Wiley and Sons, New York 1976).

- <sup>48</sup>J. A. Gardner and M. Cutler, Proceedings of the 7th International Conference on Liquid and Amorphous Semiconductors, ed. W. Spear (University of Edinburgh, 1977), p. 838.
- <sup>49</sup>S. G. Bishop, U. Strom, and P. C. Taylor, *J. Non-Cryst. Solids* 32, 359 (1979).
- <sup>50</sup>D. W. Bullett, *Phys. Rev. B* 14, 1683 (1976).
- <sup>51</sup>R. A. Street and G. Lucovsky, *S. S. Comm.* 31, 289 (1979).
- <sup>52</sup>S. G. Bishop and N. J. Shevehik, *Phys. Rev. B* 12, 1567 (1975).
- <sup>53</sup>J. D. Joannopoulos and F. Yndurain, *Phys. Rev. B* 10, 5164 (1974).
- <sup>54</sup>R. Zallen and M. Slade, *Phys. Rev. B* 9, 1627 (1974).
- <sup>55</sup>A. C. Larson and D. T. Cromer, *Acta Cryst. B* 29, 1583 (1973).
- <sup>56</sup>W. B. Pollard and J. D. Joannopoulos, *Phys. Rev. B* 19, 4217 (1979).
- <sup>57</sup>D. Adler, *J. Non-Cryst. Solids* 35-36, 819 (1980).
- <sup>58</sup>C. K. Chiang, C. R. Fincher, Jr., Y. W. Park, A. J. Heeger, H. Shirikawa, E. J. Louis, S. C. Gau, and A. G. MacDiarmid, *Phys. Rev. Lett.* 39, 1098 (1977).
- <sup>59</sup>*Physics Today* 32 (9), 19 (1979).
- <sup>60</sup>M. J. Rice, *Phys. Lett.* 71, 4 (1979); 71, 152 (1979).
- <sup>61</sup>W. P. Su, J. R. Schrieffer, and A. J. Heeger, *Phys. Rev. Lett.* 42, 1698 (1979).
- <sup>62</sup>C. R. Fincher, Jr., M. Ozaki, M. Tanaka, D. Peebles, L. Lauchlan, A. J. Heeger, and A. G. MacDiarmid, *Phys. Rev.* B20, 1589 (1979).

- <sup>63</sup>T. Tani, P. M. Grant, W. D. Gill, G. B. Street, and T. C. Clarke, *Solid State Commun.* 33, 499 (1980).
- <sup>64</sup>C. B. Duke, A. Paton, W. R. Thomas, E. W. Plummer, A. J. Heeger, and A. G. MacDiarmid, *Chem. Phys. Lett.* 59, 146 (1978).
- <sup>65</sup>A. Karpfen and J. Petkov, *Solid State Commun.* 29, 251 (1979).
- <sup>66</sup>R. H. Baughman, S. L. Hau, G. P. Pez, and A. J. Signorelli, *J. Chem. Phys.* 68, 5405 (1978).
- <sup>67</sup>P. A. Lee, T. M. Rice, and P. W. Anderson, *Phys. Rev. Lett.* 13, 462 (1973); M. J. Rice and S. Strassler, *Solid State Commun.* 13, 125 (1973).
- <sup>68</sup>H. Shirakawa, T. Ito, and S. Ikeda, *Polym. J.* 4, 460 (1973); S. Lefrant, L. S. Lichtman, H. Temkin, D. B. Fithcen, D. C. Miller, G. E. Whitewell, and J. M. Burlitch, *Solid State Commun.* 29, 1919 (1979).
- <sup>69</sup>For example, F. Inagaki, M. Tasumi, and T. Miyazawa, *J. Raman Spectrosc.* 3, 335 (1975).
- <sup>70</sup>E. J. Mele and M. J. Rice, *Solid State Commun.* 34, 339 (1980).
- <sup>71</sup>J. Bart and C. H. MacGillvary, *Acta Crystallogr. Sect. B* 24, 1569 (1968).
- <sup>72</sup>J. D. Joannopoulos and M. L. Cohen, *Solid State Physics* (Academic, New York, 1971), Vol. 31, p. 71.
- <sup>73</sup>W. P. Su, *Solid State Commun.* 35, 899 (1980).

