PHYSICAL REVIEW B VOLUME 17, NUMBER 4 15 FEBRUARY 1978

Chemical-bond approach to the electric susceptibility of semiconductors*

R. N. Nucho, J. G. Ramos,† and P. A. Wolff
Department of Physics, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 16 May 1977)

A simple model-independent method is developed to relate chemical bonds to the dielectric constant and other physical properties of tetrahedral semiconductors with the minimum number of parameters possible. For this purpose, we express $\epsilon_1(0)$, via the Kramers-Kronig relation, as a function of the zeroth and the first moments of $\epsilon_2(\omega)$. The first moment is determined by the f sum rule while the zeroth moment can be calculated if the valence- and conduction-band wave functions are known. Since conduction bands are inadequately described by models that are analytically simple, we bypass the problem by using completeness to eliminate the conduction band entirely. The result is an expression for $\epsilon_1(0)$ which involves only valence-band wave functions. Since working in a localized representation is more convenient than in the Bloch representation, we introduce a generalized Wannier function of bonding character for the valence bands. Realizing that this is appropriate for only those semiconductors like diamond in which the bonding-antibonding coupling is weak, we build into our Wannier function the lacking antibonding character via a power-series expansion in the quantity V_1/V_2 (Hall-Weaire parameters). Using Herman-Skillman values for the atomic orbitals, we obtain numerical results that agree with experiment to about 10%.

I. INTRODUCTION

The chemical-bond approach to the study of solids is an old one. In its earlier applications, it was only used in a qualitative way to explain the structure and stability of crystals. More recently,[1-9] there have been several attempts to make these ideas quantitative. The motivation for this interest stems from the multiple attractions of a chemical-bond approach. Such theories emphasize the bond aspect of crystal structure, a concept which is obscured in conventional band-structure theory, yet which accounts, in a qualitative way, for many chemical properties such as covalency, polarity, and metallicity. The chemical-bond approach is graphic, satisfying to the intuition, and is much simpler than band theory where difficult sums over all k space must be dealt with.

The newer theories of chemical bonding, most notably those of Hall-Weaire,[2-5] Phillips,[6,7] and Harrison,[8,9] generally assume a simple-model Hamiltonian for the underlying electronic structure of the crystal in question. The great virtue of these models is their simplicity and the possibility of dealing with them analytically. These features permit easy comparison between theoretical expressions and experimental results. However, because of the semiempirical nature of these theories, their success hinges upon a judicious choice of the experimental parameters which they contain, a problem which has led to some discrepancies between different calculations. In order to choose these parameters, a physical quantity which characterizes the chemical theory is needed. Such a quantity must be "global" in character, depending upon the overall distribution of quantum-mechanical states rather than on properties of a particular group of states. The dielectric constant is precisely of this nature, and is in fact the crucial element in both the Phillips and the Harrison models. After the parameters have been carefully chosen, these models can be used to predict other physical quantities such as average optical gaps, the macroscopic transverse charge (for ionic compounds), cohesive energies, and elastic constants, as well as other chemical properties like those mentioned earlier. In the Phillips theory, for instance, it is shown that the dielectric constant can be used to determine the ionic character of $A^N B^{8-N}$ compounds. Ionicity, in turn, enables one to make predictions concerning the coordination number and the type of bonding of these crystals. Compounds with ionicities less than the critical value 0.785 are predominantly covalent, while those with larger values are predominantly ionic, the ionicity scale varying between 0 and 1. This example clearly illustrates the value of the chemical bond approach in correlating trends within families of related compounds, as well as the important role of the dielectric constant. Moreover, such theories do have a certain degree of flexibility. As Weaire and Thorpe[4] have shown, this type of approach is well suited to the study of amorphous materials as well as periodic ones. Although there exists little hope for diagonalizing the Hamiltonian exactly in the amorphous case, one may expect to calculate properties involving, for instance, a trace over a complete set of valence-band states. The static dielectric constant, we will show, is in fact such a quantity. Of

course there are limitations. The band structures, for instance, that come out of such an analysis are much simplified,[8] and the resulting effective masses, naturally, are incorrect. High accuracy, on the whole, cannot be claimed. On the other hand, the overall picture of crystal structure and the new understanding of chemical trends that does emerge is well worth the effort.

