Band Structures and Pseudopotential Form Factors for Fourteen Semiconductors of the Diamond and Zinc-blende Structures

MARVIN L. COHEN and T. K. BERGSTRESSER
Department of Physics, University of California, Berkeley, California
(Received 30 August 1965)

Pseudopotential form factors and band structures are determined for 14 semiconductors of the diamond and zinc-blende structures: AlSb, CdTe, GaAs, GaP, GaSb, Ge, InAs, InP, InSb, Si, Sn, ZnS, ZnSe, and ZnTe. Experimental values of the splitting of energy levels in the crystal are used. The form factors appear to be accurate to 0.01 Ry and yield energy bands which agree with experiment to within ~0.01 Ry near the band gap and ~0.04 Ry over a range of 1 Ry. For some of these substances these are the first band structures to be calculated.

I. INTRODUCTION

Since the introduction of the pseudopotential method, and the subsequent proof of the Phillips Cancellation Theorem by Cohen and Heine, this method has become an important tool both for the investigation of electronic band structures of solids and for understanding the behavior of electrons in crystals. Pseudopotential theory has developed considerably in the last few years, and pseudopotentials have been determined for a large number of crystals. There are several methods for determining pseudopotentials or their form factors. These potentials can be derived from atomic wave functions, they can be computed from free-atom term values, or they can be obtained empirically from crystalline energy levels. This latter approach, the empirical pseudopotential method, has been successfully used both to obtain very accurate band structures for Si and Ge, and also to interpret optical experiments.

In this paper we extend the empirical pseudopotential method to include semiconductors with the zincblende structure and present band structures and pseudopotential form factors for AlSb, CdTe, GaAs, GaP, GaSb, Ge, InAs, InP, InSb, Si, Sn, ZnS, ZnSe, and ZnTe. The form factors appear to be accurate to 0.01 Ry, and they yield energy bands with splittings which agree with experiment to within ~0.01 Ry near the band gap and 0.04 Ry over a range of 1 Ry. The resulting band structures have been very useful for interpreting reflectivity and photoemission experiments.

In this paper we present little discussion or interpretation of the data. The experimental splittings which we have used are given in Table I. Some of the splittings are consistent with other authors, while others are not. The latter result either from our reinterpretation of the data or the availability of more recent experimental data. The experimental splittings should not be taken as absolutely rigid. Some of these splittings will undoubtedly change as even more accurate experimental data become available.

The paper is organized in the following way. In Sec. II we discuss the pseudopotential Hamiltonian and the mechanics of its solution. We describe the determination of the pseudopotential form factors in Sec. III. In Sec. IV we present the results and a discussion of them and of our approximations.

II. THE PSEUDOPOTENTIAL HAMILTONIAN

The pseudopotential Hamiltonian for an electron in the crystal consists of a kinetic-energy term plus a weak potential which depends on position.

\[ H = -\frac{(\hbar/2m)^2}{\mathbf{V}^2} + V(r). \]  

(1)

The potential \( V \) can be expanded in reciprocal lattice vectors \( \mathbf{G} \) and can be expressed as the product of a structure factor \( S(\mathbf{G}) \) times a pseudopotential form factor \( V_\alpha \). It is convenient to break up the potential into a symmetric and an antisymmetric part.

\[ V(r) = \sum_{|\mathbf{G}| \leq G_0} (S^s(\mathbf{G})V_\alpha^s + S^A(\mathbf{G})V_\alpha^A)e^{-\alpha|x-r|}. \]

