BAND STRUCTURE OF INDIUM ANTIMONIDE

EVAN O. KANE

General Electric Research Laboratory, Schenectady, New York

Abstract—The band structure of InSb is calculated using the $k \cdot p$ perturbation approach and assuming that the conduction and valence band extrema are at $k = 0$. The small band gap requires an accurate treatment of conduction and valence band interactions while higher bands are treated by perturbation theory. A highly nonparabolic conduction band is found. The valence band is quite similar to germanium. Energy terms linear in k which cannot exist in germanium are estimated and found to be small, though possibly of importance at liquid-helium temperature. An absolute calculation of the fundamental optical absorption is made using the cyclotron resonance mass for n-type InSb. The agreement with experimental data for the fundamental absorption and its dependence on n-type impurity concentration is quite good. This evidence supports the assumptions made concerning the band structure.

1. INTRODUCTION AND CONCLUSIONS

THERE is now a considerable amount of experimental and theoretical information available concerning the properties of indium antimonide so that it has become possible to give a reasonably consistent picture of the band structure of this material. We will first present a summary of the information that is already known and then discuss what further properties can be deduced from the analysis presented in this paper.

Cyclotron resonance measurements[1] indicate that the conduction band minimum in indium antimonide is at $k = 0$ and is doubly degenerate* with an effective mass of $0 \cdot 013\ m_0$. Magnetoresistance measurements[2, 3] also support these conclusions.

The fundamental optical absorption measurements of FAN and GOBELI[4] show that the valence band energy at $k = 0$ is within a few hundredths of a volt of the thermal gap. These measurements were made at room temperature and gave an optical gap for direct transitions of $0 \cdot 175$ eV. ROBERTS and QUARRINGTON[5] have shown that the optical gap is strongly temperature dependent, having a value $0 \cdot 23$ eV at $T = 0^\circ$K.

The following information about the conduction and valence bands in the neighborhood of $k = 0$ can be obtained from the work of DRESSELHAUS and PARMENTER[6] on the symmetry properties of the zinc-blende structure.

The energy of the valence band in the neighborhood of $k = 0$ will be proportional to k, rather than to k^2 as it is for germanium. The valence-band maxima will most probably occur in the (111) directions away from $k = 0$. For energies very near the maxima, the energy surfaces will be ellipsoids of revolution with their axes of revolution in the (111) directions. The fourfold degeneracy of the bands at $k = 0$ is completely removed for a general k direction. The maxima in the (111) direction are singly degenerate.

For larger values of k the linear terms will be dominated by terms proportional to k^2. These k^2 terms give energy bands entirely similar to germanium. The fourfold degeneracy at $k = 0$ is broken up into a doubly degenerate heavy-mass band and a doubly degenerate light-mass band. The double degeneracy is split by an amount proportional to k by the linear k terms. The double degeneracy of the conduction band is removed by terms proportional to k^3.

A number of experiments suggest the value

---

* This double degeneracy is due to spin.

0·18 m_0 as the "average"* valence band mass. Thermoelectric power measurements by H. WEISS[7] give 0·18 m_0 at room temperature. The binding energy of holes, 0·007 eV as determined by HROSTOWSKI et al.,[8] suggests that the "average" mass is less than the density of states mass for germanium, 0·35 m_0, since holes in germanium are bound by 0·01 eV. HROSTOWSKI found that an "average" mass of 0·18 m_0 gave a good fit to his data on the deionization of holes. Cyclotron resonance measurements[1] have shown the presence of mass peaks at 0·18 m_0 and $> 1·2\ m_0$. These peaks were tentatively identified with holes.

The existence of light holes is suggested by the data of HROSTOWSKI et al.[8] on the variation of Hall coefficient with magnetic field. Their data are similar to the data of WILLARDSON et al.[9] which have been interpreted in terms of light and heavy holes.

There is a region of optical absorption of the order 100 cm^{-1} extending several hundredths of a volt beyond the edge for direct transitions. Two analyses of the absorption in this region have been made according to a method first proposed by MACFARLANE and ROBERTS.[10] The two analyses reach different conclusions concerning the band structure. BLOUNT et al. suggest that the valence band maximum is far from $k = 0$. POTTER stresses the importance of optical phonons and suggests that the valence band maximum is 0·015 eV above the energy at $k = 0$, but at a value of k much nearer zero than claimed by BLOUNT et al.

The effect of impurities on the band structure has been calculated by STERN and TALLEY.[11] They made the approximation that their impurities were periodically arranged so that they could calculate the band structure of metallic hydrogen with an appropriate effective mass and dielectric constant. The neglect of randomness is, of course, a strong objection to this procedure. Randomness can be taken into account by a perturbation approach as has been done by PARMENTER.[12]

---

* For the case of non-spherical energy surfaces the concept of an "average mass" has no precise significance since it will in general be different for every different experiment. The above summary of the experimental information for the valence band uses the word "average" loosely, but it should be understood that a more detailed analysis of each experiment in terms of a definite model is necessary in order to derive maximum information.

Perturbation theory cannot apply to bound or quasibound states but may be applicable to free states. The limitations of perturbation theory have not yet been adequately discussed for this problem. The perturbation calculation gives an appreciably smaller effect than the Stern and Talley approach.

The present paper enlarges upon the $k \cdot p$ approach of DRESSELHAUS and PARMENTER.[6] The spin-orbit interaction is also treated as a perturbation.[13] It has been assumed in this paper that because of the smallness of the energy gap at $k = 0$ it would be a good approximation to ascribe all of the unusually low effective mass of the conduction band to the mutual interaction of conduction and valence bands.

A modification of the usual perturbation technique is used whereby the mutual interaction of the conduction and valence bands via the $k \cdot p$ interaction and the k-independent spin-orbit interaction are treated exactly. The effects of higher and lower bands are treated by perturbation theory.

For energies above $k = 0$ of the order of a fraction of the band gap the conduction band becomes non-parabolic. At this point straightforward perturbation theory breaks down but the present technique remains valid. Because of the small band gap, the non-parabolic nature of the conduction band will be important at high temperature or high electron concentration (degeneracy).