We mentioned above three chemical-bond models. These models, although highly successful in describing chemical trends, are less satisfying in other respects. The Hall-Weaire model, for instance, yields a fair description of the valence band but a poor one of the conduction band. The situation is illustrated by the following example. Weaire and Thorpe[5] found that the parameters V_1 and V_2 of the Hall-Weaire model which gave a good fit for the valence-band density of states for Ge were $V_1 = -2.5$ eV and $V_2 = -6.75$ eV. We used[10] these parameters to determine the Hall-Weaire wave functions and energy levels for Ge and to calculate the imaginary part of the frequency-dependent dielectric function. We found that the resulting peak in $\epsilon_2(\omega)$ was at about $2V_2 = 13.50$ eV, far displaced from the experimental value of 4.3 eV. Since the form of $\epsilon_2(\omega)$ that we used in this calculation depends symmetrically on valence and conduction bands, we attribute this discrepancy to the poor conduction bands. The Phillips model, although it does yield good results, is based on a simplified Penn band-structure model[11] which bears little relation to that of real semiconductors. Finally, in Harrison's model, the shortcoming is the introduction of an arbitrary "scale factor" in the dielectric constant which we believe results, in part, from neglect of the coupling between bonding and antibonding states, as our forthcoming analysis would seem to indicate.

Although these models have strikingly different characteristics, they have also been remarkably successful in explaining chemical properties. This fact suggests the possibility of a more "model-independent" description of such properties. The aim of this paper is to construct a theory with that purpose. In particular, we will develop a relation between the static dielectric constant and localized chemical orbitals—namely, the valence-band Wannier functions—with the minimum number of parameters possible. We claim an accuracy of no better than 10% for reasons we shall discuss further on. We have achieved this aim for diamond, with no parameters at all. For the other group-IV semiconductors, we have introduced a single parameter which is a measure of the antibonding character of the valence band and which we relate to parameters of Hall-Weaire-like models. We also confirm the general expression for $\epsilon_1(0)$ used by Phillips.

In Sec. II we develop a relation between $\epsilon_1(0)$ and both conduction- and valence-band wave functions. We then find that $\epsilon_1(0)$ can be written as a trace over valence-band states, and that, as a consequence, we are not limited by any definite representation. In Sec. III we choose to use a localized basis and introduce the appropriate generalized Wannier function for the problem. Section IV is devoted to obtaining a general expression for the dielectric constant and to a comparison of numerical results with experimental values. We discuss the effect of high-lying core d states on $\epsilon_1(0)$ in Sec. V. Finally, Sec. VI consists of a comparison of our results to those of other models. The study of the III-V semiconductors in a similar framework will be presented in a subsequent paper.

II. THE FORMALISM

To find an expression for the low-frequency dielectric constant we start from the Kramers-Kronig relation at $\omega = 0$,

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

The standard form[12] for $\epsilon_2(\omega)$ in terms of position vector matrix elements is

$$\epsilon_2(\omega) = \frac{4\pi^2 e^2}{V\hbar} \sum_{\substack{kk' \\ nn'}} |\langle nk|\hat{\epsilon}\cdot\vec{r}|n'k'\rangle|^2 \times \delta\left(\omega - \frac{E_{n'}(k')}{\hbar} + \frac{E_n(k)}{\hbar}\right) , \tag{2a}$$

where $|nk\rangle$ and $|n'k'\rangle$ are Bloch functions representing filled and empty states, respectively. A summation over spin is implicit in the above expression. Since $n \neq n'$, the matrix element is diagonal in k space [see Eq. (B3)]. We have chosen this form for convenience. Realizing that simple models give a poor description of the conduction band (empty states), we are led to manipulate the above expressions in order to get rid of the conduction bands and to express Eq. (1) in terms of valence bands alone.

First note that Eq. (2a) can be written as the sum of the following two terms:

$$\epsilon_2^{\text{valence}}(\omega) = \frac{4\pi^2 e^2}{V\hbar} \sum_{\substack{kk' \\ vc}} |\langle vk|\hat{\epsilon}\cdot\vec{r}|ck'\rangle|^2 \times \delta\left(\omega - \frac{E_c(k')}{\hbar} + \frac{E_v(k)}{\hbar}\right) \tag{2b}$$

and

$$\epsilon_2^{\text{core}}(\omega)=\frac{4\pi^2e^2}{V\hbar}\sum_{\substack{k k'\\ \sigma c}}|\langle\sigma k|\hat{\epsilon}\cdot\vec{r}|ck'\rangle|^2 \times\delta\left(\omega-\frac{E_c(k')}{\hbar}+\frac{E_\sigma(k)}{\hbar}\right), \quad (2c)$$

where σ, v, and c refer to core, valence, and conduction bands, respectively, and $\hat{\epsilon}$ is the arbitrary polarization unit vector. Equation (2c) describes transitions between core and conduction states. Since the latter occur at high energies, and since there is a factor $1/\omega$ in the integrand of Eq. (1), it is permissible to drop the contribution to the dielectric constant from core transitions without losing more than a few percent in accuracy.[13] Henceforth, when we write $\epsilon_2(\omega)$, we shall mean only the part describing valence- to conduction-band transitions.