The cubic semiconductors of the diamond or zinc-blende type have the fcc structure with two atoms per unit cell. We take the origin of coordinates to be halfway between these two atoms, whose positions are denoted by \( r_1 \) and \( r_2 \), so that \( r_1 = a(1,1,1) \), and \( r_2 = -a \), where \( a \) is the length of the unit cube. In this case

\[ S^s(\mathbf{G}) = \cos \mathbf{G} \cdot \mathbf{r}, \quad S^A(\mathbf{G}) = \sin \mathbf{G} \cdot \mathbf{r}. \]
In terms of atomic potentials,
\[ V_\sigma^2 = \frac{1}{2} \left( V_1(G) + V_2(G) \right), \quad V_\sigma^4 = \frac{1}{2} \left( V_1(G) - V_2(G) \right), \]
\[ V_1(G) = \frac{2}{\Omega} \int V_1(r) e^{-i\mathbf{Q} \cdot \mathbf{G} \cdot \mathbf{r}} \, d\mathbf{r}, \]
and similarly for \( V_2 \), where \( V_1 \) and \( V_2 \) are the pseudo-potentials due to single atoms in the lattice and \( \Omega \) is the volume of the unit cell. For the diamond structure, \( V_\sigma^2 = V_1(G) = V_2(G) \), and \( V_\sigma^4 = 0 \).

The basis states used to form the Hamiltonian matrix consist of plane waves with wave vector \( \mathbf{G} + \mathbf{k} \), where \( \mathbf{G} \) is a reciprocal lattice vector and \( \mathbf{k} \) is a wave vector lying within the first Brillouin zone. All vectors \( \mathbf{G} \), such as...
that \((G+k)^2 \leq E_i\), form the basis, and the second-order contribution from all vectors \(G\), such that \(E_i < (G+k)^2 \leq E_2\), are added to the matrix according to L"owdin's method,\(^8\) as modified by Brust.\(^9\) Convergence to within 0.1 eV is obtained with \(E_1 = 7\), and \(E_2 = 21\).

Here, \(G^2\) is measured in units of \((2\pi/a)^2\).

We solve the Schrödinger equation by determining the roots of the secular equation derived from the Hamiltonian matrix. For the diamond structure the matrix involved is real and is \(\sim 20 \times 20\). The zinc-blende matrix is complex since the matrix elements of the antisymmetric potential are imaginary. The eigenvalues of this matrix are found by diagonalizing a matrix twice as large as the original complex one. The matrices for nonsymmetry points are generally about the same size as those for symmetry points; hence, it is just as easy to compute the energy levels at a general point in the zone as it is to compute them at a symmetry point. It is therefore possible to compute the joint density of states for these crystals. Wave functions are also easily obtainable, making it possible to compute oscillator strengths and hence the zero wave vector dielectric function.\(^8\)\(^9\)

III. DETERMINATION OF THE FORM FACTORS

The first five of the reciprocal lattice vectors have squared magnitudes of 0, 3, 4, 8, and 11. Only these are allowed to have a nonzero potential. The symmetric structure factor is zero for \(G^2 = 4\), and the antisymmetric structure factor is zero for \(G^2 = 0\) and \(G^2 = 8\). We keep \(V_0, V_8\) = 0, since it merely adds a constant to all energy levels. Hence, there are only three symmetric and three antisymmetric form factors to be determined. These form factors are taken in a local and static approximation to be independent of momentum and energy so that they produce average potentials for the whole Brillouin zone and the entire range of energy of the valence and conduction bands. Our task is to determine the experimental data and to determine the approximate form factors which give band structures consistent with these data.

Table I contains our estimate of the best values for the energy differences between various electronic states in the Brillouin zone. Reflectivity and photoemission measurements have been most useful in yielding experimental information about the band structures. Reflectivity spectra have been taken of a large number of substances, including many of the cubic semiconductors, and the edges and peaks may be interpreted on the basis of direct transitions to yield values for the energy level splittings. The joint interband density of states \(J\) possesses a characteristic structure at a critical point\(^14\)

in the Brillouin zone, and \(J\) is simply related\(^9\) to the imaginary part of the dielectric constant \(\epsilon_2\) which is related\(^10\) to the reflectivity \(R\). In fact, the structure in \(R\) and the structure in \(\epsilon_2\) usually are within 0.1 eV of each other, with some exceptions at higher energies. In particular, we find a large shift for peaks arising from states near the \(X\) point in the zone. This peak in \(R\) is 0.2 to 0.4 eV higher than the corresponding peak in \(\epsilon_2\). In cases where \(\epsilon_2\) is not available, its value for the \(X\) peak is estimated from the peak in \(R\). Another small adjustment must be made to extrapolate the data to low temperature. The interpretation would be much strengthened by the availability of more reflectivity data, especially taken at low temperatures, where there is more and clearer structure. Two new techniques have recently become available for investigation of band structures, electroreflectivity\(^16\) and piezoreflectivity.\(^17\) These experiments have yielded quite sharp structure in Ge and Si which agree very well with our calculations. It is expected that these techniques will have a profound influence on the determination of critical point energies.