- <sup>74</sup>P. M. Grant and I. P. Batra, *Solid State Commun.* 29, 225 (1979).
- <sup>75</sup>M. J. Rice and S. Strassler, *Solid State Commun.* 13, 1389 (1973).
- <sup>76</sup>J. J. Ritsko, E. J. Mele, A. J. Heeger, A. G. MacDiarmid, and M. Ozaki, *Phys. Rev. Lett.* 44 1351 (1980).
- <sup>77</sup>J. D. Joannopoulos and F. Yndurain, *Phys. Rev. B* 10, 5164 (1974).
- <sup>78</sup>J. D. Joannopoulos, *Phys. Rev. B* 16, 2764 (1977).
- <sup>79</sup>J. C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974).
- <sup>80</sup>D. J. Chadi, *Phys. Rev. Lett.* 41, 1062 (1978).
- <sup>81</sup>D. D. Wagman, W. H. Evans, V. B. Parker, I. Halow, S. M. Bailey, and R. H. Schum, *Selected Values of Chemical Thermodynamic Properties*, U. S. National Bureau of Standards Technical Note 270-3 (U.S.G.P.O., Washington, DC., 1968), p. 56.
- <sup>82</sup>K. K. Yee and R. F. Barrow, *J. Chem. Soc. Faraday Trans. 2* 68, 1181 (1972).
- <sup>83</sup>G. Herzberg, *Spectra of Diatomic Molecules* (Van Nostrand, Princeton, 1950), pp. 101-102, 568.
- <sup>84</sup>P. Hohenberg and W. Kohn, *Phys. Rev.* 136, B864 (1964).
- <sup>85</sup>W. Kohn and L. J. Sham, *Phys. Rev.* 140, A1133 (1965).
- <sup>86</sup>L. J. Sham and W. Kohn, *Phys. Rev.* 145, 561 (1966).

- <sup>87</sup>J. C. Slater, Phys. Rev. 81, 385 (1951); 82, 538 (1951). See also J. C. Slater, The Self-Consistent Field for Molecules and Solids (McGraw Hill, New York, 1974).
- <sup>88</sup>E. P. Wigner, Phys. Rev. 46, 1002 (1934).
- <sup>89</sup>D. Pines, Elementary Excitations in Solids (W. A. Benjamin, New York, 1963), pp. 94, 164.
- <sup>90</sup>A. K. Rajagopal, Adv. Chem. Phys. 41, 59 (1980).
- <sup>91</sup>U. von Barth and L. Hedin, J. Phys. C 5, 1629 (1972).
- <sup>92</sup>L. Hedin and B. I. Lundquist, J. Phys. C 4, 2064 (1971).
- <sup>93</sup>O. Gunnarsson and B. I. Lundquist, Phys. Rev. B 13, 4247 (1976).
- <sup>94</sup>J. F. Janak, V. L. Moruzzi, and A. R. Williams, Phys. Rev. B 12, 1257 (1975).
- <sup>95</sup>U. von Barth, Phys. Rev. A 20, 1693 (1979).
- <sup>96</sup>J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 8, 1777 (1973).
- <sup>97</sup>D. R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. 43, 1494 (1979).
- <sup>98</sup>J. A. Verges and C. Tejedor, Phys. Rev. B 20, 4251 (1979).
- <sup>99</sup>Th. Starkloff and J. D. Joannopoulos, J. Chem. Phys. 68(2), 579 (1978).
- <sup>100</sup>A. Zunger, J. Vac. Sci. Technol. 16, 1337 (1979).
- <sup>101</sup>J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B 14, 556 (1976).
- <sup>102</sup>J. Ihm, A. Zunger and M. L. Cohen, J. Phys. C 12, 1979.
- <sup>103</sup>M. T. Yin and M. L. Cohen, unpublished.

