Electronic Spectra of Crystalline Germanium and Silicon*†

DAVID BRUST‡

Argonne National Laboratory, Argonne, Illinois

(Received 9 December 1963)

A detailed calculation of the energy bands of germanium and silicon has been performed by use of the pseudopotential method. The first three potential coefficients have been determined empirically, and all higher ones set equal to zero. This potential was used to compute the energy eigenvalues at $\sim$50 000 points throughout the Brillouin zone. By use of this sample, we calculated the imaginary part of the dielectric constant in the optical and near ultraviolet where direct transitions between the valence and low-lying conduction bands dominate the response. Photoelectric yield curves were obtained for comparison with recent experiments. In all cases agreement of theory and experiment was reasonable. Energy contours were constructed in several of the principal symmetry planes. These were used to explain the structure in the optical properties of Ge and Si in terms of transitions near certain important critical points. Effective masses and the static dielectric constant were also computed.

I. INTRODUCTION

RECENTLY, very precise reflectivity measurements have been taken on[1] Ge and[2] Si in the optical and near ultraviolet region of the electromagnetic spectrum. When the reflectance, $|r(\omega)|^2$, is plotted as a function of ω, the resulting curves show detailed structure. The energies involved (1.5 eV$<\hbar\omega<10$ eV) are such that direct interband transitions between states lying near to the forbidden gap are expected to dominate the dielectric properties of these materials.

Before attempting an explanation of the observed data, it is necessary to know the behavior of the valence and low-lying conduction bands throughout the Brillouin zone. Prior theoretical work on the energy bands of[3-5] Ge and[6-9] Si has been confined to calculations at points and along lines of particularly high symmetry. This is inadequate for our purposes; we find it necessary to extend the energy-band calculations so as to sample all of the zone.

The previous energy-band work will, however, be used as a starting point. First of all, we shall employ a simplified version of the orthogonalized plane-wave method (OPW method) utilized in Refs. 3–9. Secondly, we shall make use of the eigenvalues at symmetry points in order to deduce the numerical values of the pseudopotential matrix elements used as computational parameters in this simplified approach (see Sec. III of this paper). A reasonably precise knowledge of these eigen-

values (or rather the differences between them) is necessary to undertake the present computations.

For our purposes the energy levels can be put into two categories: levels which are sensitive to small changes in the crystal potential and those which are not. The former have, in the past, been computed *ab initio* with an uncertainty of order 3 eV,[10] while the latter were obtained to within $\sim\frac{1}{2}$ eV. The principal features (peaks and edges) of the optical curves are separated by less than 1 eV. It is necessary, therefore, to know the levels with a considerably greater accuracy than the current *a priori* calculations permit. Fortunately, a good deal of very fine experimental information is available which, when combined with the results of band theory, allows a sharp definition of both the sensitive and insensitive levels. Ehrenreich, Philipp, and Phillips[11] have given a good account of most of the main energy gaps at the symmetry points. Table I is largely taken from their work.

TABLE I. Some of the major energy gaps at the symmetry points $\Gamma = (2\pi/a)(0,0,0)$, $L = (2\pi/a)(\frac{1}{2},\frac{1}{2},\frac{1}{2})$, and $X = (2\pi/a)(1,0,0)$. The labeling of states is after Bouckaert, Smoluchowski, and Wigner.[a] The experimental reference list is by no means exhaustive. Furthermore, energy-band calculations, not included in the table, played an important role in making the identifications.

Gap	Energy (eV) Ge	Energy (eV) Si	Refs. Ge	Refs. Si
$L_{3'} \rightarrow L_1$	2.1 2.3	3.7	1, 13, 16	17, 18, 19
$L_{3'} \rightarrow L_3$	5.9 6.1	5.5	1	2
$\Gamma_{25'} \rightarrow \Gamma_{2'}$	0.8	$\sim$3.0	17, 20	17, 21
$\Gamma_{25'} \rightarrow \Gamma_{15}$	3.1	3.5	19, 22	2, 21
$X_4 \rightarrow X_1$	4.5	4.5	1	2

[a] L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. 50, 58 (1936).

* Based on work performed under the auspices of the U. S. Atomic Energy Commission.

† Submitted in partial fulfillment of the requirements for the degree of Ph.D. in Physics, University of Chicago, Chicago, Illinois.

‡ Present address: Department of Physics, Purdue University, Lafayette, Indiana.

[1] H. R. Philipp and E. A. Taft, Phys. Rev. 113, 1002 (1959).

[2] H. R. Philipp and E. A. Taft, Phys. Rev. 120, 37 (1960).

[3] F. Herman and J. Callaway, Phys. Rev. 89, 518 (1952).

[4] F. Herman, Physica 20, 801 (1954).

[5] F. Herman, Phys. Rev. 93, 1214 (1954).

[6] T. Woodruff, Phys. Rev. 103, 1159 (1956).

[7] F. Bassani, Phys. Rev. 108, 263 (1957).

[8] F. Bassani, Nuovo Cimento 13, 244 (1959).

[9] L. Kleinman and J. C. Phillips, Phys. Rev. 118, 1152 (1960).

[10] F. Bassani and M. Yoshimine, Phys. Rev. 130, 20 (1963).

[11] H. Ehrenreich, H. R. Philipp, and J. C. Phillips, Phys. Rev. Letters 8, 59 (1962).

DAVID BRUST

It should be emphasized at this point that Ehrenreich *et al.* have deduced some of the energy gaps of Table I with the aid of several natural assumptions. They assume, for example, that the main features of the optical curves can be associated with direct transitions between levels at symmetry points. One of the primary purposes of the present work is to attempt a justification of such assumptions and working hypotheses. In order to illustrate further, Ehrenreich *et al.*[11] make use of certain characteristic properties of the reflectivity, as discussed by Phillips,[12] in order to make their assignments in a definite way. For example, the edge near 2.2 eV in Ge has been resolved into a doublet[13] (see Table I). This corresponds to the expected 0.2 eV spin-orbit splitting of the $L_{3'}$ level[14] provided the optical edge is taken as arising from transitions near $L_{3'} \to L_1$. This assignment was considered as reasonable since both band theory and g-factor information[15,16] indicate an $L_{3'} \to L_1$ splitting ~ 2 eV.

With the $L_{3'} \to L_1$ gap determined in Ge it was possible to deduce it in Si.[12] This was done by extrapolating the indirect gap ($\Gamma_{25'} \to L_1$) in Ge-rich Ge-Si alloys[17,18] to its value in pure Si, and also by extrapolating the $L_{3'} \to L_1$ optical edge[19] to give a value for the $L_{3'} \to L_1$ splitting in Si.

The $\Gamma_{25'} \to \Gamma_{2'}$ splitting was found in Ge by observing the threshold for direct optical transitions.[17,20] Its value could be deduced in Si by the extrapolation of alloy data[17] and also (see Ref. 12) by cyclotron resonance measurements in strained p-type Si.[21] The last-mentioned experiments as well as optical reflectivity were used to find a value for $\Gamma_{25'} \to \Gamma_{15'}$ in Si. Again cyclotron resonance[22] and extrapolation techniques[19] were used to get $\Gamma_{25'} \to \Gamma_{15}$ in Ge.[11,12] Finally, $X_4 \to X_1$ and $L_{3'} \to L_3$ were determined by comparing band-theory results (these two gaps are both insensitive to the potential) with reflectivity data.

To sum up, the energy levels of Table I provide a model on the basis of which we can begin our detailed calculations. Hopefully we will be able to explain the frequency dependence of the optical properties. This will enable us to establish the validity of Table I and provide us with a much broader understanding of the band structure than has heretofore been possible. Furthermore, the present work in the course of checking the level assignments of Ehrenreich *et al.*[11] and Phillips[12] will confirm or redefine their diagnostic techniques.

Spin-orbit effects will not be accounted for in this paper. They are observable in Ge, as mentioned, but are $<10\%$ of the band separations. These effects, if desired, can be treated as a perturbation, and their influence on the line shape determined.

II. OPTICAL PROPERTIES

A. General Discussion

If $\epsilon_2(\omega)$ is used to denote the imaginary part of the complex dielectric constant, we have from Eq. (9) of Ref. 23

$$\epsilon_2(\omega) = \frac{4\pi^2 e^2 \hbar}{3m^2\omega^2} \sum_{n,s} \int_{B.Z.} \frac{2}{(2\pi)^3}$$

$$\times \delta(\omega_{n,s}(\mathbf{k}) - \omega) |M_{n,s}(\mathbf{k})|^2 d^3k. \quad (1)$$

The subscripts n and s refer to filled and unfilled bands, respectively, and $\omega_{n,s}(\mathbf{k}) = (E_s(\mathbf{k}) - E_n(\mathbf{k}))/\hbar$. $|M_{n,s}(\mathbf{k})|^2 = |(U_{\mathbf{k},n} | \nabla | U_{\mathbf{k},s})|^2$, where $U_{\mathbf{k},n}$ and $U_{\mathbf{k},s}$ represent the periodic parts of Bloch functions. Expression (1) neglects lifetime broadening effects such as those resulting from phonon and impurity scattering. If we ignore for a moment the matrix element in the integral (and the factor $4\pi^2 e^2\hbar/3m^2\omega^2$) then the contribution from a given pair of bands to $\epsilon_2(\omega)$ is simply

$$J_{n,s}(\omega) = \int_{B.Z.} \frac{2}{(2\pi)^3} \delta(\omega_{n,s}(\mathbf{k}) - \omega) d^3k. \quad (2)$$

The quantity $J_{n,s}(\omega)$ is the joint density of states for the two bands indexed by n and s. $J_{n,s}(\omega)\Delta\omega$ is equal to the number of pairs of states in bands n and s with $\hbar(\omega - \Delta\omega/2) \leqslant (E_s(\mathbf{k}) - E_n(\mathbf{k})) < \hbar(\omega + \Delta\omega/2)$. Later in our discussion of momentum matrix elements we shall see that $|M_{n,s}(\mathbf{k})|^2$ can, to a good approximation, be treated as a constant. This implies that a knowledge of the relevant $J_{n,s}(\omega)$ is all that one requires to understand the properties of $\epsilon_2(\omega)$.

In the Introduction it was stated that the reflectivity data were used as an aid in establishing the model term scheme of Table I. Actually the reflectance is a function of both $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$. However, by examining the figures of Refs. 1 and 2 and comparing them with $\epsilon_2(\omega)$ for Ge and Si presented in Sec. VIII of this paper one sees that $R(\omega) \sim \epsilon_2(\omega)$. All of the principal features in $|r(\omega)|^2$ are reproduced at the same energies in $\epsilon_2(\omega)$.

B. Theory of Optical Structure

At this point it is clear that to understand $\epsilon_2(\omega)$ (we shall direct our attention to the imaginary part of the dielectric constant hereafter), we must carefully examine the functions $J_{n,s}(\omega)$. Expression (2) can be

[12] J. C. Phillips, Phys. Rev. **125**, 1931 (1962).
[13] J. Tauc and A. Abraham, Phys. Chem. Solids **20**, 190 (1961).
[14] L. Liu, Phys. Rev. **126**, 1317 (1962).
[15] L. Roth and B. Lax, Phys. Rev. Letters **3**, 217 (1959).
[16] G. Feher, D. K. Wilson, and E. A. Gere, Phys. Rev. Letters **3**, 25 (1959).
[17] R. Braunstein, A. R. Moore, and F. Herman, Phys. Rev. **109**, 695 (1958).
[18] E. R. Johnson and S. M. Christian, Phys. Rev. **95**, 560 (1954).
[19] J. Tauc and A. Abraham, Phys. Rev. Letters **5**, 253 (1960).
[20] S. Zwerdling and B. Lax, Phys. Rev. **106**, 51 (1956).
[21] J. C. Hensel and G. Feher, Phys. Rev. Letters **5**, 307 (1960).
[22] G. Dresselhaus, A. F. Kip, and C. Kittel, Phys. Rev. **98**, 368. (1955).