Photoemission-yield experiments, while not yet as extensive as reflectivity, can in principle give more information than reflectivity. The extra degrees of freedom gained by varying the work function and by measurement of energy distributions allows one to fix the levels of the conduction band on an absolute energy scale as well as to determine energy differences between the conduction and valence bands. The photoelectron energy distributions may even show structure that the reflectivity or yield does not show. Photoemission experiments have provided independent confirmation of the same picture of the band structure that has been obtained from reflectivity studies.\(^10\) Transmission and other experiments give additional information. See, for example, Refs. 9–12 for further discussion and references to some of these experiments.

The model Hamiltonian, Eq. (1), has no spin-orbit term, and our computed band structures neglect the effects of spin-orbit coupling. Therefore, we remove the spin-orbit splittings from the experimental data before quoting energies. We take the value of a level, degenerate without this splitting, as the average of its component levels, weighted with their degeneracies.

The method attempts to obtain band structures which are consistent with experiment. Consider first a group IV semiconductor, for example Ge. We do not find that the seven experimental values can be exactly fit within the constraint of using three form factors; attempting to fit some values exactly results in decreasing agreement with other values, so that a choice

we use a symmetric potential that is the average of the Ge and the Sn potential. Similarly for GaP and ZnS, we use the average of the Si and Ge potentials. For InP we were able to use the Ge symmetric potential while AlSb can be fitted more accurately with the average of the Si and the Sn symmetric potentials. The lattice constant of InP is fairly close to that of Ge while that of AlSb lies halfway between Si and Sn.

IV. RESULTS

Table II lists the pseudopotential form factors obtained for these 14 cubic semiconductors. The symmetric form factors for the heteropolar compounds are obtained from those of the homopolar substances as explained in Sec. III. One might expect a slight change in the symmetric potential upon going, for example, from Ge to GaAs. Other symmetric potentials were tried for all of the heteropolar compounds, and no significant improvement of the fit to experiment was found. For all of these, \( V_{s}^{s} \) is large and negative. The form factor passes through zero in the region of \( G^{2} = 8 \), and is positive but small for \( G^{2} = 11 \). The over-all sign of the antisymmetric form factors has no effect and is taken to be positive. The antisymmetric potential is stronger for the more ionic crystals: It is zero for the homopolar substances, while in the II-VI compounds it is about as strong as the symmetric potential.

Table III lists the derivatives of some of the energy levels with respect to the form factors for the series Sn, InSb, CdTe. These derivatives are not much different within the three classes: homopolar, III-V and II-VI. The most sensitive level is the \( \Gamma_{s} \) level, followed by \( L_{s} \). However, these become less sensitive with increasing antisymmetric potential. The least sensitive levels are the \( L_{s} \) and \( X_{s} \) levels in the valence band. All the levels except these are sensitive to the antisymmetric form factors. Varying one antisymmetric form factor moves