In diagonalizing the mutual interaction of the conduction and valence bands the sixfold degeneracy of the valence band is split into three twofold bands corresponding to the heavy-mass band, the light-mass band and the spin-orbit split-off band of germanium. The conduction band is also doubly degenerate. With part of the valence-band degeneracy thus removed, the interaction with higher bands is calculated by second-order perturbation theory. The second-order terms involving the $k \cdot p$ interaction alone shift the bands by an amount proportional to k^2. The remaining twofold degeneracy of the bands is removed by terms proportional to k^3. The twofold degeneracy of the valence band is also removed by second-order terms involving the $k \cdot p$ interaction and the k-independent spin-orbit interaction, giving a splitting proportional to k. These second-order terms are larger than the first-order linear k terms

considered by DRESSELHAUS[6] which resulted from the k-dependent spin-orbit interaction.

The above procedure results in simpler formulae for the valence bands than were obtained by DRESSELHAUS. The simplicity results from having removed part of the degeneracy before applying perturbation theory. Before perturbation theory is applied the bands are spherical. The accuracy of the simpler expressions is less than that of the more complicated expressions. Considering the k^2 terms alone, the present procedure will be accurate if the mutual interaction energy of conduction and valence bands is large compared to the interaction energy with higher bands. This is the case in InSb. The present method always breaks down for energies very close to $k = 0$ where the linear k terms dominate the k^2 terms used to remove part of the initial valence band degeneracy.

The spin-orbit splitting of the valence band is not known but a rough estimate from atomic spectra puts it at about 0.9 eV. According to theory the light mass should be $0.015\ m_0$ with a small uncertainty due to the unknown spin-orbit splitting. The effect of higher bands on the light mass value should also be small. The heavy mass, however, is determined entirely by higher bands, receiving no contribution from the conduction band interaction. It is reasonable to assume tentatively that the $0.18\ m_0$ average valence band mass is the average value for the heavy mass band. No estimate of the warping of the heavy mass band has yet been obtained. There appears to be no place in the present theory for the $> 1.2\ m_0$ peak seen in cyclotron resonance.

The magnitude of the linear k terms is estimated very roughly in the present paper. The linear k terms cause the maximum in the valence band to be $\sim 10^{-4}$ eV above the energy at $k = 0$ for a value of $k \sim 0.3$ per cent of the way to the zone boundary in the (111) direction. For energies much greater than 10^{-4} eV the k^2 terms are dominant and the valence band of indium antimonide resembles that of germanium. The above estimate of the magnitude of the linear k terms is very close to the value given by DRESSELHAUS.[6] However the present estimate refers to a second order process involving the $k \cdot p$ interaction and the k-independent spin-orbit interaction. These terms are actually larger than the first-order terms from the k-dependent spin-orbit inter-

action which were somewhat overestimated by DRESSELHAUS.

The fundamental optical absorption is calculated in the approximation of direct transitions. The optical matrix element is determined from the electron cyclotron mass so that no arbitrary constants are involved. The agreement with the experimental data of FAN and GOBELI[4] supports the major assumptions made in this paper.

The effect of impurities in shifting the optical absorption edge has been computed. The effects of a change in band structure are negligible using PARMENTER'S[12] perturbation theory calculation. The agreement with the experimental data of FAN and GOBELI[4] for 10^{18} impurities per c.c. is satisfactory.

The indirect transitions analysis of POTTER[10] would agree with the picture of the band structure presented here if it were assumed that the linear k terms were a good deal larger than the theoretical estimate. At the present time the theoretical analysis of indirect transitions does not appear sufficiently straightforward to warrant drawing strong conclusions.

The work of HROSTOWSKI et al.[8] at 77°K appears to suggest the existence of light holes. Since the mass ratio of light and heavy holes is similar in germanium and indium antimonide the variation of Hall coefficient with magnetic field should be similar. If the size of the linear k terms were as great as POTTER'S analysis requires, the band structure at energies appropriate to 77°K would be so unlike germanium that we could not speak of "light holes".

2. BAND STRUCTURE
General.

The point of view in the present calculations is similar to the point of view taken in an earlier paper[13] in discussing the valence band structure of germanium. The difference is that here the interaction of the valence bands and the conduction band via the $k \cdot p$ interaction and the k-independent spin-orbit interaction is taken into account exactly. The effect of higher bands is taken into account by second order perturbation theory. This refinement was necessary because the band gap is so small in indium antimonide that the parabolic bands given by a straight perturbation treatment would be an inaccurate

approximation to use in analysing many experiments of practical interest.

The Schroedinger equation for the cell-periodic functions $u_k(r)$ is

$$\{(p^2/2m)+V+(\hbar/m)k \cdot p+(\hbar/4m^2c^2)[\triangledown V \times p] \cdot \sigma +$$
$$+(\hbar^2/4m^2c^2)[\triangledown V \times k] \cdot \sigma\}u_k(r) = E'_k u_k(r) \tag{1}$$

$$E'_k = E_k-(\hbar^2/2m)k^2 \tag{2}$$

where E_k is the energy of a state with wave vector k. We assume we know the solution of the equation

$$\{(p^2/2m)+V\}U_i = E_i U_i \tag{3}$$

and use the complete set of U_i as a basis for a representation. Group theory gives us the symmetry properties of the functions U_i. The functions corresponding to the conduction band are singly degenerate (doubly degenerate counting spin) and will be denoted by $S\uparrow$ and $S\downarrow$. The symbol $\uparrow$ means spin up and $\downarrow$ means spin down. The designation S refers to the fact that the functions have the symmetry properties of s functions under the operations of the tetrahedral group.[6] The functions S belong to symmetry type Γ_1 in the notation of reference 6. The valence band functions are triply degenerate (sixfold with spin) and will be written $X\uparrow$, $Y\uparrow$, $Z\uparrow$, $X\downarrow$, $Y\downarrow$, $Z\downarrow$. Again the notation serves to indicate that the functions have the symmetry properties of the atomic p-functions x, y, z under the operations of the tetrahedral group. These functions have symmetry Γ_4 in the notation of reference 6.

The third term in eq. (1) will be referred to as the $k \cdot p$ interaction, the fourth term as the k-independent or atomic-like spin-orbit interaction, and the fifth term as the k-dependent spin-orbit interaction. For the present we shall neglect the fifth term. Later we treat it as a perturbation and estimate its size. It turns out to be very small because the crystal momentum $\hbar k$ is very small compared to the atomic momentum p in the far interior of the atom where most of the spin-orbit interaction occurs.

The fifth term gives an energy linear in k in the valence band. Energies linear in k also result from second order perturbation terms containing $k \cdot p$ and the k-independent spin-orbit interaction.

These second-order terms are actually larger than the first-order terms of the k-dependent spin-orbit interaction. The magnitudes of these terms will be discussed later.