[23] M. H. Cohen, Phil. Mag. **3**, 762 (1958).

transformed to

$$J_{n,s}(\omega) = \frac{2}{(2\pi)^3} \int_{\omega_{n,s}(\mathbf{k})=\omega} ds/|\nabla_{\mathbf{k}}\omega_{n,s}(\mathbf{k})|, \quad (3)$$

where the integral extends over the surface defined by $\omega_{n,s}(\mathbf{k})=\omega$ and ds is an infinitesimal element of area on that surface. The points $\mathbf{k}_0$ where $|\nabla_{\mathbf{k}}\omega_{n,s}(\mathbf{k})|_{\mathbf{k}=\mathbf{k}_0}=0$ are significant and are called critical points (c.p.). Van Hove[24] wrote the following expansion, valid about a normal c.p.,

$$\omega_{n,s}(\mathbf{k}) = \omega_c + \sum_{\alpha=1}^{3} a_\alpha \epsilon_\alpha \Delta k_\alpha{}^2, \quad (4)$$

where $\epsilon_\alpha = \pm 1$, $\Delta\mathbf{k} = \mathbf{k} - \mathbf{k}_0$. Here $\omega_c = \omega_{n,s}(\mathbf{k}_0)$, and a coordinate transformation which preserves d^3k has been made so as to get the functional dependence expressed by (4). He found that the critical points produce analytic singularities in what we call $J_{n,s}(\omega)$. The mathematical behavior of $J_{n,s}(\omega)$ near the "Van Hove" singularities is described by the following list:

M_0 $\epsilon_1 = \epsilon_2 = \epsilon_3 = 1$,

$$J_{n,s}(\omega) = \begin{cases} C + O(\omega-\omega_c) & \text{when } \omega < \omega_c \\ C + A(\omega-\omega_c)^{1/2} + O(\omega-\omega_c) & \text{when } \omega > \omega_c, \end{cases} \quad (5)$$

M_1 $\epsilon_1 = \epsilon_2 = -\epsilon_3 = 1$,

$$J_{n,s}(\omega) = \begin{cases} C - A(\omega_c-\omega)^{1/2} + O(\omega_c-\omega) & \text{when } \omega < \omega_c \\ C + O(\omega-\omega_c) & \text{when } \omega > \omega_c, \end{cases} \quad (6)$$

M_2 $\epsilon_1 = \epsilon_2 = -\epsilon_3 = -1$,

$$J_{n,s}(\omega) = \begin{cases} C + O(\omega-\omega_c) & \text{when } \omega < \omega_c \\ C - A(\omega-\omega_c)^{1/2} + O(\omega-\omega_c) & \text{when } \omega > \omega_c, \end{cases} \quad (7)$$

M_3 $\epsilon_1 = \epsilon_2 = \epsilon_3 = -1$,

$$J_{n,s}(\omega) = \begin{cases} C + A(\omega_c-\omega)^{1/2} + O(\omega-\omega_c) & \text{when } \omega < \omega_c \\ C + O(\omega-\omega_c) & \text{when } \omega > \omega_c. \end{cases} \quad (8)$$

In each of these expressions C and A represent constants. The symbol M_0 designates a minimum in $\omega_{n,s}(\mathbf{k})$, M_3 denotes a maximum, and M_1 and M_2 refer to saddle points. In Figs. 1(a)–1(d) we plot the behavior of $J_{n,s}(\omega)$ near the Van Hove singularities. One can already see the way in which a critical point can produce *edges* in the joint density of states and hence in $\epsilon_2(\omega)$. In Fig. 1(e) we show the possibility of two nearly degenerate critical points of the M_1 and M_2 type producing a peak-like structure in $J_{n,s}(\omega)$. Later work by Phillips[25] showed that the periodicity of the energy bands requires the existence of a minimal set of critical points. From group theoretical arguments he demonstrated that

[24] L. Van Hove, Phys. Rev. **89**, 1189 (1953).
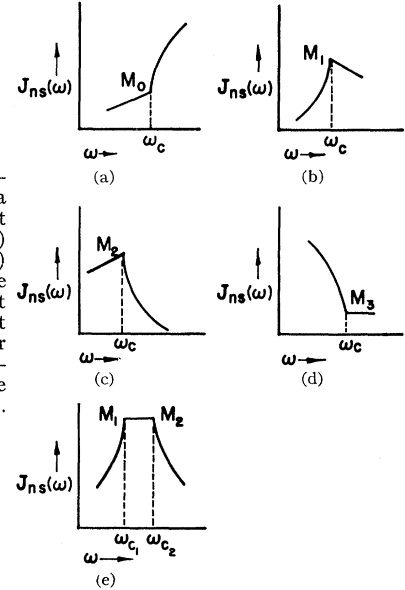[25] J. C. Phillips, Phys. Rev. **104**, 1263 (1956).

FIG. 1. Joint density of states near a normal critical point (a) M_0, (b) M_1, (c) M_2, (d) M_3, (e) a nearly degenerate pair M_1 and M_2. It should be noted that the slope of the linear portion is not determined by the nature of the critical point.

critical points occur at symmetry points. In the simplest cases (e.g., lattice vibration spectra), almost all the critical points may occur at symmetry points. In the absence of detailed information about $E(\mathbf{k})$ throughout the Brillouin zone, the empirical analysis of the previous section was confined to symmetry points. For the complicated band structures of Ge and Si this simplification (which has also been made in order to interpret the ultraviolet spectra of the noble metals) is certainly not valid. Moreover, the constants in (5)–(8) which indicate the strength of a given edge can only be obtained from values of $E(\mathbf{k})$ throughout the Brillouin zone.

C. Band-Structure Approach to Optical Properties

We now wish to do a band-structure calculation with the view of deriving $\epsilon_2(\omega)$. A description of the logical framework to be employed is the following. First a method of computing the bands will be proposed. The method should have a small number of disposable parameters. In order to determine these parameters, the term scheme discussed in Sec. I will be used as a starting point. Having set the parameters, we shall then proceed to find the bands throughout the Brillouin zone. From the band structure $\epsilon_2(\omega)$ can be computed. If the dielectric function agrees with experiment, we shall conclude that both the model of Sec. I and the computed band structure are correct. Should there be substantial disagreement then a consistent alternate to the starting term scheme will be tried, and a corrected band structure found. A schematic diagram of this is shown in Fig. 2.

Since the optical structure is associated with critical points, we must find them in the computed bands. Of course, we know that they will appear at the symmetry points, and they will have energies which correspond to
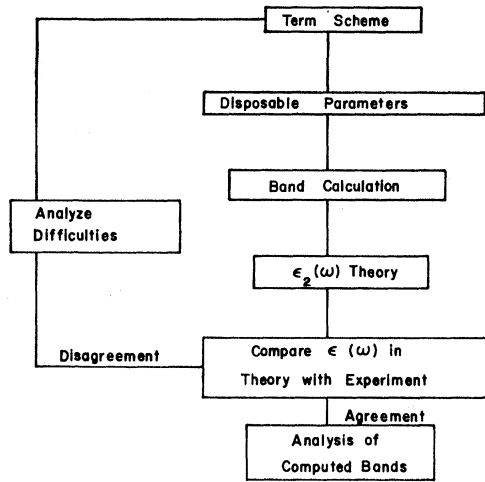
FIG. 2. Logical flow diagram.

optical edges. We cannot be sure in advance that they will be of the right analytical class or will contribute edges of the correct strength. Furthermore, there may be critical points away from the principal symmetry points, in which case they will have to be located.

III. PSEUDOPOTENTIAL METHOD

We mentioned earlier that for certain levels the OPW technique is only accurate to ~ 3 eV.[10] Furthermore, the unwieldy nature of the orthogonalization terms makes it unsuitable for extensive calculations. These considerations imply that the OPW method is unsuited for our needs.

Research workers using the OPW method noticed that the effect of orthogonalizing the plane waves to core orbitals was to greatly diminish the magnitude of the off-diagonal matrix elements in the secular equation. This led Phillips and Kleinman[26] to write the wave functions in the form

$$\psi_{n,\mathbf{k}}(\mathbf{r})=\phi_{n,\mathbf{k}}(\mathbf{r})-\sum_t \langle \theta_{t,\mathbf{k}}(\mathbf{r}) | \phi_{n,\mathbf{k}}(\mathbf{r})\rangle \theta_{t,\mathbf{k}}(\mathbf{r}). \quad (9)$$

Here $\theta_{t,\mathbf{k}}$ is the core function constructed from the tth core orbital, and $\phi_{n,\mathbf{k}}$ is a smooth wave function satisfying the wave equation

$$(P^2/2m+V_c+V_R)\phi_{n,\mathbf{k}}=E_n(\mathbf{k})\phi_{n,\mathbf{k}}, \quad (10)$$

where V_c is the crystal potential. The operator V_R has the character of a repulsive potential. Cohen and Heine[27] show that $\phi_{n,\mathbf{k}}$ and V_R are not unique. By a variational argument they demonstrated that the latter can be chosen so as to minimize

$$V_{\mathrm{eff}}=V_c+V_R, \quad (11)$$

in which case V_R has the form

$$V_R\phi_{n,\mathbf{k}}=-\sum_t(\theta_{t,\mathbf{k}}|V_c\phi_{n,\mathbf{k}})\theta_{t,\mathbf{k}}. \quad (12)$$

From (11) it is apparent that if the set of core orbitals were complete, perfect cancellation of $(V_c\phi_{n,\mathbf{k}})$ by $(V_R\phi_{n,\mathbf{k}})$ would result (the bands would then be completely free electron). The core orbitals form a surprisingly good basis set in the core region. This implies that the Fourier coefficients of $V_{\mathrm{eff}}{}^{Gi}=V_c{}^{Gi}+V_R{}^{Gi}$ are small for large values of $\mathbf{G}_i=(a/2\pi)\mathbf{K}_i$ where $\mathbf{K}_i$ is a reciprocal lattice vector.[28] This suggests introducing a pseudopotential with Fourier coefficients $V_p{}^{Gi}\simeq V_{\mathrm{eff}}{}^{Gi}$ which are zero for $|\mathbf{G}_i|$ greater than the first two or three values.[29] A pseudopotential of this form was first introduced by Phillips.[30] He, however, chose $V_p{}^{Gi}$ = const for large $\mathbf{G}_i$. Bassani and Celli,[31] on the other hand, took $V_p{}^{Gi}=0$ for $\mathbf{G}_i$ such that $|\mathbf{G}_i|^2>11$ which materially improved the convergence in agreement with the variational arguments of Cohen and Heine.

In this paper we shall adopt a pseudopotential having the form used by Bassani and Celli. For both Ge and Si this will be of the form

$$V_p=\sum_j V_p{}^{Gj}\exp[(2\pi i/a)\mathbf{G}_j\cdot\mathbf{r}],$$

$$V_p{}^{Gj}=V_p(3)\cos\left[(\mathbf{G}_{j,1}+\mathbf{G}_{j,2}+\mathbf{G}_{j,3})\frac{\pi}{4}\right] \quad \text{if} \quad |\mathbf{G}_j|^2=3,$$

$$=V_p(8)\cos\left[(\mathbf{G}_{j,1}+\mathbf{G}_{j,2}+\mathbf{G}_{j,3})\frac{\pi}{4}\right] \quad \text{if} \quad |\mathbf{G}_j|^2=8,$$

$$=V_p(11)\cos\left[(\mathbf{G}_{j,1}+\mathbf{G}_{j,2}+\mathbf{G}_{j,3})\frac{\pi}{4}\right] \quad \text{if} \quad |\mathbf{G}_j|^2=11,$$

$$=0 \quad \text{if} \quad |\mathbf{G}_j|^2>11.$$

Here $\mathbf{G}_{j,1}$ is the projection of $\mathbf{G}_j$ on the $(1,0,0)$ direction (direction normal to a square face). There are then three adjustable parameters $V_p(3)$, $V_p(8)$, and $V_p(11)$. We should point out that V_{eff} is both l and $\mathbf{k}$ dependent. That is, it is different depending on whether $\phi_{n,\mathbf{k}}$ has s or p symmetry, and also varies throughout the Brillouin zone. This nonlocal character is a consequence of the OPW method. We notice, however, that V_p is a local (function of $\mathbf{r}$ only) potential. This simplifying feature cannot be regarded as deleterious. The energy eigenvalues which result from using V_p are generally within 2 eV of those deriving from the use of V_{eff}. The latter, one remembers, gives energies certain to only 3 eV furthermore, V_p will be adjusted at the outset to agree with what we believe to be a large number of symmetry term values. We would then expect V_p to give eigenvalues at other points of the zone with an error much

[26] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959).
[27] M. H. Cohen and V. Heine, Phys. Rev. **122**, 1821 (1961); also, see Ref. 30.