Let us next consider the form of $\epsilon_2(\omega)$ for a group-IV semiconductor, for instance Si. This function is illustrated in Fig. 1. The steep character of this curve, as well as its narrow width, suggest an expansion of the factor $1/\omega$ in Eq. (1) about some average frequency $\bar{\omega}$ in the following manner:

$$\frac{1}{\omega}=\frac{1}{\bar{\omega}}+\frac{1}{\omega}-\frac{1}{\bar{\omega}}=\frac{1}{\bar{\omega}}+\frac{1}{\bar{\omega}^2}(\bar{\omega}-\omega)+\frac{(\bar{\omega}-\omega)^2}{\bar{\omega}^2\omega}. \quad (3)$$

This procedure is essentially a moment method for evaluating $\epsilon_1(0)$. Indeed, if we use Eq. (3) to calculate $\epsilon_1(0)$ via Eq. (1), we see that the first term in Eq. (3) yields the zeroth moment of $\epsilon_2(\omega)$, the second term is related to both the zeroth and the first moments, and the third is an exact sum of all other moments that contribute to $\epsilon_1(0)$. We then choose $\bar{\omega}$ to eliminate the contribution of the second term in Eq. (3) and thus find

$$\bar{\omega}=\int_0^\infty \omega\epsilon_2(\omega)\,d\omega \Big/ \int_0^\infty \epsilon_2(\omega)\,d\omega = \frac{\pi}{2}\omega_p^2 \Big/ \int_0^\infty \epsilon_2(\omega)\,d\omega . \quad (4)$$

For the cases of diamond and silicon, ω_p is the valence-electron plasma frequency and we have made use of the well-known sum rule in evaluating Eq. (4). Combining Eqs. (1) and (3) we obtain

$$\epsilon_1(0)=1+\left(\frac{2}{\pi}\right)\frac{1}{\bar{\omega}}\int_0^\infty \epsilon_2(\omega)\,d\omega + \left(\frac{2}{\pi}\right)\int_0^\infty \epsilon_2(\omega)\frac{(\bar{\omega}-\omega)^2}{\bar{\omega}^2\omega}\,d\omega . \quad (5)$$

We have verified numerically from the data of Ehrenreich and Philipp,[14] that, for silicon, the contribution of the third term in Eq. (5) is less than 10%. This fact is a consequence of the narrowness of the $\epsilon_2(\omega)$ curve and leads us to make the approximation of dropping this term and to claim an accuracy no better than 10% in our subsequent calculation of $\epsilon_1(0)$. With the help of Eq. (4) we can rewrite Eq. (5) in the following form:

$$\epsilon_1(0)=1+\omega_p^2/\bar{\omega}^2 . \quad (6a)$$

Equation (6a) has the same form as the formula of Phillips.[6,7] $\bar{\omega}_{\text{Phillips}}$ corresponds to the position of the resonance in the $\epsilon_2(\omega)$ curve, whereas our $\bar{\omega}$ is proportional to the first moment of $\epsilon_2(\omega)$. We can interpret $\bar{\omega}$ as a mean optical gap.

For semiconductors which contain high-lying core d states, such as Ge and Sn, the peaks in $\epsilon_2(\omega)$ are broader, due mainly to the mixing of these core d states with the valence states.[15] As a result, the last term in Eq. (5) no longer is small. In fact, for the case of Ge, we estimate it to be about 27% of the total contribution to $\epsilon_1(0)$, again using the data of Ehrenreich and Philipp. We will consequently have to correct for this effect by introducing a factor D which we can calculate and which we identify with the similar factor introduced by Van Vechten.[13] The evaluation of the factor D will be discussed in Sec. V, where it will be shown that Eq. (5) can be