The linear k terms are responsible for the fact that the valence band of indium antimonide is more complex than that of germanium.[6] We include these terms by perturbation theory. This procedure is only valid for energies such that the k^2 terms which remove the valence band degeneracy are large compared to the k terms. For lower energies the more complicated expressions given by DRESSELHAUS[6] must be used.

Valence and Conduction Band Interactions

We take the k vector in the z direction and consider the Hamiltonian corresponding to the first four terms of eq. (1). The mutual interaction of the conduction and valence bands leaves the bands doubly degenerate. We take as a basis $|iS\downarrow\rangle$, $|(X-iY)\uparrow/\sqrt{2}\rangle$, $|Z\downarrow\rangle$, $|(X+iY)\uparrow/\sqrt{2}\rangle$, $|iS\uparrow\rangle$, $|-(X+iY)\downarrow/\sqrt{2}\rangle$, $|Z\uparrow\rangle$, $|(X-iY)\downarrow/\sqrt{2}\rangle$. The first four functions are respectively degenerate with the last four. The 8×8 interaction matrix may be written

$$\begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix} \tag{4}$$

$$H = \begin{bmatrix} E_s & 0 & kP & 0 \\ 0 & E_p-\Delta/3 & \sqrt{2}\Delta/3 & 0 \\ kP & \sqrt{2}\Delta/3 & E_p & 0 \\ 0 & 0 & 0 & E_p+\Delta/3 \end{bmatrix}.$$

The positive constant Δ is the spin-orbit splitting of the valence band. The real quantity P is defined by

$$P = -i(\hbar/m) \langle S|p_z|Z \rangle$$
$$\Delta = \frac{3\hbar i}{4m^2c^2} \left\langle X \left| \frac{\partial V}{\partial x}p_y - \frac{\partial V}{\partial y}p_x \right| Y \right\rangle. \tag{5}$$

E_s and E_p refer to the eigenvalues of the Hamiltonian of eq. (3). E_s corresponds to the conduction band and E_p to the valence band. Symmetry properties have been used. If the k vector is not in the z direction, the Hamiltonian is more complicated but it can be transformed to the form of eq. (4) by a rotation of the basis functions.

$$\begin{bmatrix} \uparrow' \\ \downarrow' \end{bmatrix} = \begin{bmatrix} e^{-i\phi/2}\cos\theta/2 & e^{i\phi/2}\sin\theta/2 \\ -e^{-i\phi/2}\sin\theta/2 & e^{i\phi/2}\cos\theta/2 \end{bmatrix} \begin{bmatrix} \uparrow \\ \downarrow \end{bmatrix}$$

(6)

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

(7)

$$S' = S. \tag{8}$$

The angles θ, ϕ are the usual polar angles of the k vector referred to the crystal symmetry axes x, y, and z with θ measured from z and ϕ measured from x. This transformation would be obvious if the functions X, Y, Z transformed like the spherical harmonics x, y, z under the full spherical group rather than just under the tetrahedral group.

The four double roots of the secular equation resulting from eq. (4) can be written

$$E' = 0 \tag{9}$$

$$E'(E'-E_G)(E'+\Delta) - k^2 P^2(E'+2\Delta/3) = 0 \tag{10}$$

where

$$E_s = E_G$$
$$E_p = -\Delta/3 \tag{11}$$

E' is defined in eq. (2). E_G is the band gap at $k = 0$.

For small values of k^2 the solutions of eq. (9) and (10) give parabolic bands.

$$E_c = E_G + \frac{\hbar^2 k^2}{2m} + \frac{P^2 k^2}{3}\left(\frac{2}{E_G} + \frac{1}{E_G+\Delta}\right).$$

$$E_{v1} = \hbar^2 k^2/2m$$

$$E_{v2} = \frac{\hbar^2 k^2}{2m} - \frac{2P^2 k^2}{3E_G}$$

(12)

$$E_{v3} = -\Delta + \frac{\hbar^2 k^2}{2m} - \frac{P^2 k^2}{3(E_G+\Delta)}.$$

The above equations give the expressions for the conduction band, the heavy-mass band, the light-mass band, and the split-off band in that order. The equations are, of course, not complete because the effects of higher bands have not yet been included. The expressions are probably fairly good for all bands except the heavy-mass band E_{v1}.

A different approximation can be made in simplifying the solutions of eq. (10), namely that $\Delta \gg kP, E_G$. The solutions then become

$$E_c = \frac{\hbar^2 k^2}{2m} + \frac{E_G + (E_G^2 + 8P^2 k^2/3)^{\frac{1}{2}}}{2} \tag{13}$$

$$E_{v1} = \hbar^2 k^2/2m$$

$$E_{v2} = \frac{\hbar^2 k^2}{2m} + \frac{E_G - (E_G^2 + 8P^2 k^2/3)^{\frac{1}{2}}}{2}$$

$$E_{v3} = -\Delta + \frac{\hbar^2 k^2}{2m} - \frac{P^2 k^2}{3(E_G+\Delta)}.$$

These equations exhibit simply the nonparabolic nature of the bands E_c and E_{v2} when $|(Pk)| > E_G$.

The doubly degenerate wave functions which result from the diagonalization of the Hamiltonian of eq. (4) may be written

$$\phi_{i\alpha} = a_i[iS\downarrow]' + b_i[(X-iY)\uparrow/\sqrt2]' + c_i[Z\downarrow]'$$

$$\phi_{i\beta} = a_i[iS\uparrow]' + b_i[-(X+iY)\downarrow/\sqrt2]' + c_i[Z\uparrow]'$$

$$\phi_{v1\alpha} = [(X+iY)\uparrow]'/\sqrt2$$

$$\phi_{v1\beta} = [(X-iY)\downarrow]'/\sqrt2. \tag{14}$$

a_i, b_i, c_i are real coefficients; the index i refers to the bands E_c, E_{v2}, E_{v3}. The primes denote that the wave functions for a general k direction are rotated according to eqs. (6) and (7). The general expressions for a_i, b_i, c_i may be written

$$a_i = kP(E_i' + 2\Delta/3)/N$$

$$b_i = (\sqrt2\Delta/3)(E_i' - E_G)/N \tag{15}$$

$$c_i = (E_i' - E_G)(E_i' + 2\Delta/3)/N$$

where N is a normalizing factor equal to the square root of the sum of the squares of the numerators. The E_i' are the solutions of eq. (10). For small k eqs. (15) become

$$a_c = 1, \quad b_c = c_c = 0$$

$$a_{v2} = 0, \quad b_{v2} = (1/3)^{\frac{1}{2}}, \quad c_{v2} = (2/3)^{\frac{1}{2}} \tag{16}$$

$$a_{v3} = 0, \quad b_{v3} = (2/3)^{\frac{1}{2}}, \quad c_{v3} = -(1/3)^{\frac{1}{2}}.$$