[28] In these units the center of the hexagonal face is located at $(\frac{1}{2},\frac{1}{2},\frac{1}{2})$ and the shortest nonzero $\mathbf{G}_i$ has length $=\sqrt{3}$.
[29] We are here distinguishing between V_{eff} which is uniquely defined by (7) and (8) and V_p which is an approximation to V_{eff}. Inserting V_{eff} into (6) gives exactly the same results as the OPW method from which it is derived, whereas V_p does not.
[30] J. C. Phillips, Phys. Rev. **112**, 685 (1958).
[31] F. Bassani and V. Celli, Phys. Chem. Solids **20**, 64 (1961).

less than the OPW method. The pseudopotential in a sense circumvents the task of finding a crystal potential replete with a correct exchange interaction, Coulomb interaction, etc. It fits, instead, the most important terms in the potential to experiment.

One more point should be made. Bassani and Celli do distinguish between s- and p-like states by allowing $V_s(0)$ to differ from $V_p(0)$. In this way they can raise s-like states with respect to p-like states. We have found that this is not necessary to get an adequate fit to the experimental interpretation of the levels.

IV. DIAGONALIZATION PROCEDURE

We seek to determine the energy eigenvalues of our model wave equation

$$H_p\phi_{n,k}=E_{n,k}\phi_{n,k}, \tag{13}$$

where $H_p=(-\hbar^2/2m)\nabla^2+V_p$. To do this we must in principle solve the infinite secular equation

$$|H_p{}^{i,i}-E\delta_{i,j}|=0. \tag{14}$$

Here $H_p{}^{i,i}=\langle\phi_{k+K_i}|H_p|\phi_{k+K_j}\rangle$ and $\phi_{k+K_i}=e^{i(k+K_i)\cdot r}$. That is, we use a plane-wave representation in which to expand the $\phi_{n,k}$. The lowest four levels will represent valence states and the higher one's conduction states (one remembers that the core levels have been eliminated from the problem by the orthogonalization procedure). Only the low-lying conduction levels are of interest so that we shall ask for only the first eight eigenvalues of H_p (four valence and four conduction).

Assuming that the ϕ_{k+K_i} have been ordered so that for $i>j$, $|k+K_i|^2\geqslant|k+K_j|^2$, we may truncate the secular Eq. (14) so as to include contributions from only those plane waves for which i, $j\leqslant N$ (N some integer). We expect the low-lying levels to converge to their final values as N is made large. If we take as our convergence criterion that $|E_{n,k}{}^{N=N_0}-E_{n,k}{}^{N=\infty}|<0.1$ eV then N_0 would have to be ~50. The time required by the IBM 704 to diagonalize a 50×50 matrix at a large number of points was judged as too great to make this approach feasible. By augmenting the calculation with perturbation theory it is possible, however, to reduce greatly the computing time.

A form of perturbation theory, e.g., Löwdin[32] was used instead of the Rayleigh-Schrödinger method. We seek the eigenvalues of the secular equation

$$|U_p{}^{n,m}-E\delta_{n,m}|=0, \tag{15}$$

$$U_p{}^{n,m}=H_p{}^{n,m}+\sum_{\gamma=N+1}^{\Gamma}\frac{H_p{}^{n,\gamma}H_p{}^{\gamma,m}}{E-H_p{}^{\gamma,\gamma}}, \tag{16}$$

with n, $m\leqslant N$ and $\gamma>N$ but $\leqslant\Gamma$ where N and Γ are integers. The indexes n and m refer to plane waves being treated exactly; whereas Γ refers to higher plane waves being treated only through perturbation theory. This

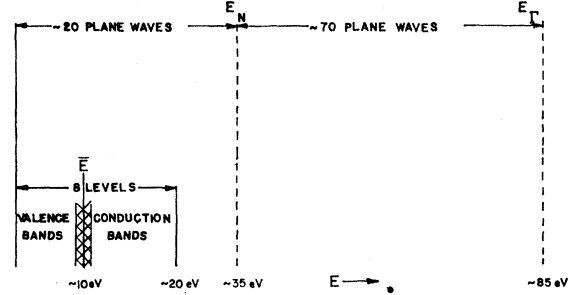[32] P. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

FIG. 3. The approximate relationship of the quantities entering the band calculation. $\bar{E}$ represents an average energy for the valence and conduction bands. E_N is a cutoff energy such that all plane waves with kinetic energy $(\hbar^2/2m)|k+K_i|^2\leqslant E_N$ are treated exactly, whereas those having a kinetic energy between E_N and E_Γ are accounted for by perturbation theory.

procedure is especially convenient for handling degenerate and quasidegenerate cases as it is not necessary to find linear combinations of zero-order degenerate states as in the standard method. This is particularly advantageous when automatic computing machinery is employed.

Before proceeding it was necessary to eliminate the eigenvalue dependence of the matrix elements in (15). If this were not done, we should have to solve for each of the eigenvalues separately and to use an iteration procedure for each of them. We wished, instead, to diagonalize (15), and obtain all eight interesting levels at once.

In handling the matrix elements the following substitutions were made: off the diagonal $\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,m})/(E-H_p{}^{\gamma,\gamma})\rightarrow\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,m})/(E-H_p{}^{\gamma,\gamma})$, that is, $E\rightarrow\bar{E}$. Here $\bar{E}$ is an average of the eight energy levels at each point in the reduced zone. On the diagonal $\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,n})/(E-H_p{}^{\gamma,\gamma})\rightarrow\sum_\gamma(H_p{}^{n,\gamma}H_p{}^{\gamma,n})/(H_p{}^{n,n}-H_p{}^{\gamma,\gamma})$, i.e., $E\rightarrow H_p{}^{n,n}=(\hbar^2/2m)|k+K_n|^2$.
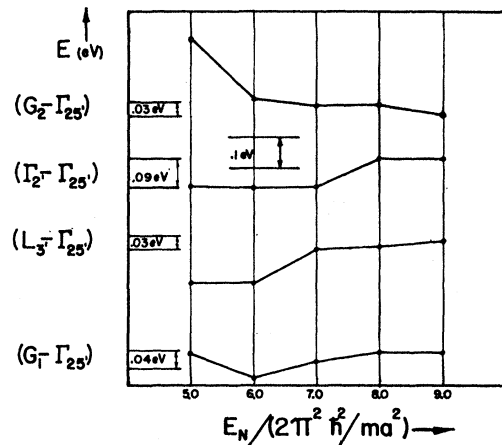

FIG. 4. Shows the convergence properties of the worst cases among those tested. G_1 and G_2 refer to states at the general point of the zone with $k=(0.60,0.35,0.10)$. The energies (in eV) at the left are an estimate of the convergence for $E_N=7.0$ (the value actually used). They were gotten by comparing with the results for $E_N=9.0$.
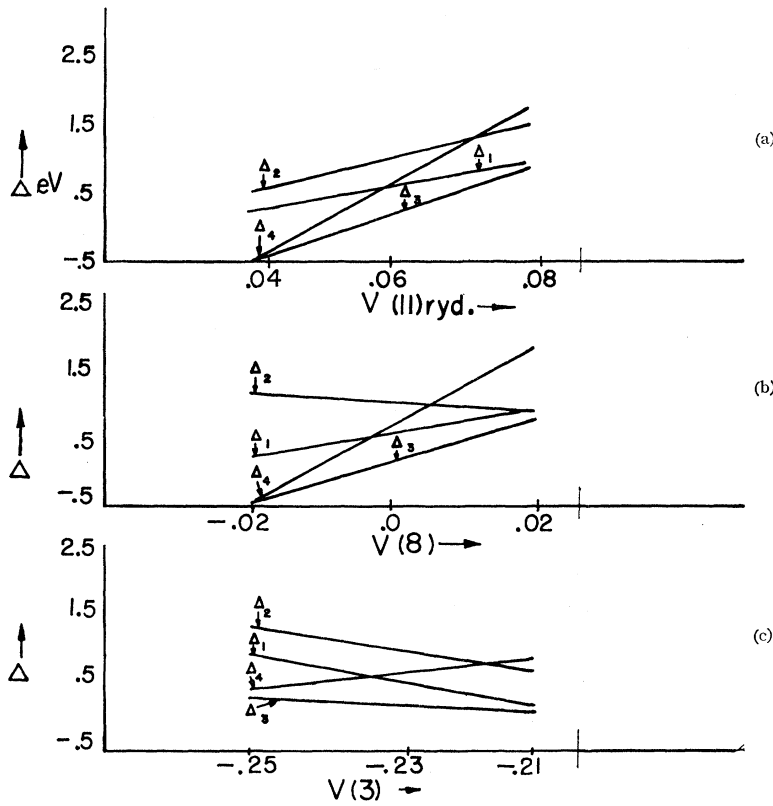
FIG. 5. Variation of some of the principal band gaps with the potential coefficients. In eV: $\Delta_1 = (X_1 - X_4) - 3.0$, $\Delta_2 = (L_3 - L_{3'}) - 4.5$, $\Delta_3 = (L_1 - L_{3'}) - 1.5$, $\Delta_4 = (\Gamma_{2'} - \Gamma_{25'})$; (a) $V(3) = -0.21$ Ry, $V(8) = 0.0$ Ry; (b) $V(3) = -0.23$ Ry, $V(11) = 0.06$ Ry; (c) $V(8) = 0.0$ Ry, $V(11) = 0.06$ Ry.

We here, in essence, replace E by the kinetic energy of the principal plane waves making up the expansion of $\phi_{n,k}$. This allows a closer representation of E than any one choice of $\bar{E}$ permits. It was felt that the more accurate representation of the diagonal matrix elements would improve the convergence.

The error $\Delta U_p{}^{n,m}$ in a matrix element resulting from these approximations is

$$\Delta U_p{}^{n,m} \approx \sum_\gamma \frac{H_p{}^{n,\gamma} H_p{}^{\gamma,m}}{(E - H_{\gamma,\gamma}{}^p)^2} \delta E,$$

$$\delta E = \bar{E} - E, \quad n \neq m,$$

$$\delta E = H_p{}^{n,m} - E, \quad n = m.$$

(17)

Since E refers to the energy of one of the lowest eight eigenvalues, one sees that by increasing N and hence the minimum value of $H_p{}^{\gamma,\gamma}$ that $\Delta U_p{}^{n,m} \to 0$.

Now suppose that the plane waves ϕ_{k+K_n} being treated exactly are those obeying the condition $(\hbar^2/2m)|k+K_n|^2 \leqslant E_N$. Here E_N is a cutoff energy which we shall choose below. By this definition N will be somewhat dependent on k. In a like manner the plane waves ϕ_{k+K_γ} entering the calculation through perturbation theory satisfy $E_N < (\hbar^2/2m)|k+K_\gamma|^2 \leqslant E_\Gamma$.

The actual choice of E_N, E_Γ, and $\bar{E}$ depend on an intelligent use of trial and error. The values ultimately

chosen were

$$E_N = 7.0(2\pi^2\hbar^2/ma^2) \sim 35 \text{ eV},$$

$$E_\Gamma = 19.0(2\pi^2\hbar^2/ma^2) \sim 85 \text{ eV},$$

$$\bar{E} = 2.0(2\pi^2\hbar^2/ma^2) \sim 10 \text{ eV}.$$

With this choice N was ~ 20 and $\Gamma \sim 90$. In Fig. 3 we show these values in comparison with the band energies.

The question of convergence was examined by observing the change in the eigenvalues as E_N was increased. In Fig. 4 the levels showing the poorest convergence among the test cases are plotted as a function of E_N. One should note that convergence is good to ~ 0.05 eV which is within the limits of the accuracy being sought.

V. PSEUDOPOTENTIAL PARAMETERS

We now turn to a brief discussion concerning the choice of the pseudopotential coefficients. The work of Bassani and Celli[31] on Ge was found to be a convenient starting point. Since the present work makes no distinction between s and p symmetry, it was necessary to redefine their potential slightly. In determining the Fourier coefficients particular emphasis was placed on getting good agreement with the proposed interpretation of the optical experiments. That means we sought to make the direct splittings such as $L_{3'} \to L_3$, $L_{3'} \to L_1$, etc., come as close as possible to the experimental

TABLE II. Energy of principal gaps (in eV).