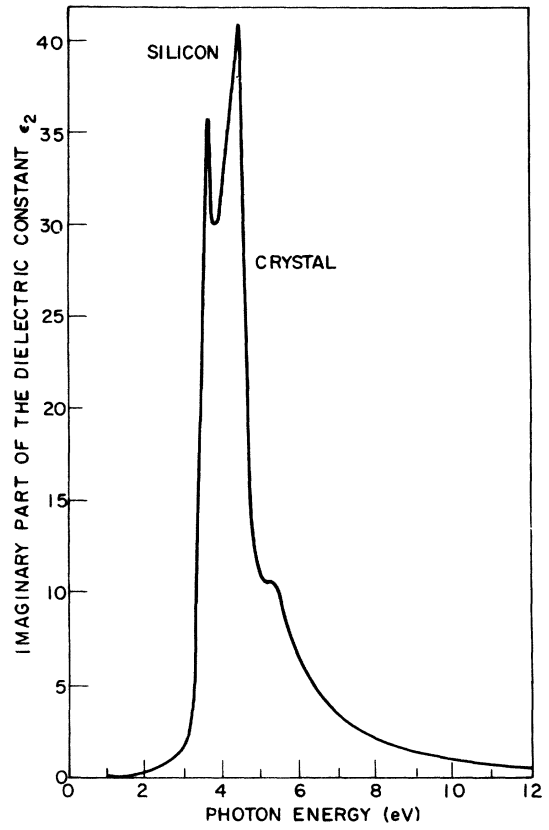

FIG. 1. Imaginary part of the frequency-dependent dielectric function (courtesy of Philipp and Ehrenreich).

written

$$\epsilon_1(0)=1+(\omega_p^2/\bar{\omega}^2)D\ . \tag{6b}$$

Clearly, if we restrict ourselves to semiconductors in which $D=1$, our problem will be reduced to a calculation of the integral $\int\epsilon_2(\omega)\,d\omega$ only. Substitution of the form given in Eq. (2b) for $\epsilon_2(\omega)$ then yields

$$\int_0^\infty \epsilon_2(\omega)\,d\omega=\frac{4\pi^2e^2}{V\hbar}\sum_{\substack{kk'\\ vc}}|\langle vk|\hat{\epsilon}\cdot\vec{r}|ck'\rangle|^2\ . \tag{7}$$

Completeness states that

$$\sum_{ck'}|ck'\rangle\langle ck'|=1-\sum_{v'k'}|v'k'\rangle\langle v'k'|-\sum_{\sigma k'}|\sigma k'\rangle\langle\sigma k'|\ .$$

This relation allows us to rewrite Eq. (7) as

$$\int_0^\infty \epsilon_2(\omega)\,d\omega=\frac{4\pi^2e^2}{V\hbar}\left(\sum_{vk}\langle vk|(\hat{\epsilon}\cdot\vec{r})^2|vk\rangle-\sum_{\substack{vv'\\ kk'}}|\langle vk|\hat{\epsilon}\cdot\vec{r}|v'k'\rangle|^2\right)\ . \tag{8}$$

Here we have neglected the core contribution which we show to be negligible in Appendix B. From Eqs. (4), (6a), and (8) we see that we have succeeded in expressing $\epsilon_1(0)$ in terms of valence-band wave functions alone.

III. GENERALIZED WANNIER FUNCTION

We wish now to relate $\epsilon_1(0)$ to localized chemical bonds. It is clear, then, that we need some localized representation rather than the Bloch one. The aim of this section is to develop localized wave functions (Wannier functions) for the valence band which have the character of bonding orbitals in the chemist's sense. What restrictions are we forced to comply with in defining our localized functions? To answer this question, we refer back to Eq. (8) and note that both terms can be viewed as a trace over valence bands only. In particular, the second can be written

$$\sum_{vk}\langle vk|\left(\sum_{v'k'}\hat{\epsilon}\cdot\vec{r}\,|v'k'\rangle\langle v'k'\,|\hat{\epsilon}\cdot\vec{r}\right)|vk\rangle\ .$$

Thus, we may evaluate Eq. (8) in any representation which can be obtained from the Bloch functions through a unitary transformation which mixes *only valence-band states*. This is a central point of our analysis which will become clearer as we proceed. In particular, we will later use this flexibility to define localized functions that have precisely the character of chemical bonds.