In the approximation $\Delta \gg kP$, E_G the eqs. (15) become

$$a_c = \left(\frac{\eta+E_G}{2\eta}\right)^{\frac{1}{2}}; \quad b_c = \left(\frac{\eta-E_G}{6\eta}\right)^{\frac{1}{2}};$$

$$c_c = \left(\frac{\eta-E_G}{3\eta}\right)^{\frac{1}{2}}$$

$$a_{v2} = -\left(\frac{\eta-E_G}{2\eta}\right)^{\frac{1}{2}}; \quad b_{v2} = \left(\frac{\eta+E_G}{6\eta}\right)^{\frac{1}{2}};$$

$$c_{v2} = \left(\frac{\eta+E_G}{3\eta}\right)^{\frac{1}{2}}$$

$$a_{v3} = 0; \quad b_{v3} = (2/3)^{\frac{1}{2}}; \quad c_{v3} = -(1/3)^{\frac{1}{2}} \quad (17)$$

$$\eta = [E_G{}^2 + 8k^2P^2/3]^{\frac{1}{2}}.$$

Perturbations From Other Bands

The perturbing effects of higher and lower bands on the conduction and valence bands can be taken into account by second order perturbation theory. We take as our representation the complete set of functions U_i of the Hamiltonian operator of eq. (3) except that for the valence and conduction bands we use the wave functions of eq. (14).

We define a set of constants analogous to those used by DRESSELHAUS *et al.*[14]

$$A = \frac{\hbar^2}{m^2} \sum_{n,l} \frac{|\langle X|p_x|\Gamma_1(nl)\rangle|^2}{E_0-E_n}$$

$$B = \frac{\hbar^2}{m^2} \sum_{n,l} \frac{|\langle X|p_x|\Gamma_3(nl)\rangle|^2}{E_0-E_n}$$

$$C = \frac{\hbar^2}{m^2} \sum_{n,l} \frac{|\langle X|p_y|\Gamma_4(nl)\rangle|^2}{E_0-E_n}$$

$$D = \frac{\hbar^2}{m^2} \sum_{n,l} \frac{|\langle X|p_y|\Gamma_5(nl)\rangle|^2}{E_0-E_n}$$

$$F = \frac{\hbar^2}{m^2} \sum_{n,l} \frac{|\langle S|p_x|\Gamma_4(nl)\rangle|^2}{E_0-E_n}$$

$$G = \frac{\hbar^2}{m^2} \sum_{n,l} \frac{\langle S|p_x|\Gamma_4(nl)\rangle\langle\Gamma_4(nl)|p_y|Z\rangle}{E_0-E_n}. \quad (18)$$

The symbol Γ_i (nl) refers to the lth function of the nth band having symmetry type Γ_i in the notation of reference 6. The summation, of course, does not extend over the conduction and valence bands. The energy E_0 is the valence band energy at $k=0$. More accurate expressions for the perturbation energies would result if we used the appropriate E_i given by eqs. (9) and (10) in place of E_0. The number of unknown parameters would be much larger if this were done.

In determining the effects of higher bands on the wave functions of eq. (14) we consider second-order perturbation matrix elements between degenerate wave functions $\phi_{i\alpha}$, $\phi_{i\beta}$ but not between functions ϕ_i, ϕ_j whose degeneracy has been removed by the interactions already discussed. In terms of the above constants the perturbing matrix elements can be written

$$H_{\alpha\alpha}(i) = H_{\beta\beta}(i)$$

$$= a_i{}^2Fk^2 + b_i{}^2 \left\{ (C+D)k^2 + \left(\frac{3B}{2}-2C\right)\frac{k_x{}^2k_y{}^2+k_x{}^2k_z{}^2+k_y{}^2k_z{}^2}{k^2} \right\} +$$

$$+ c_i{}^2 \left\{ (A+B)k^2 + (-3B+4C)\left(\frac{k_x{}^2k_y{}^2+k_x{}^2k_z{}^2+k_y{}^2k_z{}^2}{k^2}\right) \right\} \quad (19)$$

$$H_{\alpha\beta}(i) = 2\sqrt{2}a_ib_iG \left\{ \frac{kk_z(k_y{}^2-k_x{}^2)+ik_xk_y(3k_z{}^2-k^2)}{k(k_x{}^2+k_y{}^2)^{\frac{1}{2}}} \right\}$$

$$|H_{\alpha\beta}(i)| = \frac{2\sqrt{2}|a_ib_iG|}{k}\{(k_x{}^2+k_y{}^2)(k_x{}^2+k_z{}^2)(k_y{}^2+k_z{}^2)-8k_x{}^2k_y{}^2k_z{}^2\}^{\frac{1}{2}}. \quad (20)$$

The index i refers to the bands E_c, E_{v2} and E_{v3}. The perturbation energy is then

$$H_{\alpha\alpha}(i) \pm H_{\alpha\beta}(i) \qquad (21)$$

which is to be added to the energy determined from the cubic equation (10). The term $H_{\alpha\beta}$ removes the double degeneracy of the bands except for the (100) and (111) directions.

For the heavy-mass band E_{v1}

$$H_{\alpha\alpha} = (C+D)k^2 +$$

$$+ \left(\frac{3B}{2} - 2C\right) \frac{k_x^2 k_y^2 + k_x^2 k_z^2 + k_y^2 k_z^2}{k^2}$$

$$H_{\alpha\beta} = 0. \qquad (22)$$

The four constants A, B, C, D occur in three combinations which correspond to the constants L, M, N of reference 14. The constant D arises from spherical harmonics f and higher, so that it probably will be quite small.

The expression for the energy of the heavy mass band E_{v1} is simpler than the expression given by DRESSELHAUS et al.[14] for germanium. The reason for the simplification is that the degeneracy be-

tween the low-mass and high-mass valence bands at $k = 0$ is removed by the conduction band interaction alone in the present treatment. In the more accurate treatment of reference 14, all bands act to remove this degeneracy. The terms neglected in the present treatment are of the order of $E_G A/P^2$, $E_G B/P^2$, $E_G C/P^2$, $E_G D/P^2$ where P is the matrix element of eq. (5). The expressions given by the present treatment are identical with the expressions of reference 14 for the (100) and (111) directions.