Gap	Bassani (Ge)	Brust (Ge)	Brust (Si)
$\Gamma_{25'} \rightarrow \Gamma_{2'}$	0.8	0.6	3.8
$\Gamma_{25'} \rightarrow \Gamma_{15}$	3.4	3.6	3.4
$L_{3'} \rightarrow L_1$	2.0	1.8	3.1
$L_{3'} \rightarrow L_3$	5.4	5.4	5.4
$X_4 \rightarrow X_1$	4.4	3.6	4.0

identification. Since at our first trial we sought to test the proposed energy model by attempting to get the correct structure in the optical properties, it was natural to make the splittings agree as closely as possible with the model.

In order to reduce the labor required to get a good set of pseudopotential parameters the graph shown in Fig. 5 was prepared. This indicates how the principal band gaps vary with changes in the potential coefficients. By using this a good fit could be obtained with only one or two trials for both Ge and Si. In this way we found for Ge: $V(3) = -0.23$ Ry, $V(8) = 0.0$ Ry, and $V(11) = 0.06$ Ry; while for Si: $V(3) = -0.21$ Ry, $V(8) = 0.04$ Ry, and $V(11) = 0.08$ Ry. The term values resulting from these choices are given in Table II where our Ge results are compared with those of Bassani and Celli.

The values of Table II are somewhat different from those of Table I. It is not possible to get an exact fit with a three-parameter potential. If one required a better set of values, additional coefficients would have to be introduced into the potential.

It was of interest to compute the energy bands along symmetry directions. These are shown in Fig. 6, and should be compared with earlier results obtained by Herman on Ge and by Kleinman and Phillips for Si. We should point out that the good behavior of our bands was used as an indication of both the convergence of our computational procedure and of the correctness of the pseudopotential approach.

We are now in a position to go on to a calculation of $\epsilon_2(\omega)$. Before doing this, however, we shall examine the result of computing effective masses according to the present scheme.

VI. EFFECTIVE MASSES

We wish to determine the effective masses of the conduction electrons within the framework of the pseudopotential method. Doing this for Ge by the formulas of $\mathbf{k} \cdot \mathbf{p}$ perturbation theory we have for Ge, e.g., Liu[14]

$$\left(\frac{m_e}{m_t^*}\right) = 1 + \frac{2}{m} \frac{|\langle L_1 | P_{y'} | L_{3'}{}^{y'}\rangle|^2}{(E_{L_1} - E_{L_{3'}})}. \quad (18)$$

Here $P_{y'}$ is a momentum operator in a direction transverse to the (1,1,1) direction. Using our pseudopotential
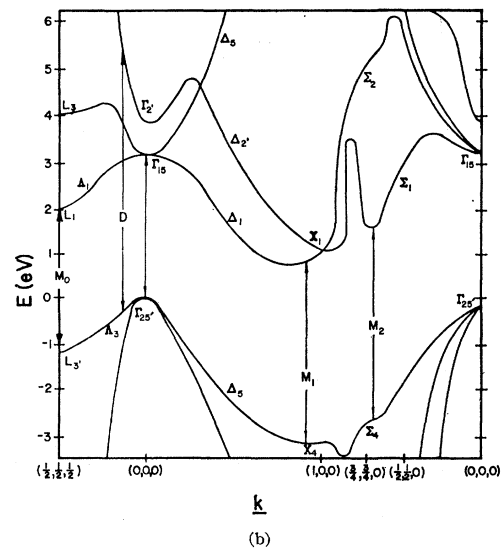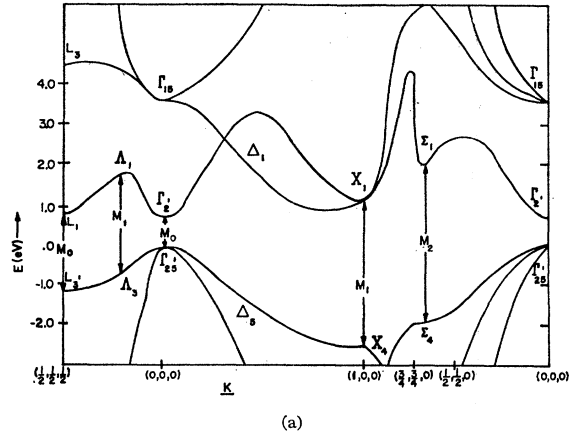

(a)

(b)

FIG. 6. Pseudopotential energy bands along Δ, Λ, and Σ symmetry directions. Some of the principal transitions have been marked.

wave functions the momentum matrix element was found to be $0.73\hbar^2(2\pi/a)^2$. The energy denominator is 1.8 eV. Putting these values into (18) one finds $m_t^* = 0.12m_e$ compared with the experimental value of $0.082m_e$. Considering the approximations, used to get m_t^* the result is reasonably close to experiment. This value was checked by directly computing $E(\mathbf{k})$ in a transverse direction and using

$$m_e/(m_t^*) = [2m_e(E(\Delta\mathbf{k} + \mathbf{k}_L) - E(\mathbf{k}_L))]/\hbar^2(\Delta\mathbf{k})^2.$$

This also gave us $0.12m_e$, which justifies the use of a two-band model in the $\mathbf{k} \cdot \mathbf{p}$ calculation.

For Si where the minima correspond to Δ_1 states, we have from Dresselhaus[33]

$$\left(\frac{m_e}{m_t^*}\right) = 1 + \sum_1 \frac{|\langle \Delta_1 | P_y | \Delta_5{}^i\rangle|^2}{(E_{\Delta_1} - E_{\Delta_5}{}^i)}. \quad (19)$$

[33] G. F. Dresselhaus, Ph.D. thesis, University of California (unpublished).
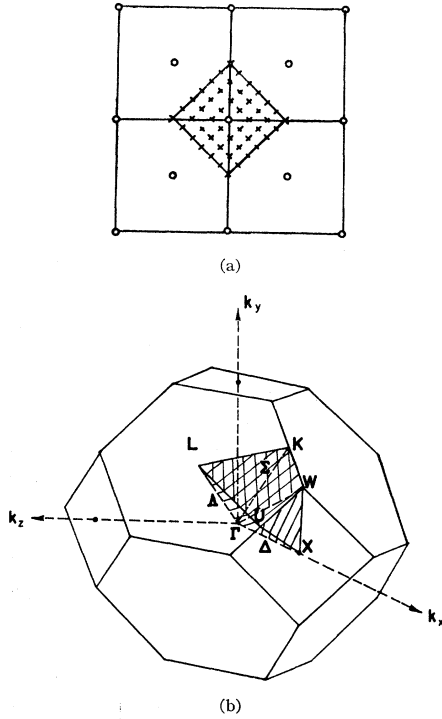
(a)

(b)

FIG. 7. Scheme for sampling zone (a) mesh of sampling points in two dimensions, (b) section of zone actually sampled.

In this case we must use a 3-band model taking two Δ_5 states (one valence and one conduction). We now get $m_t^* = 0.17m_e$ compared with $0.19m_e$ from experiment. If only the Δ_5 state in the valence band is used, we get $m_t^* = 0.14m_e$. By examining $E(\mathbf{k})m_t^* = 0.17m_e$ was found, showing that a 3-band $\mathbf{k} \cdot \mathbf{p}$ calculation is adequate for our purposes.

VII. SAMPLING PROCEDURES

With a view toward finding $\epsilon_2(\omega)$ we discuss the calculation of the joint density-of-states functions $J_{ns}(\omega)$ defined by Eq. (2). Since in practice it is possible to solve for the eigenvalues of H_p at only a finite number of points in the reduced zone, we must replace the integral in (2) by a finite sum. This was done by defining the quantity

$$K_{ns}(\omega_i) = \sum_{\mathbf{k}_l} \delta^{\Delta\omega}(\omega_{ns}(\mathbf{k}_l) - \omega_i). \qquad (20)$$

Here $\mathbf{k}_l$ is a set of uniformly spaced sampling points lying within the first Brillouin zone. The delta function is defined by

$$\delta^{\Delta\omega}(\omega_{ns} - \omega_0) = 1 \quad \text{if} \quad |\omega_{ns} - \omega_0| \leqslant (\Delta\omega/2) = 0 \text{ otherwise}. \qquad (21)$$

The function $K_{ns}(\omega_i)$ is then defined for a set of values ω_i such that $\omega_{i+1} = \omega_i + \Delta\omega$. The meaning of $K_{ns}(\omega_i)$ is clear. It measures the number of pairs of states in bands n and s with direct energy gap $\hbar\omega_{ns} = \hbar\omega_i \pm \hbar(\Delta\omega/2)$ as

found in the sample. The relationship between $K_{n,s}(\omega_i)$ and $J_{ns}(\omega)$ is

$$K_{ns}(\omega_i) = \frac{\Delta\omega}{\Delta^3 k} \frac{(2\pi)^3}{2} J_{ns}(\omega = \omega_i), \qquad (22)$$

where $\Delta^3 k$ is the elementary volume surrounding the sampling points.

In making a choice for $\Delta\omega$ we were guided by the consideration that most of the experimental structure in $\epsilon_2(\omega)$ has a half-peak width $\sim\frac{1}{2}$ eV. We consequently took $\Delta\omega = 0.1$ eV$/\hbar$ as a reasonable value.

The reciprocal lattice vectors of the diamond structure form a bcc lattice. The most natural choice for the sampling vectors $\mathbf{k}_l$ is a mesh of points forming a bcc array. This was reduced in size from the host lattice by a scaling factor of 36. That is, the $\mathbf{k}_l$ are determined by $\mathbf{k}_l = \mathbf{K}_l/36$ where $\mathbf{K}_l$ are reciprocal lattice vectors. This idea is demonstrated for a two-dimensional body centered structure in Fig. 7(a). With this reduction in scale the calculation gives results for $36^3 = 46\,656$ $\sim 50\,000$ points in the reduced zone.

Advantage was taken of the 48-fold symmetry of the Brillouin zone of the diamond structure. It is then necessary to examine only that 1/48th part of the zone indicated in Fig. 7(b). This region is defined by the conditions

$$0 \leqslant (\mathbf{k}_l)_z \leqslant (\mathbf{k}_l)_y \leqslant (\mathbf{k}_l)_x, \qquad (23)$$

and $\mathbf{k}_l$ contained within the first Brillouin zone. The contribution to $K_{ns}(\omega_i)$ from a point $\mathbf{k}_l$, in the interior of the sampling region is then taken to be 48. The number 48 is referred to as the weighting factor W of the point $\mathbf{k}_l$.

For points not in the interior of the sampling region W is easily found. We simply find the number of regions between which the point $\mathbf{k}_l$ is shared. If we call this number C, then $W = 48/C$. For points not on the Brillouin zone surface C is found as follows:

$$C = \begin{Bmatrix} 1 \\ = 2 \end{Bmatrix}, \quad \text{if} \quad (\mathbf{k}_l)_x \neq (\mathbf{k}_l)_y \neq (\mathbf{k}_l)_z \begin{matrix} \neq 0 \\ = 0 \end{matrix},$$

$$C = \begin{Bmatrix} 2 \\ = 8 \end{Bmatrix}, \quad \text{if} \quad (\mathbf{k}_l)_x \neq (\mathbf{k}_l)_y = (\mathbf{k}_l)_z \begin{matrix} \neq 0 \\ = 0 \end{matrix},$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (24)$$

$$C = \begin{Bmatrix} 2 \\ = 4 \end{Bmatrix}, \quad \text{if} \quad (\mathbf{k}_l)_x = (\mathbf{k}_l)_y \neq (\mathbf{k}_l)_z \begin{matrix} \neq 0 \\ = 0 \end{matrix},$$

$$C = \begin{Bmatrix} 6 \\ = 48 \end{Bmatrix}, \quad \text{if} \quad (\mathbf{k}_l)_x = (\mathbf{k}_l)_y = (\mathbf{k}_l)_z \begin{matrix} \neq 0 \\ = 0 \end{matrix}.$$

If $\mathbf{k}_l$ lies on the first Brillouin zone surface then it is shared with adjoining zones, and the above value of C is multiplied by 2, 3, or 4 according to whether $\mathbf{k}_l$ is on a face, edge, or corner. For example, for the point Γ, $C = 48$ and $W = 1$. With this technique H_p had only to be diagonalized at ~ 1200 points.

When the data first came from the computer, it was observed that the scatter in the 0.1-eV histograms, $K_{ns}(\omega_i)$, was excessive. We found, however, that by smoothing $K_{ns}(\omega_i)$ according to

$$\bar{K}_{ns}(\omega_i) = \tfrac{1}{3}[K_{ns}(\omega_{i+1}) + K_{ns}(\omega_i) + K_{ns}(\omega_{i-1})]; \quad (25)$$

the scatter was materially reduced. This method is discussed by Hartree.[34] When the smoothing procedure is combined with the results of critical point analysis as in the next section, excellent results are achieved. In the Appendix we examine the effects of smoothing as well as the sources of the statistical noise.