### Table II. Pseudopotential form factors, in rydbergs, derived from the experimental energy band splittings.

<table>
<thead>
<tr>
<th>Substances</th>
<th>( V_{s}^{s} )</th>
<th>( V_{s}^{a} )</th>
<th>( V_{t}^{s} )</th>
<th>( V_{t}^{a} )</th>
<th>( V_{l}^{s} )</th>
<th>( V_{l}^{a} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>-0.21</td>
<td>+0.04</td>
<td>+0.08</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ge</td>
<td>-0.23</td>
<td>+0.01</td>
<td>+0.06</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sn</td>
<td>-0.20</td>
<td>+0.00</td>
<td>+0.04</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GaP</td>
<td>-0.22</td>
<td>+0.03</td>
<td>+0.07</td>
<td>+0.12</td>
<td>+0.07</td>
<td>+0.02</td>
</tr>
<tr>
<td>GaAs</td>
<td>-0.23</td>
<td>+0.01</td>
<td>+0.06</td>
<td>+0.07</td>
<td>+0.05</td>
<td>+0.01</td>
</tr>
<tr>
<td>AlSb</td>
<td>-0.21</td>
<td>+0.02</td>
<td>+0.06</td>
<td>+0.06</td>
<td>+0.04</td>
<td>+0.02</td>
</tr>
<tr>
<td>InP</td>
<td>-0.23</td>
<td>+0.01</td>
<td>+0.06</td>
<td>+0.07</td>
<td>+0.05</td>
<td>+0.01</td>
</tr>
<tr>
<td>GaSb</td>
<td>-0.22</td>
<td>0.00</td>
<td>+0.05</td>
<td>+0.06</td>
<td>+0.05</td>
<td>+0.01</td>
</tr>
<tr>
<td>InAs</td>
<td>-0.22</td>
<td>0.00</td>
<td>+0.05</td>
<td>+0.08</td>
<td>+0.05</td>
<td>+0.03</td>
</tr>
<tr>
<td>InSb</td>
<td>-0.22</td>
<td>+0.00</td>
<td>+0.04</td>
<td>+0.06</td>
<td>+0.05</td>
<td>+0.01</td>
</tr>
<tr>
<td>ZnS</td>
<td>-0.22</td>
<td>+0.03</td>
<td>+0.07</td>
<td>+0.24</td>
<td>+0.14</td>
<td>+0.04</td>
</tr>
<tr>
<td>ZnSe</td>
<td>-0.23</td>
<td>+0.01</td>
<td>+0.06</td>
<td>+0.18</td>
<td>+0.12</td>
<td>+0.03</td>
</tr>
<tr>
<td>ZnTe</td>
<td>-0.22</td>
<td>0.00</td>
<td>+0.05</td>
<td>+0.13</td>
<td>+0.10</td>
<td>+0.01</td>
</tr>
<tr>
<td>CdTe</td>
<td>-0.20</td>
<td>+0.00</td>
<td>+0.04</td>
<td>+0.15</td>
<td>+0.09</td>
<td>+0.04</td>
</tr>
</tbody>
</table>

must be made of the values that are to be fit more closely. We regard the levels near the band gap and the splittings of less than 1 eV to be the most important ones to fit. These are fit to 0.1 or 0.2 eV; they can be fit exactly by determining the form factors to thousandths of a rydberg in the potential and sacrificing the larger splittings somewhat. Inclusion of more recent data has caused the present determination of the form factors of Ge to differ slightly from an earlier result.8,9

Having found the potential for Ge, we next consider GaAs. We keep the same symmetric potential and add a small antisymmetric potential. This potential has roughly the form of a Coulomb potential, which would be given by a simple point-ion approximation. In order to get the potential for ZnSe, we keep the same symmetric potential and increase the antisymmetric form factors of GaAs by a factor of \( \sim 2.3 \) to within \( \sim 0.01 \) Ry. The splittings for GaAs and for ZnSe are fitted as well as those for Ge. This same process is done for the next row of the periodic chart: Sn, InSb, CdTe. The “skew” compounds are done next. For InAs, GaSb, and ZnTe,
each level in the same direction, except $X_3$. In this table, all the levels are measured with respect to the level $\Gamma_{2\nu}$. On an absolute scale the $X_1$ level is least sensitive to changes in the form factors.

The calculated band structures are given in Figs. 1 through 14. The symmetry assignments are according to Herring and to Parmenter. The bands are computed along the symmetry directions $\Lambda$, $\Delta$, $\Sigma$, and the line between $X$ and $U(K)$. The last two are done by going directly from (1,1,0) to $\Gamma$. Most of the structure in the optical data appears to arise from states near these symmetry lines.