The discussion up to this point has been general. To illustrate, we shall choose to work in a tight-binding framework. This restriction is not necessary, as will be discussed at the end of this section, but is convenient for us. Moreover, it is well known that tight-binding descriptions do give an adequate picture of the valence bands of a crystal,[8] which are the states that concern us here. Thus, we consider the standard primitive cell for tetrahedral semiconductors containing two basic atoms at site i, with four sp^3 hybrids ($\psi_j^{\rm I}$) pointing from atom I to the nearest-neighbors (atom II) along the directions j ($j=1,\ldots,4$) and four other sp^3 hybrids ($\psi_j^{\rm II}$) pointing from these nearest neighbors to atom I. Let $\phi_j^a(\vec{r}-\vec{R}_i)$ and $\phi_j^b(\vec{r}-\vec{R}_i)$, respectively, be the antibonding and the bonding combinations[8] of these hybrids for a given bond j:

$$\phi_j^a(\vec{r}-\vec{R}_i)=[2(1-S)]^{-1/2}\times[\psi_j^{\rm I}(\vec{r}-\vec{R}_i)-\psi_j^{\rm II}(\vec{r}-\vec{R}_i-\vec{\delta}_j)]\ , \tag{9a}$$

$$\phi_j^b(\vec{r}-\vec{R}_i)=[2(1+S)]^{-1/2}\times[\psi_j^{\rm I}(\vec{r}-\vec{R}_i)+\psi_j^{\rm II}(\vec{r}-\vec{R}_i-\vec{\delta}_j)]\ , \tag{9b}$$

where $\vec{R}_i$ is a lattice vector for site i and locates atoms of type I, and $\vec{\delta}_j$ is a nearest-neighbor vector joining atom I with atom II along the direction j. S is the overlap integral between two hybrids which combine to form a bond. S is defined in Eq. (21a). It is well known[4,7,8,9,17] that the wave functions of Eqs. (9) diagonalize the simplest tight-binding Hamiltonian, which only includes the coupling (V_2) between every two hybrids forming a bond. Consequently, Eqs. (9a) and (9b) are orthonormal and correspond, respectively, to two sets of degenerate energy levels $E_a=-V_2$ and $E_b=V_2$. With these definitions we can now construct the usual Bloch-like tight-binding sums

$$\chi_j^a(\vec{r},\vec{k})=\sum_i\frac{e^{i\vec{k}\cdot\vec{R}_i}}{\sqrt{N}}\,\phi_j^a(\vec{r}-\vec{R}_i)\ , \tag{10a}$$

$$\chi_j^b(\vec{r},\vec{k})=\sum_i\frac{e^{i\vec{k}\cdot\vec{R}_i}}{\sqrt{N}}\,\phi_j^b(\vec{r}-\vec{R}_i)\ . \tag{10b}$$

The choice of sp^3 hybrids for the atomic orbitals results from various considerations, the most important of which is that they do yield reasonable valence bands as shown by Harrison[8] and by Kane.[16] Harrison used neutral sp^3 atomic orbitals to calculate his matrix elements, starting from an extension of the Hall model, while Kane used Gaussian wave functions of sp^3 character and a pseudopotential Hamiltonian. They both found that the valence bands were adequately represented by such a description. For nontetrahedral structures, of course, the choice of sp^3 hybrids will no longer be

a valid one.

With the choice of basis defined by Eqs. (10), the matrix of the most general tight-binding Hamiltonian can be divided into four blocks, as follows:

$$H=\begin{pmatrix} H^{aa} & H^{ab} \\ H^{ba} & H^{bb} \end{pmatrix} , \tag{11}$$

where each block is a 4×4 matrix for tetrahedral structures. In the limit of the simplest one-parameter Hamiltonian including only V_2, expression (11) reduces to a diagonal matrix with $H^{aa}=E_a$ and $H^{bb}=E_b$. Switching on the other interactions $(V_1, V_3, V_4, \ldots, V_n)$ introduces the remaining terms. These interactions appear in both H^{aa} and H^{bb}, and as we shall discuss later, are small compared to $E_a - E_b$. Now the diagonal blocks in (11) couple bonding orbitals to bonding orbitals or antibonding orbitals to antibonding orbitals, while the off-diagonal blocks mix bonding and antibonding states. This distinction is important because the two types of terms play quite different roles.[17] The diagonal blocks couple degenerate levels, hence all small terms in H^{aa} or H^{bb} cannot be treated by perturbation theory but must be diagonalized exactly, thus broadening the bonding and antibonding levels, respectively, into bands E_n^b and E_m^a. This procedure is usually hard to do. Fortunately we will not have to carry through this calculation, as will be shown below, because we are ultimately taking a trace which is independent of representation. The off-diagonal blocks, on the other hand, couple nondegenerate levels, and we can thus hope to treat them by perturbation theory. Henceforth, let H_0 denote the diagonal blocks of H, and H_1 the off-diagonal blocks. Thus, we can, in principle diagonalize H_0 and determine the eight zero-order Bloch eigenfunctions