VIII. DISCUSSION OF LINE SHAPES

In discussing the results we shall begin by examining the histogram for bands 4 and 5 (the bands are ordered

(a)

(b)
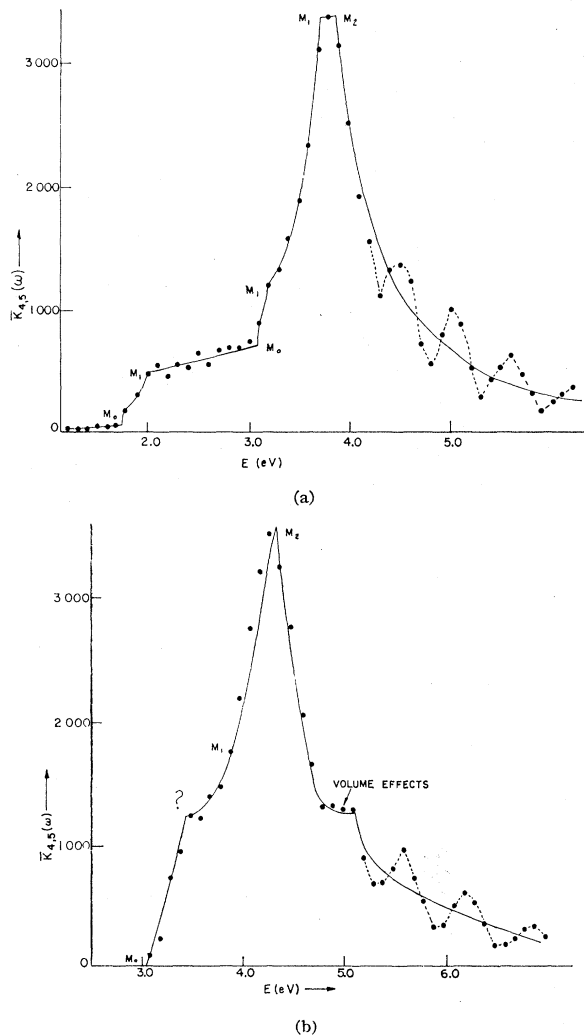
FIG. 8. 4–5 histograms. Volume effects refer to structure which could not be traced to critical points. (a) Ge, (b) Si.

[34] D. R. Hartree, *Numerical Analysis* (Oxford University Press, London, 1952), pp. 249–251.
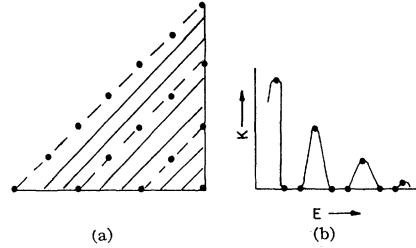
FIG. 9. (a) Sampling planes (dashed lines) are parallel to energy contours (solid lines); (b) resulting density of states showing the spurious peaks.

in the conventional way with $n=4$ the highest valence band, and $s=5$ the lowest conduction band). From energy considerations it can be guessed that this will make the most important contribution to $\epsilon_2(\omega)$. Figure 8 shows $\bar{K}_{4,5}(\omega)$ for both Ge and Si. One notices that the graphs show a group of periodic peaks (dashed lines) beginning at 4.5 eV in Ge and at 5.6 eV in Si. A close study of the data disclosed that these arise in an artificial way as boundary sampling errors. Near the point W one finds that the planes of sampling points lie parallel to the surfaces on which $E_{4,5}(\mathbf{k})$ is constant. This condition is drawn for a two-dimensional example in Fig. 9, where one easily sees how these artificial sampling peaks can be formed. Since they were judged to be spurious, the peaks have been eliminated from the joint density of states (solid lines in Fig. 8).

The most striking result is the large central peak. We remember that in our starting model the large optical peak was associated with transitions of the type $X_4 \rightarrow X_1$. To understand how it arises in $\bar{K}_{4,5}$ we inspect Fig. 10 where $E_{4,5}(\mathbf{k})$ is plotted in the basal plane (ΓKX plane). By looking at the energy contours near the point X in Fig. 10(b), one discovers that there exists a critical point of the M_1 class with a direct energy gap of 3.6 eV. Furthermore, on the Σ axis near the point $\mathbf{k} = (0.6,0.6,0.0)$ there occurs another critical point. This belongs to the M_2 class, and has an energy gap of 3.8 eV. The large peak in $\bar{K}_{4,5}$ of Ge, results then from these two nearly degenerate critical points corresponding to $X_4 \rightarrow X_1$ and $\Sigma_4 \rightarrow \Sigma_1$ transitions. This idea has been demonstrated previously in Fig. 1(e). The occurrence of a critical point on Σ is a new result.

The analysis of Fig. 10(b) for Si leads to the same general conclusions. At the point X there again occurs an M_1 critical point with energy of 4.0 eV. On the Σ axis near $\mathbf{k} = (0.4,0.4,0.0)$ we find an M_2 critical point this time with an energy gap of 4.4 eV.

In Fig. 8 the reader may notice that sharp corners have been drawn into the data. The analysis discussed earlier indicates that this is the correct behavior near a critical point.

Returning to Fig. 8(a), we see a sharp corner of the M_1 type in $\bar{K}_{4,5}$ of Ge at 2.0 eV. It is clear that this is identifiable with the spin-orbit split optical edge at 2.2 eV. Figure 11(a) shows the contours of $E_{4,5}(\mathbf{k})$ in the
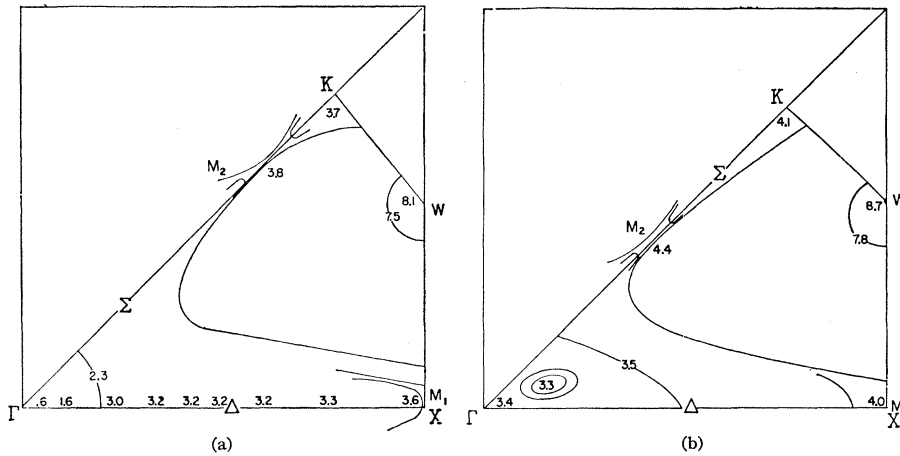
FIG. 10. $E_{4,5}(\mathbf{k})$ contours in ΓKX plane. The energies indicated on graph are all in eV. A few of the principal contours are drawn showing behavior near critical points. (a) Ge, (b) Si.

plane defined by ΓKL. At L there is an M_0 critical point with a gap of 1.8 eV. Looking along Λ, we detect another critical point near $\mathbf{k}=(0.17,0.17,0.17)$. This one is classified as an M_1 with an energy of 2.0 eV. The edge results from $\Lambda_3 \rightarrow \Lambda_1$ transitions, while the threshold is produced by $L_{3'} \rightarrow L_1$ transitions. This interpretation is again a new one.

Both $\Lambda_3 \rightarrow \Lambda_1$ and $L_{3'} \rightarrow L_1$ are spin-orbit split by approximately the same amount. Although the predicted breadth of the 2-eV edge given by the separation of the M_0 and M_1 singularities (0.2 eV) is in good agreement with experiment, more decisive confirmation of the predicted M_0 threshold can be obtained in materials where the spin-orbit splitting and the M_1 peak are somewhat larger. This is the case for GaAs[35] and still more clearly ZnTe.[36] The M_0 thresholds are well resolved approximately 0.2 eV below the M_1 edges. It is striking that such fine details appear to carry over not only to slightly ionic crystals such as GaAs but also to strongly ionic ones such as ZnTe.

In the 4–5 histogram of Si there is an edge at 3.5 eV. This is identified with the peak at 3.5 eV in $\epsilon_2(\omega)$. Figure 11(b) shows that the threshold for this edge is once more an M_0 critical point at L. The energy of the $L_{3'} \rightarrow L_1$ gap is here 3.1 eV (in Si this appears to be the

minimum vertical separation between the valence and conduction bands). The edge itself is a bit more difficult to classify. The energy contours indicate that it comes from points near the zone center. There appear to be several critical points clustered about Γ degenerate to within 0.04 eV. Our convergence, however, is not reliable for such small energy differences. We prefer to say simply that the edge is due to a critical point at $(\Gamma_{25'} \rightarrow \Gamma_{15})$ with an energy gap of 3.5 eV.

Finally returning to the Ge histogram of Fig. 8(a), we notice a shoulder near 3.2 eV. This is presumably caused by a point of inflection in $E_{4,5}(\mathbf{k})$ near the point $\mathbf{k}=(0.4,0,0)$. This has the same effect as a weak pair of critical points of the M_0 and M_1 category. We mention this as a possible explanation of a similar shoulder in $\epsilon_2(\omega)$ of Ge near the same energy. In our data it is less pronounced than in experiment owing to the nearness of the large peak which masks the shoulder.

That exhausts all of the interesting structure in the 4–5 histograms. We wish to check that our critical-point analysis is complete. To do this we use a rule discussed by Phillips,[25] which says that

$$N(M_0)-N(M_1)+N(M_2)-N(M_3)=0. \qquad (26)$$

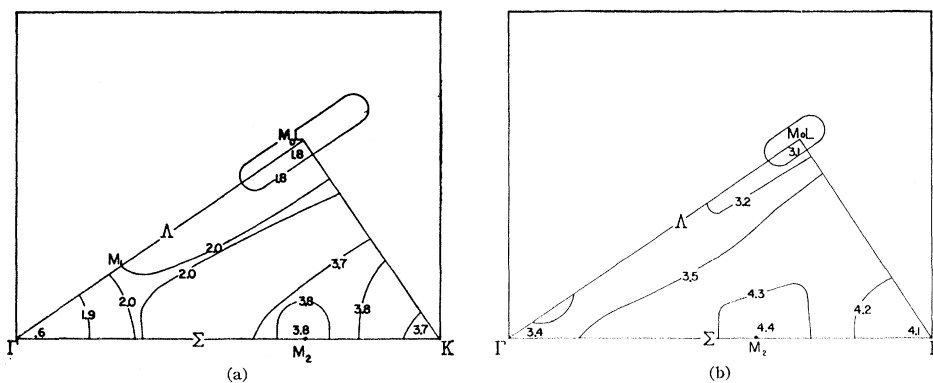Here $N(M_0)=$ number of M_0 critical points; $N(M_1)$

FIG. 11. $E_{4,5}(\mathbf{k})$ in ΓLK plane. (a) Ge, (b) Si.

[35] D. L. Greenaway, Phys. Rev. Letters 9, 97 (1962).
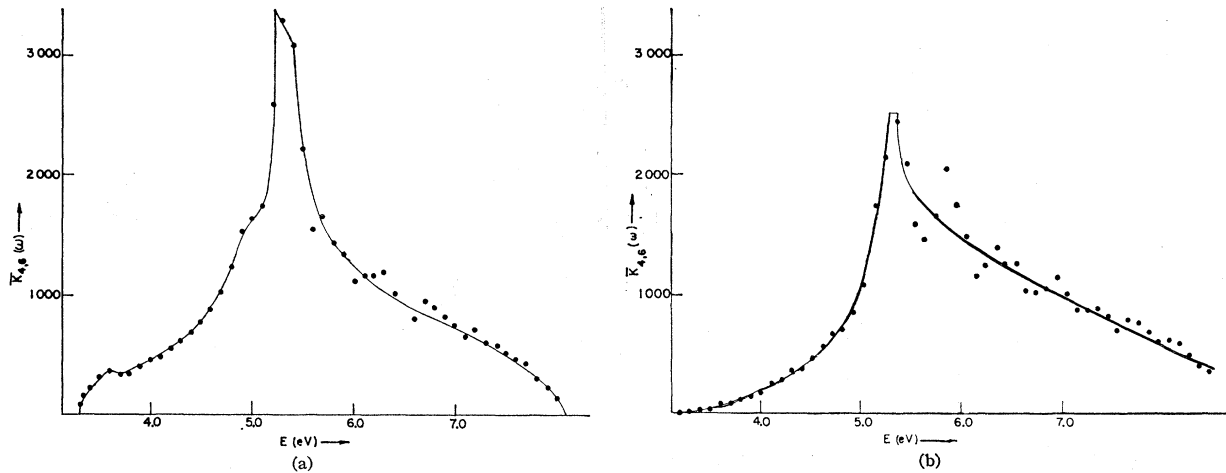[36] M. Cardona and D. L. Greenaway, Phys. Rev. 131, 98 (1963).