Linear Terms

Energies linear in k are possible in the valence bands E_{v1} and E_{v2}. These linear terms act to split the double degeneracy of the bands E_{v1} and E_{v2}. No such terms are possible in the conduction band. Linear terms are altogether absent in germanium because of inversion symmetry. In the absence of spin-orbit interaction linear k terms are absent by time reversal symmetry even without inversion symmetry. These facts have been pointed out in reference 6.

The linear k terms are represented by an interaction Hamiltonian which has the form

$$H = K \begin{bmatrix} 0 & \sqrt{3}(k_x - ik_y) & k_x + ik_y & -2k_z \\ \sqrt{3}(k_x + ik_y) & 0 & -2k_z & -(k_x - ik_y) \\ k_x - ik_y & -2k_z & 0 & \sqrt{3}(k_x + ik_y) \\ -2k_z & -(k_x + ik_y) & \sqrt{3}(k_x - ik_y) & 0 \end{bmatrix} \qquad (23)$$

with respect to a set of basis functions $|(X+iY)\uparrow/\sqrt{2}\rangle$, $|(X-iY)\downarrow/\sqrt{2}\rangle$, $|(2/3)^{1/2}Z\uparrow - (1/6)^{1/2}(X+iY)\downarrow\rangle$, $|(2/3)^{1/2}Z\downarrow + (1/6)^{1/2}(X-iY)\uparrow\rangle$.

Any set of Γ_8 functions may be shown to have a linear k Hamiltonian of this form if a representation is chosen having the same tetrahedral symmetry as the above basis functions.

The constant K arises from matrix elements of three types.

$$K_1 = -\frac{\hbar^2}{4\sqrt{3}m^2c^2} \left\langle X \left| \frac{\partial V}{\partial y} \right| Z \right\rangle \qquad (24)$$

$$K_2 = -\frac{\hbar^2}{4\sqrt{3}m^3c^2} \sum_{n,j} \frac{\left\langle X \left| p_x \right| \Gamma_3(n,j) \right\rangle \left\langle \Gamma_3(n,j) \left| \frac{\partial V}{\partial z}p_x - \frac{\partial V}{\partial x}p_z \right| Y \right\rangle}{E_0 - E_n}$$

$$K_3 = \frac{\hbar^2}{2\sqrt{3}m^3c^2} \sum_{n,j} \frac{\left\langle X \left| p_y \right| \Gamma_5(n,j) \right\rangle \left\langle \Gamma_5(n,j) \left| \frac{\partial V}{\partial z}p_x - \frac{\partial V}{\partial x}p_z \right| X \right\rangle}{E_0 - E_n}.$$

The constant K_1 results from first-order perturbation theory applied to the k-dependent spin-orbit interaction, the fifth term in eq. (1). The constants K_2 and K_3 result from second-order perturbation theory applied to the $k \cdot p$ term and the k-independent spin-orbit interaction term in eq. (1). The constants K_2 and K_3 are actually much larger than K_1 because the momenta which make most of the contribution to the spin-orbit splitting are very large compared to $\hbar k$.

For small energies, the linear k terms of eq. (23) will dominate all the other terms which have been considered, since these terms vary like k^2 or higher powers of k. DRESSELHAUS[6] has discussed the band structure appropriate to the valence band taking into account linear k terms and second order perturbation $k \cdot p$ terms. The constants L, M, N which he uses are related to the constants defined in this paper as follows:

$$L = A + B - P^2/E_G$$
$$M = C + D \qquad (25)$$
$$N = A - (P^2/E_G) - (B/2) + C - D.$$

For large energies, the k^2 and higher terms will be dominant and the terms of eq. (23) can be treated as perturbations.

The matrix elements of the linear k interaction between the heavy-mass band wave functions of eq. (14) are

$$H_{v1\alpha\alpha} = H_{v1\beta\beta} = 0$$

$$H_{v1\alpha\beta} = 3\sqrt{3}K\left[\frac{kk_z(k_x^2 - k_y^2) + ik_x k_y(k^2 + k_z^2)}{k^2(k_x^2 + k_y^2)^{\frac{1}{2}}}\right].$$

$$(26)$$

The perturbation energy is given by

$$\delta E_{v1} = \pm 3\sqrt{3}(K/k^2) \times$$
$$\times [(k_x^2 + k_y^2)(k_x^2 + k_z^2)(k_y^2 + k_z^2)]^{\frac{1}{2}}. \qquad (27)$$

The bands are split in all but the (100) direction.

The matrix elements of the linear k interaction between the light-mass bands are

$$H_{v2\alpha\alpha} = H_{v2\beta\beta} = 0$$

$$H_{v2\alpha\beta} =$$
$$-\sqrt{3}K\left[\frac{kk_z(k_x^2 - k_y^2) + ik_x k_y(k^2 - 3k_z^2)}{k^2(k_x^2 + k_y^2)^{\frac{1}{2}}}\right]$$

$$(28)$$

referred to the basis functions

$$\Phi_{v2\alpha} = \left[\frac{(X - iY)}{6^{\frac{1}{2}}}\uparrow + \left(\frac{2}{3}\right)^{\frac{1}{2}} Z\downarrow\right]'$$
$$\Phi_{v2\beta} = \left[-\frac{(X + iY)}{6^{\frac{1}{2}}}\downarrow + \left(\frac{2}{3}\right)^{\frac{1}{2}} Z\uparrow\right]'. \qquad (29)$$

The primes denote rotation by the operators of eqs. (6) and (7). The form of the light-mass band wave functions is less general than the form given in eq. (14) and is valid only in the neighborhood of $k = 0$. The reason for the restriction is that it is only for this simpler form that the three interactions of eq. (24) give the same result. The perturbation energy is

$$\delta E_{v2} = \pm\sqrt{3}(K/k^2) \times$$
$$\times [(k_x^2 + k_y^2)(k_x^2 + k_z^2)(k_y^2 + k_z^2) - 8k_x^2 k_y^2 k_z^2]^{\frac{1}{2}}.$$

$$(30)$$

The bands are split in all but the (100) and (111) directions.

The linear terms give a maximum in the valence band away from $k = 0$ in the (111) direction. This maximum is associated with the heavy mass band E_{v1}. The simple expression (27) may be of semiquantitative value in discussing the energy surfaces in the neighborhood of the valence band maximum because the low mass band energy at this point in k space is large compared to the linear k-term energy, as the approximation leading to (27) requires. An improved expression could be obtained by doing second order perturbation theory with the linear k interaction.