$$\Psi_m^a(\vec{r},\vec{k})=\sum_j \alpha_{jm}^a(\vec{k})\chi_j^a(\vec{r},\vec{k}) , \tag{12a}$$

$$\Psi_n^b(\vec{r},\vec{k})=\sum_j \alpha_{jn}^b(\vec{k})\chi_j^b(\vec{r},\vec{k}) , \tag{12b}$$

where we shall let $m=1,\ldots,4$ and $n=5,\ldots,8$ for convenience. Here the α's are elements of 4×4 unitary matrices. Throughout this paper we use k and $\vec{k}$ interchangeably for the vector. Substituting Eq. (12b) into the usual expression for the Wannier functions and with the help of Eq. (10b) one can see that the Wannier functions for the valence band will consist of a mixture of all bond directions at a given site [as long as

$$\sum_k \alpha_{jn}(k)\neq \delta_{j,n-4} ,$$

a condition which is generally satisfied] and will also contain contributions from other sites. Clearly, these functions will not be very localized, nor would they describe chemical bonds in the chemist's sense. We resort then to the relative freedom we have in defining our localized functions which we established at the beginning of this section and define modified Wannier functions by mixing valence-band states as follows:

$$a_j^b(\vec{r}-\vec{R}_i)=\sum_k\sum_{n=5}^{8}\frac{e^{-i\vec{k}\cdot\vec{R}_i}}{\sqrt{N}}\alpha_{nj}^{b\dagger}(\vec{k})\Psi_n^b(\vec{r},\vec{k}) . \tag{13}$$

For an unspecified set of coefficients α, expression (13) is the most general possible valence-band Wannier function. With the help of Eqs. (12b) and (10b) and the unitarity property of the α matrices, namely, that

$$\sum_n \alpha_{in}(k)\alpha_{nj}^{\dagger}(k)=\delta_{ij}$$

for α^a and α^b separately, one can verify that Eq. (13) reduces very conveniently to

$$a_j^b(\vec{r}-\vec{R}_i)=\phi_j^b(\vec{r}-\vec{R}_i) , \tag{14}$$

which associates our generalized Wannier function for the valence band, in the limit $H_1=0$, with a bonding combination of hybrids along a given bond. Similar results are true for the antibonding case.

Calculating Eq. (8) with the above defined localized functions is thus equivalent to neglecting bonding-antibonding coupling (H_1) and should yield good results for those semiconductors, like diamond, in which this coupling is small. We did this calculation and found corroborating results. Indeed, we predicted a dielectric constant that was accurate to within a few percent for diamond, but that was far too low (30%–45%) for the other three elemental semiconductors. See Sec. VI for a full discussion.

To account for this bonding-antibonding coupling, we must then include H_1 which we can deal with by perturbation theory, as indicated before. Our choice of perturbative method is suggested by the usual effective mass transformation in which the interband terms of the Hamiltonian are removed by a canonical transformation.[17,32] In our case the bonding-antibonding coupling H_1 plays the role of these interband terms. To carry this process through, proceed as follows. We know that the set of "unperturbed" eigenfunctions $\{\Psi_n\}$ defined by Eqs. (12) diagonalizes H_0. Here we let $n=1,\ldots,8$ to include both antibonding functions ($n=1,\ldots,4$) and bonding functions ($n=5,\ldots,8$). We seek the set of "perturbed" eigenfunctions $\{\Psi_n'\}_{n=1,\ldots,8}$, which diagonalizes $H=H_0+H_1$ to first order in H_1. To find them, we perform a unitary transformation $U=e^{iT}$ (with $T=T^{\dagger}$) which takes us from the $\{\Psi_n\}$ representation to the $\{\Psi_n'\}$ representation. The Hamiltonian transforms as follows[19]:

$$H'=UHU^{\dagger}=e^{iT}He^{-iT} . \tag{15a}$$

A typical matrix element is

$$H'_{mn} = \langle \Psi'_m | H_{0p} | \Psi'_n \rangle = \sum_{l,s} \langle \Psi'_m | \Psi_s \rangle \langle \Psi_s | H_{0p} | \Psi_l \rangle \langle \Psi_l | \Psi'_n \rangle \ .$$