FIG. 12. 4–6 histogram. (a) Ge, (b) Si.

=number of M_1 critical points; etc. In Ge, $N(M_0)$ has contributions from three sources: Γ (the absolute minimum $\Gamma_{25'} \rightarrow \Gamma_{2'}$), L, and Δ (the M_0 contribution of the kink). Putting in the weighting factors (W) for each of the points we have $N(M_0)=1+4+6=11$. For $N(M_1)$ the contributions are from Λ, X, and Δ (the M_1 saddle point arising from the kink). This gives $N(M_1)=8+3+6=17$. The contributions to $N(M_2)$ are all from Σ and $N(M_2)=12$. The contributions to M_3 all come from the maximum in $E_{4,5}(\mathbf{k})$ which occurs at W and $N(M_3)=6$. Forming the sum implied by (26) we get $11-17+12-6=0$ as required. We have not applied the rule to Si because of the complexities discussed earlier. We can nevertheless, be confident that all the important critical points have been found since the energy contours in Ge and Si are very similar.

Next we examine the contribution of the 4–6 histogram. In Fig. 12, $\bar{K}_{4,6}(\omega)$ is drawn. The main feature in the Ge data [Fig. 12(a)] is the large peak at 5.3 eV. From the energy contours of Fig. 13(a), one can see that the structure is associated with L. In this case the transition being $L_{3'} \rightarrow L_3$ which agrees with the interpretation of the corresponding optical peak.

The Si data are understood similarly. The $L_{3'} \rightarrow L_3$ transition is responsible for the peak at 5.3 eV. The somewhat large peak at 5.8 eV is apparently a boundary sampling error of the type discussed in connection with the 4–5 histogram. After matrix elements are accounted for, the remaining structure in $\bar{K}_{4,6}$ is too weak to make an observable contribution to the optical properties. For the same reason the remaining histograms contribute nothing to $\epsilon_2(\omega)$.

We are now in a position to use these results to construct $\epsilon_2(\omega)$. First we shall give a very brief discussion of how the matrix elements were determined.

In order to develop a reasonable way of handling $|M_{n,s}(\mathbf{k})|^2$ we computed matrix elements for $n=4$, $s=5$, and for $n=4$, $s=6$, at a number of points in the reduced zone. These were gotten by using the pseudopotential wave functions $\phi_{n,\mathbf{k}}$. Such a procedure corresponds to using the smooth part of the OPW wave functions which according to Ref. 37 gives accuracy of better than 20%.[38] The results are shown in Table III. The Si wave functions were used; however, we would get almost identical results with the Ge wave functions. In obtaining the matrix elements for the symmetry

FIG. 13. $E_{4,6}(\mathbf{k})$ in ΓLK plane. (a) Ge, (b) Si.
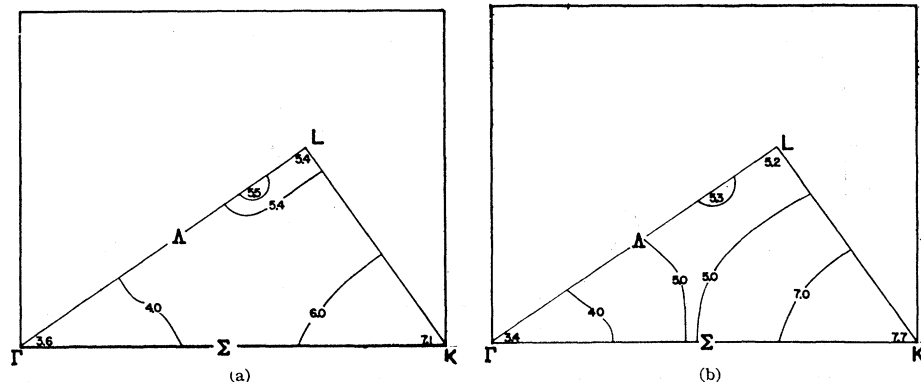
[37] L. Kleinman and J. C. Phillips, Phys. Rev. 118, 1153 (1960).
[38] The effective masses are determined by momentum matrix elements. Since our computed values for the effective masses agree well with experiment, we naturally expect the computed matrix elements also to be correct.
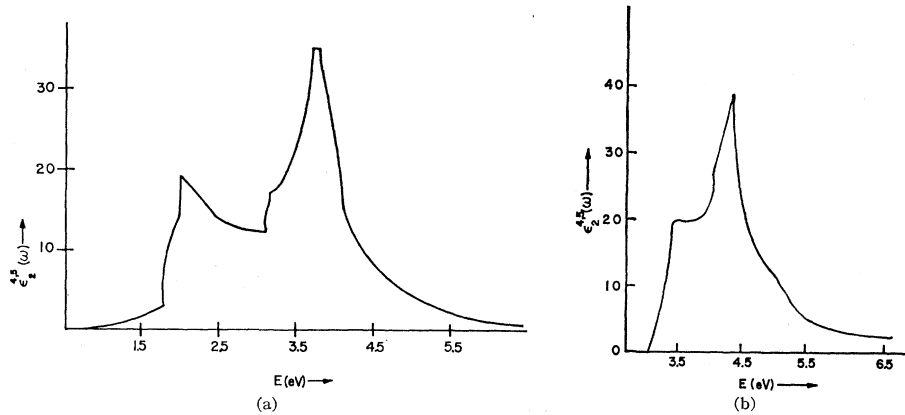
FIG. 14. Contribution of bands 4 and 5 to $\epsilon_2(\omega)$. (a) Ge, (b) Si.

points, we did the calculations at neighboring points. This was done so that degeneracies would be lifted, and the resulting matrix elements would be those of the bands being studied. If this were not done then, for example, the wave functions of bands 5 and 6 would mix at X.[39]

Looking at the values for $|M_{4,5}(\mathbf{k})|^2$ in Table III we see that it is roughly constant throughout the zone

TABLE III. Matrix elements in units of $(2\pi/a)^2$.

$\mathbf{k}$	$|M_{4,5}(\mathbf{k})|^2$	$|M_{4,6}(\mathbf{k})|^2$
(0.04,0.04,0.02)	0.90	0.01
(1.00,0.04,0.04)	1.51	0.00
(0.72,0.39,0.33)	1.33	0.09
(0.6,0.6,0.0)	1.41	0.00
(0.58,0.25,0.14)	0.94	0.22
(0.50,0.48,0.52)	0.85	0.63

varying by less than a factor of 2. We could then replace the matrix element in Eq. (1) by a $\mathbf{k}$-independent quantity $|M_{4,5}|^2$. To determine this value we took the weighted average of the values in Table III. Then $|M_{4,5}|^2 = 1.2(2\pi/a)^2$. As Table III points out, $|M_{4,6}(\mathbf{k})|^2$ is small except near the point L, precisely where the 4–6 peak originates. It is a bit difficult to settle on a way to handle the matrix element. We have

taken $|M_{4,6}|^2 = 0.6(2\pi/a)^2$. That is we take a constant value equal to $|M_{4,6}(\mathbf{k}=(\frac{1}{2},\frac{1}{2},\frac{1}{2}))|^2$. In Figs. 14 and 15 we show the contributions to $\epsilon_2(\omega)$ of the 4–5 and 4–6 histograms. In Fig. 16 we take $\epsilon_2(\omega) = \epsilon_2^{4,5}(\omega) + \epsilon_2^{4,6}(\omega)$, that is, as the sum of the contributions from the 4–5 and 4–6 histograms, and compare with experiment. In previous work (see Refs. 40 and 41) $\epsilon_2(\omega)$ was also taken as $= \epsilon_2^{4,5}(\omega) + \epsilon_2^{4,6}(\omega)$. Figure 6 shows our energy bands with the principal transitions marked.

To check our hypothesis regarding matrix elements it was decided to do Si with matrix elements included. $|M_{n,s}(\mathbf{k})|^2$ was computed along with the energy eigenvalues at each point of the zone, and put in Eq. (1). This added about 3 h of machine time to a total of about 15 h. The result for $\epsilon_2(\omega)$ shown in Fig. 16(b) includes all of the separate histograms.

As Fig. 16 indicates, our line shapes appear satisfactory. The two exceptions are the $L_{3'} \rightarrow L_3$ peaks in Ge and Si, as well as the $\Gamma_{25'} \rightarrow \Gamma_{15}$ peak in Si. The main difficulty in the first case centers around the strength. We have seen that our treatment probably overestimates the matrix element so that it is not surprising that the computed strength is too large. In Fig. 17 we examine the experimental line shape derived by taking the shaded portion of $\epsilon_2(\omega)$ as shown by the inset. With suitable magnification the theoretical and experimental line shapes are in good agreement. The $\Gamma_{25'} \rightarrow \Gamma_{15}$ peak is not well understood. Since it is so narrow, there is a
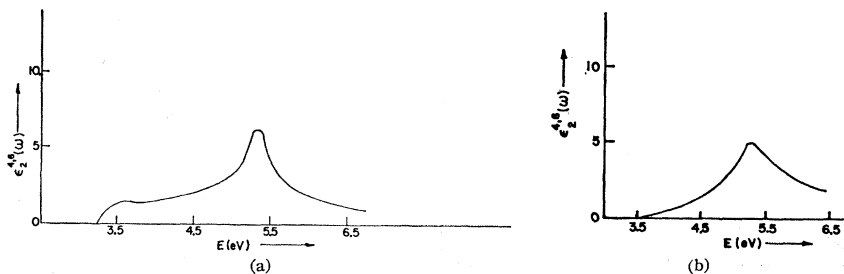

FIG. 15. Contribution of bands 4 and 6 to $\epsilon_2(\omega)$. (a) Ge, (b) Si.

[39] The program did not construct symmetrized combinations of plane waves as it was a general one designed to handle all points of the zone on an equal footing.
[40] D. Brust, J. C. Phillips, and F. Bassani, Phys. Rev. Letters 9, 94 (1962).
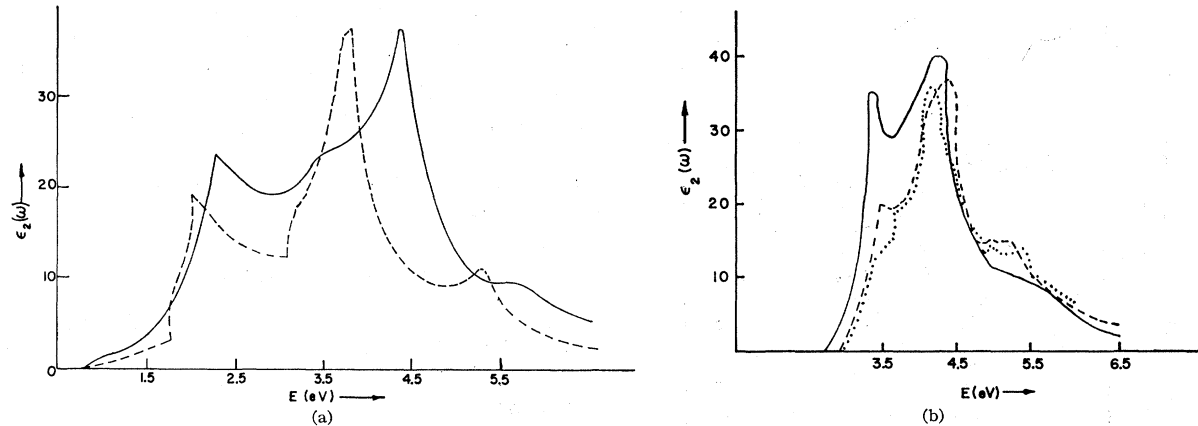[41] D. Brust, M. L. Cohen, and J. C. Phillips, Phys. Rev. Letters 9, 389 (1962).