The energy surfaces in the neighborhood of the valence band maxima are ellipsoids of revolution with the axis of revolution in the (111) direction. The maxima occur at a value of k given by

$$k_0 = \left|\frac{2\sqrt{2}K}{(\hbar^2/m) + B + (2C/3) + 2D}\right|. \qquad (31)$$

The energy of the maxima relative to $k = 0$ is

$$E_0 = \left|\frac{4K^2}{(\hbar^2/m) + B + (2C/3) + 2D}\right|. \qquad (32)$$

The longitudinal and transverse masses are given by the expressions

$$(1/m_l) = (1/\hbar^2)[(\hbar^2/m) + B + (2C/3) + 2D]$$
$$(1/m_t) = (1/\hbar^2)[(11C/3) + 3D - (B/2) + 3\hbar^2/2m].$$

$$(33)$$

The transverse mass was derived using eq. (27) and eq. (22) hence it is only approximate. The other information can be derived from the expression for the energy in the (111) direction given by DRESSELHAUS.[6] See eq. (13) of this reference.

3. ESTIMATES OF BAND STRUCTURE CONSTANTS

The best determination of the band gap, E_G, appears to be the fundamental optical-absorption edge data of ROBERTS and QUARRINGTON.[5] Their data indicate a highly temperature-dependent edge

which extrapolates to a value $E_G = 0.23$ eV at $T = 0°$K.

The spin-orbit splitting of the valence band is an unknown quantity which no experiment has indicated directly in the case of indium antimonide. The spin-orbit splittings of atomic p functions in indium and antimony are respectively 0.28 eV and ~ 0.8 eV from atomic spectra. As ELLIOTT[15] has pointed out, the atomic splittings are useful because the main spin-orbit interaction occurs so deep in the atom that the principal effect of the solid is to alter the normalization and to contaminate the wave function with higher spherical harmonics. The only reliable evidence of the magnitude of this effect comes from germanium where the one-electron atomic splitting is 0.20 eV and the solid splitting is 0.29 eV as shown by free carrier absorption.[16] The average ionicity of the valence wave functions in InSb has been estimated to be 0.35 by analysis of the reflectivity in the region of the lattice absorption.[17]

For want of a better estimate we assume that the valence electron at $k=0$ spends 35 per cent of its time on indium and 65 per cent of its time on antimony though this is not necessarily implied by the reflectivity analysis. If we make this assumption about the wave function and correct for normalization by the factor appropriate to germanium we get an estimated spin-orbit splitting of the order of 0.9 eV for indium antimonide.

The effective mass in the conduction band is $0.013\, m_0$ as given by cyclotron resonance measurements.[1] From the standpoint of the $k \cdot p$ approach it seems reasonable to attribute this unusually low mass to the strong repulsion of the closely adjacent conduction and valence bands. If we assume that the conduction-band–valence-band interaction is much larger than all other interactions responsible for the conduction electron mass we can determine the absolute value of the matrix element P in eq. (5) with the help of eq. (12). We obtain $P^2 = 0.44$ atomic units.

The average hole mass m_{v1} is $0.18\, m_0$ as determined by thermo-electric power data[7] and data on the deionization of holes.[8] Cyclotron resonance[1] measurements also suggest a mass of this same value. The value $0.18\, m_0$ is not inconsistent with the binding energy of holes.[8] If we take the (110) direction as a good "average" direction we have from eq. (22) the relation

$$\frac{\hbar^2}{2}\left(\frac{1}{m_{v1}}+\frac{1}{m_0}\right) = \left|\frac{3}{8}B+\frac{1}{2}C+D\right|. \quad (34)$$

The constant C may result from bands derived from atomic p functions, B requires d functions, D requires f functions. D is presumably small, but B and C may be comparable, according to the results in germanium.

No evidence relating to the size of the matrix elements F and G is available at present. Since F and G arise from the same band as C it is reasonable to suppose that they are comparable in magnitude. However, G would be zero by symmetry in a diamond structure. The constant G is important because it gives the splitting of the two conduction bands. A rough estimate of this effect gives a splitting of the order 0.1 eV at the Fermi level in n-type material with 10^{19} free electrons assuming $G \sim C$. A measurement of this splitting would aid in the determination of G.

Another important constant is the coefficient K of the linear k terms in the valence band structure. The constant K_2 appears to be considerably larger than K_1. An extremely rough estimate of the magnitude of K_2 was made which gave $K_2 \sim 2 \times 10^{-3}$ atomic units. This value corresponds to a maximum in the valence band $\sim 10^{-4}$ eV above the energy at $k = 0$ according to eq. (32). The value of k at the maximum is about 0.3 per cent of the value at the zone edge.

This estimate was made using published HARTREE[18] functions to calculate the spin-orbit matrix element for As and arbitrarily extrapolating to Sb by multiplying by the ratio of the atomic spin-orbit splittings of antimony and arsenic. The Γ_3 band involved in the estimate of K_2 was taken to be the $4d$ band of Sb. The $4d$ band was assumed to be about 30 eV below the valence band.* The matrix element of P between the valence band wave function and the $4d$ wave function was arbitrarily taken to be of the order $P \sim 0.7$ atomic units. The spin-orbit interaction couples the $4d$ wave function to the $5d$ part of the valence band wave function. It was assumed that 50 per cent of the valence band wave function had $5d$ character.

At present there appears to be no reliable experimental evidence which bears on the magnitude of the linear k terms. Cyclotron resonance measurements should give the most reliable information.

---

* X-ray data shows the $4d$–$5s$ energy separation to be 26 eV in antimony.

R

At temperatures low compared to the energy of the valence band maximum, cyclotron resonance should show ellipsoids of revolution in the (111) direction. The ellipsoidal masses are given approximately by eq. (33). The longitudinal mass should be close to the average valence band mass determined from higher temperatures. The size of the transverse mass cannot be estimated accurately since the relative sizes of the constants B and C is not known. If C is large compared to B the ellipsoids will be prolate with an axis ratio of the order of 5. If B is large compared to C the extremum in the (111) direction will be a saddle point rather than a maximum.

At temperatures high compared to the energy of the valence band maximum the cyclotron resonance results should be similar to p-type germanium.