From this equation we readily identify

$$U^{\dagger}_{ln} = \langle \Psi_l | \Psi'_n \rangle \ .$$

The perturbed eigenfunctions are then given by

$$|\Psi'_n\rangle = \sum_l |\Psi_l\rangle \langle \Psi_l | \Psi'_n \rangle = \sum_l U^{\dagger}_{ln} |\Psi_l\rangle = \sum_l U^{*}_{nl} |\Psi_l\rangle \ , \qquad (15b)$$

which, in its last form, corresponds to correct matrix multiplication. The point here is to choose U (i.e., T) in such a way that the bonding-antibonding coupling is removed to first order. We accomplish this by expanding H' in powers of T, obtaining

$$H' = \{H_0\} + \{H_1 + i[T, H_0]\} + \{i[T, H_1] + \tfrac{1}{2} i^2 [T, [T, H_0]]\} + \{\cdots\} \ ,$$

where the dots represent higher-order terms, the brackets indicate the successive orders, and the commutators arise from expanding the exponentials Eq. (15a). We then set

$$H_1 + i[T, H_0] = 0 \ . \qquad (16a)$$

This equation determines T which will clearly be of the same order as H_1. Thus, typical matrix elements of T, to first order in H_1, are

$$T^{aa}_{mn} = T^{bb}_{mn} = 0 \quad (\text{for } m, n \leq 4 \text{ or } m, n > 4) \ ,$$

$$T^{ab}_{mn} = (T^{ba}_{nm})^* = i \langle \Psi^a_m | H_1 | \Psi^b_n \rangle / (E^b_n - E^a_m) \qquad (16b)$$

$$(\text{for } m \leq 4 \text{ and } n > 4) \ .$$

The perturbed eigenfunctions for the valence band will then be given by Eq. (15b) with $n = 5, \ldots, 8$:

$$|\Psi^v_n\rangle = \sum_{m=1}^{8} (I + iT)^*_{nm} |\Psi_m\rangle = |\Psi^b_n\rangle + \sum_{m=1}^{4} \left(\frac{\langle \Psi^a_m | H_1 | \Psi^b_n \rangle}{E^b_n - E^a_m} |\Psi^a_m\rangle \right) \ . \qquad (17a)$$

It is thus clear that Eq. (13) will yield an expression for the perturbed Wannier function if we substitute the perturbed Bloch functions, given above, in the right-hand side of that equation. If we write the matrix element in (17a) more explicitly, we find, with the help of Eqs. (12):

$$\langle \Psi^a_m | H_1 | \Psi^b_n \rangle = \sum_{i',l} \alpha^{*a}_{i'm}(k) \alpha^b_{ln}(k) H^{ab}_{i'l}(k) \ , \qquad (17b)$$

where

$$H^{ab}_{i'l}(k) = \langle \chi^a_{i'} | H_1 | \chi^b_l \rangle \ . \qquad (17c)$$

For further simplification, let us expand the energy denominator of Eq. (17a) about the quantity Δ, the average energy separation between bonding and antibonding bands, as follows:

$$\frac{1}{E^b_n - E^a_m} = \frac{1}{\Delta} \sum_{n=0}^{\infty} \left(1 - \frac{E^b_n - E^a_m}{\Delta} \right)^n \ .$$

Then, keeping only the first term in this expansion and combing Eqs. (12), (13), (17a), and (17b), we obtain the form

$$a^v_j(\vec{r} - \vec{R}_i) = \phi^b_j(\vec{r} - \vec{R}_i) + \frac{1}{\Delta} \sum_k \sum_{n=5}^{8} \sum_{m=1}^{4} \frac{e^{-i\vec{k}\cdot\vec{R}_i}}{\sqrt{N}} \alpha^{b\dagger}_{nj} \sum_{l,l'} \alpha^{*a}_{l'm}(k) \alpha^b_{ln}(k) H^{ab}_{l'l}(k) \sum_{j'} \alpha^a_{j'm}(k) \chi^a_{j'}(\vec{r}, \vec{k}) \ ,$$

where the first part of the wave function comes from our previous result of Eq. (14). With the unitarity property of the α matrices, much simplification occurs, and the generalized Wannier function for the valence band reduces to

$$a^v_j(\vec{r} - \vec{R}_i) = \phi^b_j(\vec{r} - \vec{R}_i) + \frac{1}{\Delta} \sum_k \sum_{j'=1}^{4} \left(\frac{e^{-i\vec{k}\cdot\vec{R}_i}}{\sqrt{N}} H^{ab}_{j'j}(\vec{k}) \chi^d_{j'}(\vec{r}, \vec{k}) \right) . \qquad (18a)$$