FIG. 16. Experimental result for $\epsilon_2(\omega)$ (solid line) compared with our computed value (a) Ge $\epsilon_2(\omega) = \epsilon_2^{4,5}(\omega) + \epsilon_2^{4,6}(\omega)$ (dashed line, with edges emphasized to account for critical points), (b) Si $\epsilon_2(\omega) = \epsilon_2^{4,5}(\omega) + \epsilon_2^{4,6}(\omega)$ (dashed line), $\epsilon_2(\omega)$ taken as sum of contribution from all bands with dipole matrix elements included (dotted line).

chance that it is a density-of-states effect which we did not detect. The reader should see a further discussion concerning this point by Phillips.[42]

The results of this section are summarized in Table IV.

IX. STATIC DIELECTRIC CONSTANT

From a knowledge of $\epsilon_2(\omega)$ it is possible to calculate the real part of the dielectric constant $\epsilon_1(\omega)$ in the limit $\omega \to 0$. From the Kramers-Kronig relation the following expression can be obtained[43]

$$\epsilon_1(0) = 1 + \frac{2}{\pi} \int_0^\infty \frac{\epsilon_2(\omega)}{\omega} d\omega.$$

TABLE IV. Important critical points below 6 eV. In some cases they have been given to 0.01 eV with a reliability ~ 0.03 eV.

Type	Location	Bands	Symbol	Theoretical energy	Experimental energy
		Germanium			
M_0	(0,0,0)	$4 \to 5$	$\Gamma_{25'} \to \Gamma_{2'}$	0.6	0.8
M_0	(0.5,0.5,0.5)	$4 \to 5$	$L_{3'} \to L_1$	1.78	2.1
M_1	(0.17,0.17,0.17)	$4 \to 5$	$\Lambda_3 \to \Lambda_1$	2.01	2.3
M_1	(0.3,0.0,0.0)	$4 \to 5$	$\Delta_5 \to \Delta_1$	3.17	3.0 (?)
M_1	(1.0,0.0,0.0)	$4 \to 5$	$X_4 \to X_1$	3.6	4.3
M_2	(0.5,0.0,0.0)	$4 \to 5$	$\Delta_5 \to \Delta_1$	3.21	3.4 (?)
M_2	(0.61,0.61,0.0)	$4 \to 5$	$\Sigma_4 \to \Sigma_1$	3.8	4.4
M_0	(0.0,0.0,0.0)	$4 \to 6$	$\Gamma_{25'} \to \Gamma_{15}$	3.6	3.2 (?)
M_1	(0.56,0.56,0.39)	$4 \to 6$	$\cdots$	5.33	5.6
M_2	(0.5,0.5,0.5)	$4 \to 6$	$L_{3'} \to L_3$	5.44	5.7
		Silicon			
$\cdots$	(0.0,0.0,0.0)	$4 \to 5$	$\Gamma_{25'} \to \Gamma_{15}$	3.5	3.5
M_0	(0.5,0.5,0.5)	$4 \to 5$	$L_{3'} \to L_1$	3.15	3.0
M_1	(1.0,0.0,0.0)	$4 \to 5$	$X_4 \to X_1$	4.0	4.3
M_2	(0.4,0.4,0.0)	$4 \to 5$	$\Sigma_4 \to \Sigma_1$	4.4	4.4
M_0	(0.5,0.5,0.5)	$4 \to 6$	$L_{3'} \to L_3$	5.2	5.4

[42] J. C. Phillips, Phys. Rev. Letters **10**, 329 (1963).
[43] T. S. Moss, Proc. Phys. Soc. (London) **B66**, 141 (1953).

Using our results for $\epsilon_2(\omega)$, we can now derive $\epsilon_1(0)$. It is realized, however, that the value calculated in this way is directly proportional to the choice of matrix elements. In the case of the 4–5 bands we took this as $1.2(2\pi/a)^2$ with an assigned error $\sim 25\%$. For the other pairs of bands where the matrix element is considerably smaller we do not have a sufficiently reliable value for this purpose. Therefore, only the contribution from $\epsilon_2^{4,5}(\omega)$
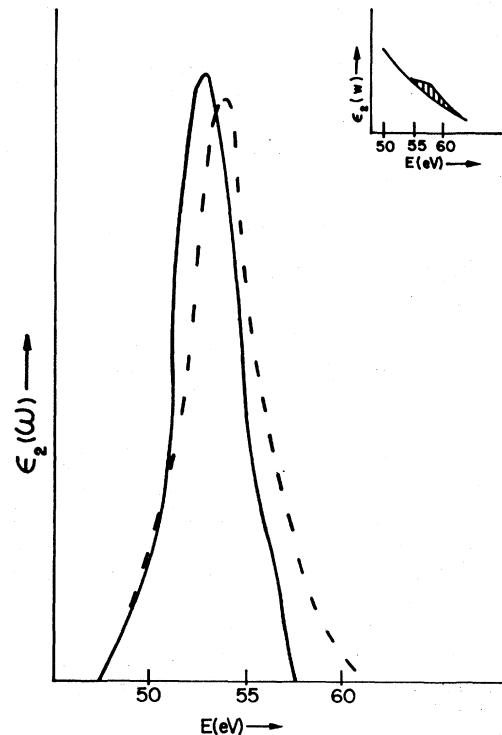

FIG. 17. Comparison of the theoretical and experimental line shapes for the $L_{3'} \to L_3$ transition in Ge. The solid line is the shaded portion of the experimental curve indicated by the inset. The shaded region has been moved 0.5 eV to the left and magnified seven times. The dashed line is a similar region under the theoretical curve.
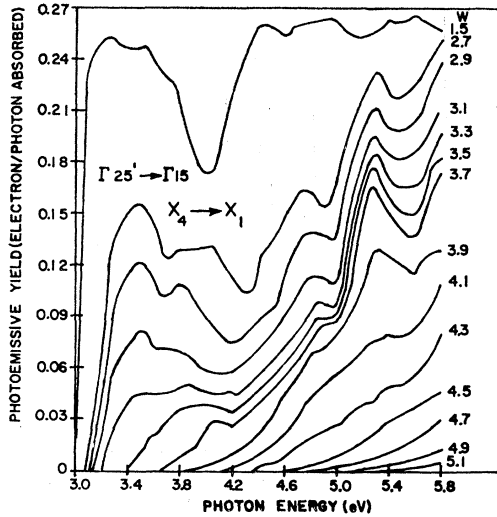
FIG. 18. Photoemissive yield: electrons/absorbed photons, as a function of photon energy, with vacuum W, eV, above the top of the valence band, as computed in present work.

was evaluated. For Ge this gave $\epsilon_1^{4,5}(0) = 12.4$ and for Si $\epsilon_1^{4,5}(0) = 7.6$, which are to be compared with the experimental values of 16.0 and 12.0, respectively. As expected $\epsilon_1^{4,5}(0) < \epsilon_1(0)$ due to the neglect of the higher bands. The 4–5 bands of Ge appear from these values to give a better result than those of Si. This, however, arises from the fact that the large peak in Ge has been erroneously shifted to a lower energy, and thus gives a bigger contribution to the integral. These values show that the 4–5 transitions, which account for most of the structure in $\epsilon_2(\omega)$, also contribute more than 50% of the value of $\epsilon_1(0)$, the macroscopic static dielectric constant.

X. g FACTOR AND SPIN-ORBIT EFFECTS

According to Roth and Lax,[15] the longitudinal component of the g tensor is

$$g_{11} = 2 - \frac{\Delta_{s0}}{\langle E_{L_1} - E_{L_{3'}} \rangle}\left(\frac{m_0}{m_t} - 1\right).$$

Here Δ_{s0} is the spin-orbit splitting of the $L_{3'}$ states. We use Lukeš and Schmidt's[44] value for Ge, $\Delta_{s0} = 0.195 \pm 0.02$ eV, and take the M_0 threshold in the optical response as the $L_{3'} \rightarrow L_1$ transition, which gives $\langle E_{L_1} - E_{L_{3'}} \rangle = 2.1$ eV. g_{11} is then found to be $= 0.96$ ($m_t = 0.082 m_0$). This is compared with the experimental value of 0.87. With the old interpretation that the edge itself is to be identified with $L_{3'} \rightarrow L_1$ we would use 2.35 eV for $\langle E_{L_1} - E_{L_{3'}} \rangle$ and get $g_{11} = 1.07$. Hence the new interpretation is seen to give better agreement with the g-factor data.

The identification of the edge with $\Lambda_3 \rightarrow \Lambda_1$ transitions allows us to resolve a problem in the spin-orbit splitting

of GaAs. In this material Δ_{s0} is observed to be $= 0.26$ eV, whereas from theory one would expect it to be 0.20 eV. The latter value is based on the assumption that the $L_{3'} \rightarrow L_1$ transition is responsible for the edge. In that case $\Delta_{s0}(L) = \frac{2}{3}\Delta_{s0}(\Gamma)$. The paradoxical result is resolved by the conclusion that the transition is associated with a point on the Λ axis where $\Delta_{s0} = 0.26$ eV, that is, falls between the extremes at Γ and L.

XI. PHOTOEMISSION RESULTS

The present work provides an interpretation not only of the dielectric properties of Ge and Si, but also of recent studies on the photoemissive properties of Si.

This work may be divided into two parts. Very precise studies[45] of weak photoelectric emission from Si near threshold on atomically clean surfaces reveal direct and indirect thresholds similar to the infrared absorption edges discussed by Hall, Bardeen, and Blatt.[46] The indirect threshold is at 5.15 eV, the direct threshold at 5.45 eV. The crystal surface is a (111) plane, and the thresholds, according to Kane,[47] are associated with symmetry line Λ, where V is directed normal to the (111) plane. The direct transition responsible for the direct threshold therefore starts *from* Λ_3 at an energy 0.3 eV below $\Gamma_{25'}$. The transition marked D in Fig. 6(b) starts from $\mathbf{k} = (0.12, 0.12, 0.12)$ and requires $E = 5.5$ eV. The agreement with Kane's theory and Allen and Gobelli's experiment is satisfying. It should be emphasized that this agreement is a by-product of our calculations for the ultraviolet absorption.

The second part of the photoelectric data concerns the broad yield curves obtained when the effective work function is reduced by covering the surface with up to one monolayer of Cs. In doing the analysis we assume that electrons which are excited into the conduction bands have an escape probability $P_s{}^W(\mathbf{k})$. Then $P_s{}^W(\mathbf{k}) = 0$ if $E_s(\mathbf{k}) < W$ and $P_s{}^W = P$ if $E_s(\mathbf{k}) > W$. Here $E_s(\mathbf{k})$ is the energy of the electron state with respect to the top of the valence band and W is the vacuum energy also measured relative to $\Gamma_{25'}$. With these assumptions the photoemissive yield function (number of electrons emitted/number of photons absorbed) is given by

$$Y_W(\omega) = \frac{\sum\limits_{n,s} \int\limits_{\text{B.Z.}} d^3k\, P_s{}^W(\mathbf{k})\delta(\omega_{n,s}(\mathbf{k}) - \omega)|M_{n,s}(\mathbf{k})|^2}{\sum\limits_{n,s} \int\limits_{\text{B.Z.}} d^3k\, \delta(\omega_{n,s}(\mathbf{k}) - \omega)|M_{n,s}(\mathbf{k})|^2},$$

(27)

in which it is assumed that the electron escapes with a negligible loss of energy.

[44] F. Lukeš and E. Schmidt, *Proceedings of the International Conference on the Physics of Semiconductors, Exeter* (Institute of Physics and the Physical Society, London, 1962), p. 389.

[45] G. W. Gobelli and F. G. Allen, Phys. Rev. **127**, 141 (1962).

[46] L. H. Hall, J. Bardeen, and F. J. Blatt, Phys. Rev. **95**, 559 (1954).

[47] E. O. Kane, Phys. Rev. **127**, 131 (1962).