The cubic of eq. (10) has been solved numerically with the parameters $P^2 = 0.44$ atomic units, $E_G = 0.23$ eV, $\Delta = 0.9$ eV. The heavy mass band E_{v1} was taken to have an average curvature corresponding to an effective mass $m_{v1} = 0.18\ m_0$. The magnitude of the heavy mass m_{v1} gives evidence of significant perturbing effects from higher bands. These higher bands will have an important

effect on the conduction band structure for energies of the order 0.5 eV above the bottom of the band. The bands resulting from the cubic equation (10) were corrected for higher bands according to eq. (19). The correction was made taking (110) as an average direction and assuming $C \gg B$, D. C can then be evaluated from $m_{v1} = 0.18\ m_0$ using eq. (22). The assumption $C \gg B$ is very questionable. If $B \gg C$, the perturbation correction is one-third of the correction assuming $C \gg B$. Actually F and G are probably of the same order as C since they come from the same band but they are not included in these calculations for lack of information.

The computed band structure is shown in Fig. 1. The conduction band before correcting for the constant C is shown as a dashed curve. The nonparabolic nature of the conduction band for energies of the order 0.5 eV above $k = 0$ is quite apparent.

The nonparabolic region of the conduction band is associated with a wave function having a large admixture of p-symmetry. The coefficients a_i, b_i, c_i of eq. (14) as determined from the Hamiltonian of eq. (4) are shown in Fig. 2.
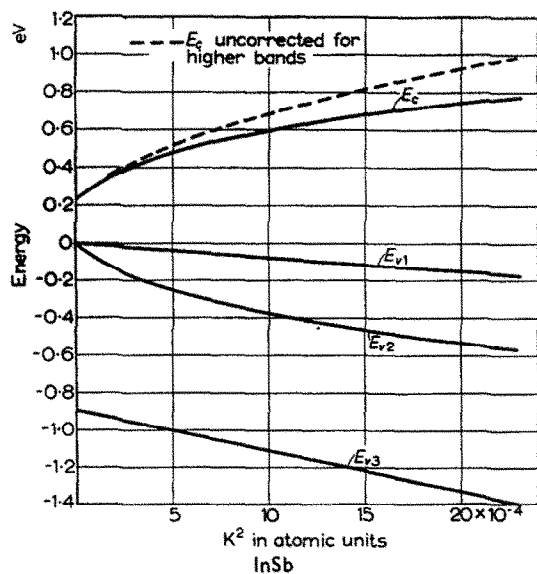

FIG. 1. Valence and conduction band energies vs. k^2 for an average direction in indium antimonide. The splitting of the double degeneracy of the bands is not given by these calculations.
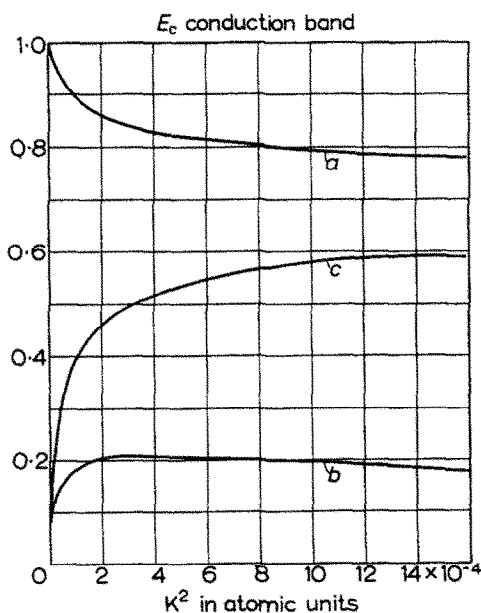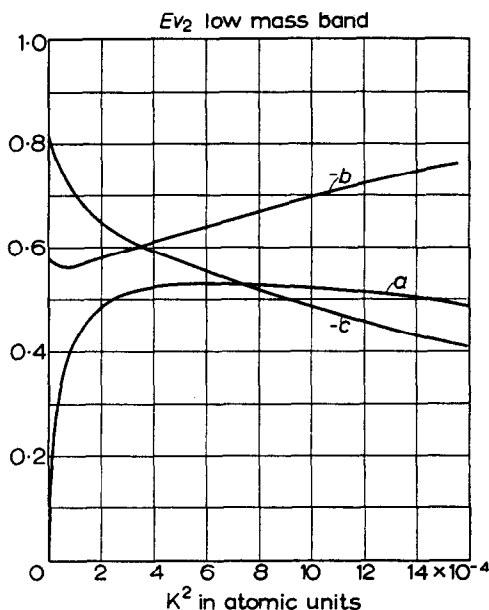

FIG. 2 (a).

FIG. 2 (*b*).
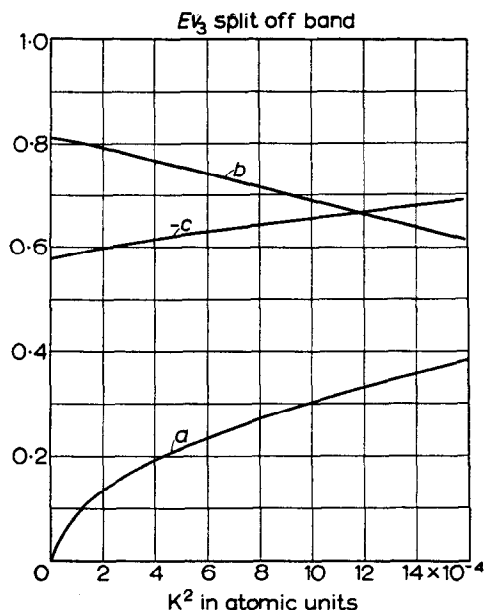
FIG. 2 (*c*).

FIG. 2. Wave function coefficients vs.k^2 in indium antimonide. See eq. (14) for definition of *a*, *b*, *c*.

4. OPTICAL ABSORPTION

Since we know the matrix element P of eq. (5) we can compute the fundamental optical absorption of InSb using the wave functions of eq. (14).

The absorption constant α may be written

$$\alpha = \frac{4\pi^2 e^2 \hbar}{m^2 cnE} \sum_j M_j^2 \rho_j(E) \qquad (35)$$

where n is the index of refraction, E is the energy of the photon absorbed, M_j^2 is the square of the optical matrix element averaged over direction and summed over degenerate bands. The summation index j refers to the three twofold valence bands E_{v1}, E_{v2}, E_{v3}. The value of M_j^2 is computed from eqs. (14) and (5)

$$M_j^2 = \frac{2m^2 P^2}{3\hbar^2} \{(a_c c_j + c_c a_j)^2 + (a_c b_j - b_c a_j)^2\}. \qquad (36)$$

The coefficients a, b, c are defined by the eqs. (15). a_c, b_c, c_c refer to the conduction band. For the heavy mass band, E_{v1}; $a = 0$, $b = 1$, $c = 0$. Values of these coefficients are given in Fig. 2.