Several trends are conspicuous in the band structures. In the conduction band the $\Gamma_{2\nu}$ level comes

---

18 C. Herring, J. Franklin Inst. 233, 525 (1942).
rapidly down in energy as one proceeds from the lighter to the heavier semiconductors. Also, the $L_1$ level comes down with respect to the $X_1$ level. The antisymmetric potential causes the $X_1$ level in a homopolar substance to be split into an $X_1$ and an $X_3$ level in heteropolar substances. This splitting appears to be a measure of $V^A$. Both the transverse and longitudinal masses of the $X_1$ level are considerably greater than those of the $X_3$ level. This is manifested in the weakness of the $X_5$-$X_3$ peak in reflectivity, seen for example in GaAs.\(^{21}\) As a result of an increasing $V^A$, the splittings between the conduction and valence bands become progressively larger, and this follows crudely the $\lambda^2$ law.\(^{22}\) However, $\lambda^2$ extrapolations should not be treated as data. The conduction bands tend to become flattened out, as do the valence bands.

---


\(^{22}\) F. Herman, J. Electronics I, 103 (1955).
In the valence band the splitting between the $X_1$ and $X_3$ levels varies with $V_A$. It is, for example, 4 eV and 9 eV in GaAs and ZnSe, respectively. This causes the opening up of a gap which extends throughout the entire zone and which is at its minimum at $X$. This gap may be observable by a photoemission experiment at high energies. The width of the upper part of the valence band, $X_1(X_3)$ to $\Gamma_{2V}(\Gamma_{15})$, becomes progressively less as $V_A$ becomes larger. For example, in GaAs this width is 25% less than in Ge and it is 55% less in ZnSe than in Ge. The lowest valence band becomes very narrow in II-VI compounds, about 1-eV wide. These values for the valence band are only approximate since they are not tied down to any experimental data. For this reason, the lower part of the valence band has not been shown in the figures. See Ref. 5 for the general form of the lower valence band. However, we find at the $L$ point in the diamond-structure substances that the lowest level is $L_4$, and $L_{2V}$ is next, whereas Ref. 5 has the reverse.

The fact that some experimental levels cannot be accurately fitted at this time by the empirical pseudopotential method leads one to suspect either the interpretation of the data or the method. An example of the first case is provided by the $\Gamma_{15}-\Gamma_{16}$ splitting in GaP, which had previously been identified with a peak in reflectivity at 3.7 eV. The method could not give such a low value for this splitting without destroying the fit of the other levels. Consequently, this value of 3.7 eV was ignored when the form factors were chosen. Later, the work of Williams on the reflectivity of the GaAs-GaP alloy system became available to the authors. This data shows that the $\Gamma_{15}-\Gamma_{16}$ splitting actually is 4.8 eV. The current identification of the $\Gamma_{16}-\Gamma_{15}$ splitting in all of the homopolar substances and III-V compounds seems well establised because photoemission data are available for some of them, in addition to the reflectivity data. In these substances the results of our calculations are satisfactory, although the calculated splittings tend to be too high in the III-V compounds. However, in the II-VI compounds the situation of the $\Gamma_{16}-\Gamma_{15}$ splitting is not yet clear and experiments determining the photoemissive yield would be most helpful.

The calculated $L_{2V}-L_3$ splitting frequently has a smaller value than the experimental estimate. Several factors may be operative here. This peak is at high energy, where the calculation is not expected to be good, and the calculated value may be too low. However, the experimental value may be taken to be too high as the displacement of the peak in $\varepsilon_2$ from the peak in $\varepsilon_1$ could be large. Also, this peak is never sharp. In GaSb the $L_3-L_1$ splitting is too big, as is the $\Gamma_{16}-L_4$ splitting, indicating that the $L_4$ level should be lower. Whether this represents a breakdown in the pseudopotential method or whether the data needs to be reinterpreted is not known at this time, but in any case, the form factors may need to be adjusted.