In performing the calculations we arbitrarily set $P=0.26$. Then, again, taking constant matrix elements we have after making a finite sample of the zone

$$Y_W(\omega_i) = \frac{\sum_{n,s} K_{n,s}{}^W(\omega_i)|M_{n,s}|^2}{\sum_{n,s} K_{n,s}(\omega_i)|M_{n,s}|^2}, \quad (28)$$

where

$$K_{n,s}{}^W(\omega_i) = \sum_{\mathbf{k}_l} P_s{}^W(\mathbf{k}_l)\delta^{\Delta\omega}(\omega_{n,s}(\mathbf{k}_l)-\omega_i). \quad (29)$$

Again we replace the quantities $K_{n,s}(\omega_i)$ and $K_{n,s}{}^W(\omega_i)$ by the three point smoothed quantities $\bar{K}_{n,s}(\omega_i)$ and $\bar{K}_{n,s}{}^W(\omega_i)$.

In Fig. 18 we have plotted the results for $Y_W(\omega)$ with ω varied between 1.5 and 5.1 eV. This is to be compared

(a)

(b)
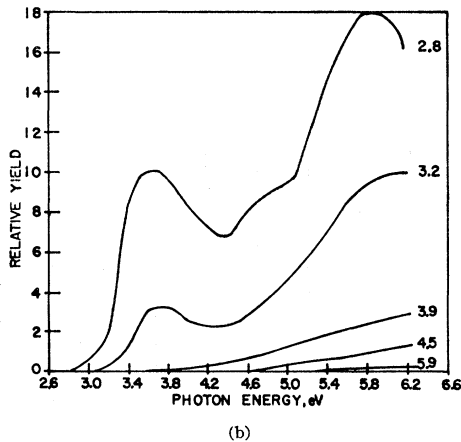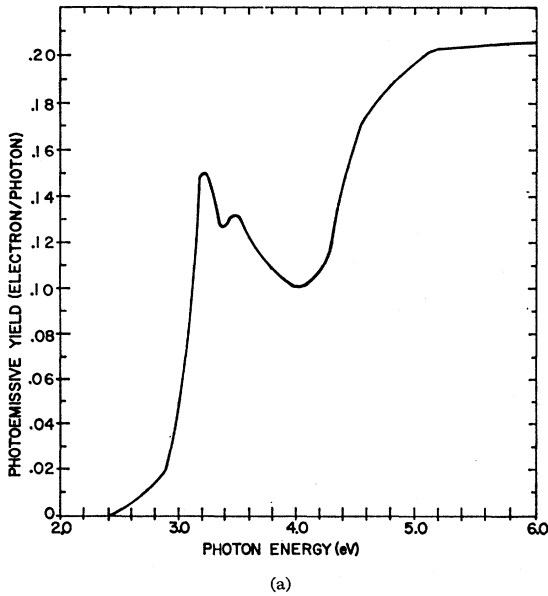
FIG. 19. Experimental results for the photoemissive yield
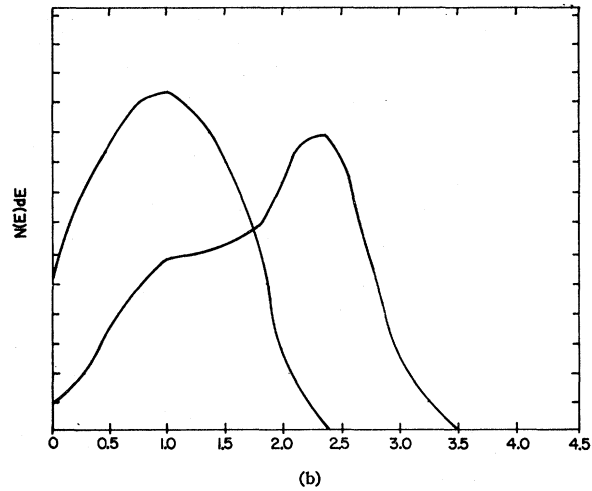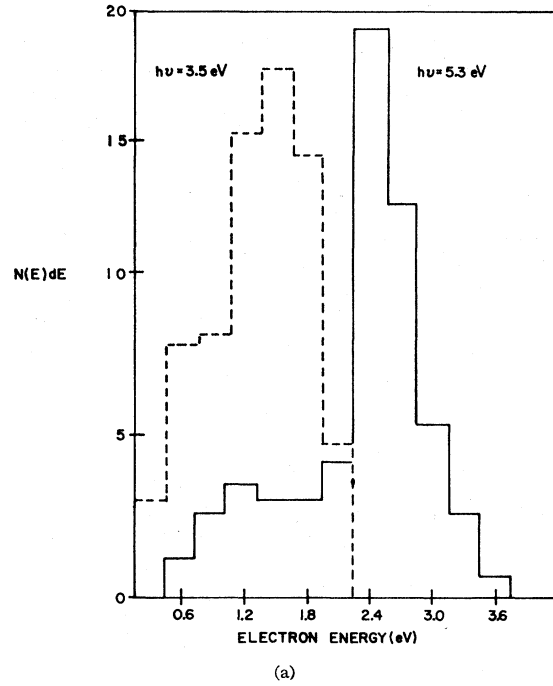(a) Spicer; (b) Gobeli and Allen.

(a)

(b)

FIG. 20. (a) The energy distribution of electrons emitted from Si for $\hbar\omega = 3.6\pm0.1$ eV and 5.3 ± 0.1 eV with $W=1.5$ eV, as computed; (b) Spicer's results for $P(E)$, $W=1.5$ eV.

with the experimental results obtained by Spicer[48] and by Gobeli and Allen[49] as shown in Fig. 19. The value of ω is experimentally varied, as we said, by covering the Si surface with a fraction of a monolayer of Cs. The over-all agreement appears satisfactory. The principal features in the $W=2.7$-eV plot are the peaks at 3.4 and 5.3 eV due to $\Gamma_{25'} \rightarrow \Gamma_{15}$ and $L_{3'} \rightarrow L_3$ transitions. The $X_4 \rightarrow X_1$ dip at 4.3 eV arises since $E(X_1)-E(\Gamma_{25'}) = 1.1$ eV $< W$. The secondary peaks at 3.8 and 4.8 eV

[48] W. E. Spicer and R. E. Simon, Phys. Rev. Letters **9**, 385 (1962).
[49] G. W. Gobeli and F. G. Allen (to be published).

have been identified as due to transitions near $\Lambda_3 \to \Lambda_1$ and points in the volume of the Brillouin zone, respectively. As W is increased the behavior of the curves is in quantitative agreement with the experiments.

In Fig. 20(a) we have drawn the energy distribution $P(E)$ for the emitted electrons with $W=1.5$ eV for the cases $\hbar\omega=3.5$ and 5.3 eV. Since our sample includes considerably fewer points, we take an energy interval of 0.3 eV. The experimental results of Spicer are drawn for comparison in Fig. 20(b). Again our results provide an interpretation of the experiment. The peak at 1.2 eV in the $\hbar\omega=3.5$-eV distribution is centered around the $\Gamma_{25'} \to \Gamma_{15}$ transition. The peak at 2.4 eV in the $\hbar\omega=5.3$-eV curve arises from $L_{3'} \to L_3$ transitions.

XII. DISCUSSION OF RESULTS

By starting with a model term scheme for the levels at Γ, X, and L, we have constructed a pseudopotential which has given the bands *throughout* the zone. The results we have found for $\epsilon_2(\omega)$ explain the line shapes which are seen in experiment. As an additional feature we can explain the results of recent studies on photoemission from Si.[50] Although our starting model had errors in it, they were subsequently straightened out by the band calculations. We also found several new critical points. This demonstrates the value of the optical data as a guide to theory. It gives us a way to start our calculation and a way to check it upon completion.

We have seen that the calculation was limited by convergence and sampling noise to errors ~ 0.05 eV. Further improvements of these quantities can be achieved by increasing the number of plane waves admitted to the expansion, and by taking more points into the zone sample. The next step would require the use of a more accurate potential.

By analyzing the optical properties of Ge and Si using a very simple model we have succeeded in exploring the electronic structure of these materials over an energy range of order 10 eV. This range is an order of magnitude larger than that studied by infrared experiments. The precision of our study enables us not only to explain all the observed structure but also to predict new features, such as M_0 thresholds.

Comparing the results for Ge and Si we find that alterations in the crystal potential are small and lead to much smaller changes in the over-all band structure than might have been supposed from infrared data. This agrees with the ultraviolet experimental results. One may say that in Ge and Si the energy surfaces in k space are determined primarily by the shape of the Brillouin zone, i.e., by symmetry considerations, and by $V(3)$, which measures the strength of the covalent bond.

---

[50] A further application has allowed us to explain some of the properties of Ge-Si alloys as well as high-pressure effects on the band structure [F. Bassani and D. Brust, Phys. Rev. **131**, 1524 (1963)].

APPENDIX

In Fig. 21 we plot $K_{4,5}(\omega_i)$. This should be compared with $\bar{K}_{4,5}(\omega_i)$ shown in Fig. 8(a) of the text, which shows the effect of smoothing.

To do a rigorous statistical analysis of the smoothing process would be exceedingly complex and not particularly rewarding. By a few heuristic arguments we can, nevertheless, estimate the random errors entering the problem. Consider first a one-dimensional situation. Suppose we have in this model a monotonically increasing $E_{ns}(\mathbf{k})$ curve. In Fig. 22(a) imagine that the solid lines divide the energy range into intervals of 0.1 eV, that is, $E_{ns}(\mathbf{k}_{i+1})=E_{ns}(\mathbf{k}_i)+0.1$ eV. Next suppose that a uniform net of sampling points is laid over the line. The result of this sample is used to approximate the true joint density of states. Suppose also that in computing $E_{ns}(\mathbf{k})$ we had convergence noise ~ 0.05 eV. There are then two sources of error in the computed joint density of states. The first results from the fact that the last sampling point in an interval is not significant. For example, in Fig. 22(b) shifting the boundary at 0.2 eV very slightly to the left will, put the last point in the 0.1–0.2 eV interval into the 0.2–0.3 eV interval. The second source of error arises because our calculation does not locate the boundary energies exactly. The idea is sketched in Fig. 22(c). One should note that the error in the energy eigenvalues is a dis-
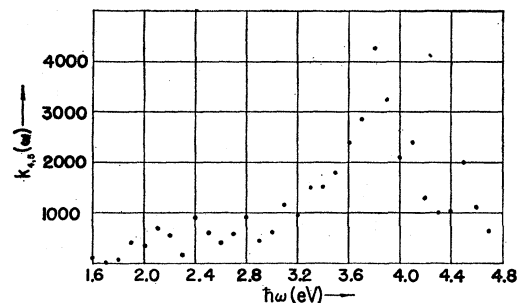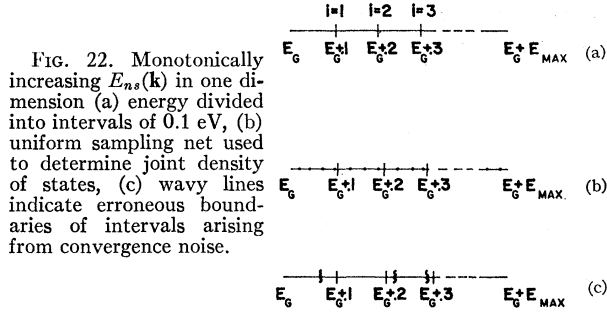

FIG. 21. $K_{4,5}(\omega_i)$, compare with $\bar{K}_{4,5}(\omega)$.

FIG. 22. Monotonically increasing $E_{ns}(\mathbf{k})$ in one dimension (a) energy divided into intervals of 0.1 eV, (b) uniform sampling net used to determine joint density of states, (c) wavy lines indicate erroneous boundaries of intervals arising from convergence noise.

continuous function of $\mathbf{k}$. This occurs because the number of plane waves in our expansion depends on $\mathbf{k}$.

If we take a sampling mesh of 1000 points and suppose that $E_{ns}(\mathbf{k})_{max} = 5$ eV, then our one-dimensional sample contains ~ 20 points per interval. These numbers roughly correspond to those of the actual problem. Then the error from boundary corrections is about 1 point per interval or 5%. The error arising from convergence noise can be ~ 10 points per interval or 50%. The effect of the smoothing is clear. We take three times as many points while leaving the absolute error constant. Hence the dominant error is reduced from about 50% to about 20%.

The situation in three dimensions is not so readily analyzed. The same sort of arguments ought to apply, however, we now have a two-dimensional surface for the boundaries of the energy intervals. The scatter resulting from boundary corrections and from convergence noise before smoothing is now estimated to be $\sim 20\%$, and after, $\sim 10\%$.