The density of states ρ_j is given by

$$\rho_j = k^2/2\pi^2 |(dE_c/dk) - (dE_j/dk)|. \qquad (37)$$

The optical absorption has been computed using the matrix element $P^2 = 0\cdot 44$ atomic units determined from the cyclotron resonance mass and using the computed band structure shown in Fig. 1. The accuracy of the optical absorption calculations is determined by the accuracy of the band structure calculations. The results are shown in Fig. 3 together with the experimental data of FAN and GOBELI.[4] The theoretical calculations apply to $T = 0^\circ$K whereas the experimental data were taken at room temperature. The entire theoretical curve was shifted by a constant energy $0\cdot 06$ eV, corresponding to a room temperature band gap of $0\cdot 17$ eV. Lattice dilation and phonon interaction effects also influence the temperature dependence. We have not included these effects; consequently the best test of the present theory will be given by low temperature optical absorption results when they become available. The theoretical curve agrees with the experimental curve except that it is about a factor of $1\cdot 5$ too low in

absolute value. It is felt that the extent of the agreement between theory and experiment constitutes good evidence in support of the major assumptions about band structure made in this paper.
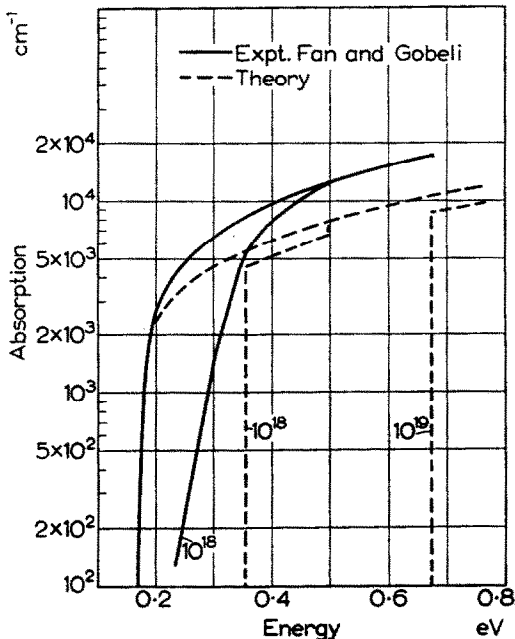
The effect of n-type impurities on the optical absorption can be taken into account in the manner suggested by BURSTEIN.[19] In impure n-type specimens no bound states exist because the impurities are too close together. At absolute zero the electrons fill a sphere of radius k_F in k space according to the formula

$$n = k_F^3/3\pi^2. \qquad (38)$$

In the approximation of direct transitions at $T = 0°K$ the electrons in the conduction band serve to cut off the absorption resulting from k values less than k_F. This effect is shown in Fig. 3 for concentrations of 10^{18} and 10^{19}. The main edge results from cutting off the absorption due to the heavy-mass band. The small additional absorption due to the light-mass band is cut off

at higher energy. The absorption has been shifted to account for the change in band gap with temperature. In this case comparison is further complicated by the effect of temperature in allowing direct transitions to states below the Fermi level as discussed by KAISER and FAN.[20] The agreement of theory and experiment on the location of the cut off appears to be fairly good. It would be very desirable to be able to make comparison to data taken at lower temperatures.

The effect of phonons and impurity scattering is to allow indirect transitions which will round off the corners shown in Fig. 3. The threshold for indirect transitions will be at an energy lower than the threshold for direct transitions by the amount

$$\Delta E = \hbar^2 k_F^2/2(0.18 m_0). \qquad (39)$$

No account has been taken of the effect of impurities on the band structure. If this effect is calculated using perturbation theory by the method of PARMENTER,[12] the change in band structure is found to be very small.

REFERENCES

1. DRESSELHAUS G., KIP A. F., KITTEL C., and WAGONER *Phys. Rev.* **98**, 556 (1955).
2. TANENBAUM M., PEARSON G. L., and FELDMAN W. L. *Phys. Rev.* **93**, 912 (1954).
3. FREDERIKSE H. P. R. and HOSLER W. R. *Bull. Amer. Phys. Soc.* **6**, 298 (June 1956).
4. FAN H. Y. and GOBELI G. W. *Bull. Amer. Phys. Soc.* **3**, 111 (March 1956).
5. ROBERTS V. and QUARRINGTON J. E. *J. Electronics* **1**, 152 (1955).
6. DRESSELHAUS G. *Phys. Rev.* **100**, 580 (1955). PARMENTER R. H. *Phys. Rev.* **100**, 573 (1955).
7. WEISS H. *Z. Naturforschung* **11a**, 131 (1956).
8. HROSTOWSKI H. J., MORIN F. J., GEBALLE T. H., and WHEATLEY G. H. *Phys. Rev.* **100**, 1672 (1955).
9. WILLARDSON R. K., HARMAN T. C., and BEER A. C. *Phys. Rev.* **96**, 1512 (1954).
10. MACFARLANE G. C. and ROBERTS V. *Phys. Rev.* **97**,

1714 (1955). BLOUNT E., CALLAWAY J., COHEN M., DUMKE W., and PHILLIPS J. *Phys. Rev.* **101**, 563 (1956). POTTER R. F. *Phys. Rev.* (to be published).

11. STERN F. and TALLEY R. M. *Phys. Rev.* **100**, 1638 (1955)

12. PARMENTER R. H. *Phys. Rev.* **97**, 587 (1955).

13. KANE E. O. *J. Phys. Chem. Solids* **1/2**, 82 (1956). This paper contains a bibliography of work on the $k \cdot p$ approach.

14. DRESSELHAUS G., KIP A. F., and KITTEL C. *Phys. Rev.* **98**, 368 (1955).

15. ELLIOTT R. J. *Phys. Rev.* **96**, 266 (1954).

16. KAHN A. H. *Phys. Rev.* **97**, 1647 (1955).

17. SPITZER W. G. and FAN H. Y. *Phys. Rev.* **99**, 1893 (1955).

18. HARTREE W., HARTREE D. R., and MANNING M. F. *Phys. Rev.* **59**, 299 (1941).

19. BURSTEIN E. *Phys. Rev.* **93**, 632 (1954).

20. KAISER W. and FAN H. Y. *Phys. Rev.* **98**, 966 (1955).