The $X_4-X_1$ or $X_7-X_1$ splitting seems to be consistently too small, as was true in the previous calculations for Ge and Si. Table III reveals that the conduction bands are rather "flexible" whereas the valence band is not. This is also visible in the band structures; the valence band retains the same general appearance throughout the list of substances considered. This indicates that the error is due primarily to the lower level and that the band structures may be corrected by shifting down the $X_{1V}$ level by a few tenths of an eV. A correction of this type would bring all the band structures much closer to experiment. However, the conduction bands may have suffered in the attempt at an over-all fit because of the inflexibility of the valence band. The next higher form factor in both the symmetric and the antisymmetric potential was used to see whether that addition would improve the fit to experiment, but it did not. Apparently, the most important correction to be added to the form factors is an energy dependence, which would act to lower $X_{1V}$ and raise

---

Table IV. Comparison of the empirical pseudopotential form factors of this paper with those derived from spectroscopic term values. (Refs. 7, 25, and 26, main text.) The former form factors are on top of the latter; both are in rydbergs.

<table>
<thead>
<tr>
<th></th>
<th>$V_{s}$</th>
<th>$V_{v}$</th>
<th>$V_{i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>−0.21</td>
<td>+0.04</td>
<td>+0.08</td>
</tr>
<tr>
<td></td>
<td>−0.17</td>
<td>+0.05</td>
<td>+0.08</td>
</tr>
<tr>
<td>Ge</td>
<td>−0.23</td>
<td>+0.01</td>
<td>+0.06</td>
</tr>
<tr>
<td></td>
<td>−0.18</td>
<td>+0.03</td>
<td>+0.07</td>
</tr>
<tr>
<td>Sn</td>
<td>−0.20</td>
<td>0.00</td>
<td>+0.04</td>
</tr>
<tr>
<td></td>
<td>+0.01</td>
<td>+0.04</td>
<td></td>
</tr>
</tbody>
</table>

Such an energy dependence is implicit in the repulsive contribution to the form factor.

The empirical pseudopotential form factors found in this paper may be compared to form factors calculated from spectroscopic term values. The basic model potential is obtained from the spectroscopic term values of the single, isolated, triply ionized ion. The spin-orbit coupling is removed by using the weighted mean of the levels within one term. Orthogonalization and correlation corrections are added, and the potential is screened by the valence electrons which are treated as a free-electron gas. Matrix elements of the potential are taken at the Fermi momentum and at the Fermi energy. The values of these matrix elements for Si, Ge, and Sn, evaluated at the appropriate wave vectors, are compared in Table IV with our form factors.

There are two principal reasons why the agreement between the two sets of form factors should be only qualitative. One is that the empirical method takes into account crystalline effects such as charge redistribution, whereas the model potential is obtained by screening the single ion with the electron gas. The difference between our form factors and Animalu’s would then exhibit those crystalline effects, except that our form factors have the constraint that $V_{i}=0$ for $G>11$, while Animalu’s form factors do not. Nonetheless, the two sets of form factors are quite similar. In each case the empirical pseudopotential method yields form factors which are more negative than the corresponding form factors from the model potential.

The semiempirical approach employed in this paper benefits from its close connection to physical reality. The form factors should be nearly equal to the actual components of the potential averaged over the states of the valence and conduction bands near the gap. The cancellation of the strong potential of the atomic cores which occurs via the core states has been enforced by the procedure; thus a weak potential is obtained, and convergence is rapid. The method is not limited to symmetry points, and a general point in the Brillouin zone is done as rapidly as a symmetry point; in the determination of the photoemission structure of Si, thousands of points in the zone were calculated. It is expected that these band structures will be useful because of their correspondence to existing data and because there has been no published band structures for some of the substances considered.

The method is easily extendible to other structures. Some of the II-VI compounds have another structure, the hexagonal wurtzite structure. In this case one may use the same form factors derived from the allotropic, cubic substance. The form factors do need to be interpolated since the reciprocal lattice vectors in the hexagonal lattice have different magnitudes. Work is in progress on the wurtzite compounds.

Acknowledgments

We would like to express our gratitude to Professor J. C. Phillips for helpful discussions on all phases of this work. We would also like to thank Dr. David Brust for providing us with a copy of his program and for helpful discussions.

---