This result clearly shows how the α coefficients drop out of the problem—sparing us the task of diagonalizing H_0. We attribute this fact both to the trace character of the calculation, which allowed us enough flexibility to introduce the $\alpha^{\dagger}$ matrix in our definition [Eq. (13)] of the Wannier function, and to the perturbative nature of the problem, which allowed us to approximate $E^b_n - E^a_m$ by Δ. To evaluate Eq. (18a), all we need, then, is an expression for H_1 in the bonding-antibonding representation. As for keeping only the first term in the energy denominator expansion, the approximation is justified when $\Delta \gg H_1$, the basis for treating H_1 by perturbation theory in the first place. The second term in the expansion is of order H_1/Δ and will thus only yield second order changes in the Wannier function.

To actually calculate the generalized Wannier function, we must know the explicit form of the Hamiltonian. The most general tightbinding Hamiltonian includes all possible interactions between orbitals. The Hall-Weaire model[4] is only a simple approximation to that Hamiltonian, but because of its analytical tractability we shall use it to calculate the Wannier function following the above

outlined procedure. The details of this calculation are presented in Appendix A. The resulting set of generalized Wannier functions for the valence band, orthogonal to order γ, is

$$a_j^v(\vec{r}-\vec{R}_i)=\frac{1}{(1+6\gamma^2)^{1/2}}\left[\phi_j^b(\vec{r}-\vec{R}_i)+\gamma\left(\sum_{j'\neq j}\phi_{j'}^a(\vec{r}-\vec{R}_i)-\sum_{j'\neq j}\phi_{j'}^a(\vec{r}-\vec{R}_i-\vec{\delta}_j+\vec{\delta}_{j'})\right)\right], \tag{18b}$$

where $\gamma=\frac{1}{4}V_1/V_2$ if S is neglected and $\gamma=\frac{1}{4}(M_1/M_2)(1-S^2)^{1/2}$ if $S\neq 0$. The quantities V_1, V_2, M_1, and M_2 are defined in Appendix A. $\vec{\delta}_j$ is the vector pointing from atom I to its nearest neighbor, atom II, along the direction $j (j=1,\ldots,4)$ (see Fig. 2). Equation (18b) contains the first two terms of a power-series expansion of the generalized Wannier function in the quantity V_1/V_2, which is of the order of 0.4 for most semiconductors. The parameter γ represents the strength of the antibonding mixing in the valence-band wave function. It can also be taken as a measure of the delocalization of the latter. The larger γ, the less localized in space the wave function is. It is interesting to define an antibonding character percentage $P_a=6\gamma^2/(1+6\gamma^2)$.

A few points are worthy of note at this time. First, as a test of our methods we have solved for the Wannier function of the linear monatomic chain using the usual definition for this function, and a Hall-Weaire-like Hamiltonian. We found qualitatively the same results as for the three-dimensional case (see Fig. 3). Secondly, a more realistic Hamiltonian than the Hall-Weaire model, with tight-binding parameters $V_1,\ldots,V_n$ would yield a more complicated generalized Wannier function. However, if one did set $V_{i>2}=0$, that function would reduce to Eq. (18b). Therefore, use of Eq. (18b) means neglecting all coupling of order $V_{i>2}/V_2$ and higher. Examining Fig. 4, we see that there are three types of coupling. Coupling between hybrids on one atom (V_1), coupling between hybrids on adjacent atoms (V_2, V_3, V_4, and V_5), and coupling between further neighbors (V_6 and above). V_1 and V_2 are accounted for in the Hall-Weaire model. We have separately shown that the parameters V_3, V_4, and V_5 do not change the form of our generalized Wannier functions. Their sole effect is to modify the average energy Δ between bonding and antibonding levels, thus yielding slightly different values for γ. However, since V_3, V_4, $V_5 \ll V_2$,[20,21] this energy separation will always be $2V_2$ as a first approximation. Alternatively, if one treats γ as a parameter, V_3, V_4, and V_5 do not matter. Finally, if we include the parameters V_6 and above, or if we keep second order terms in V_1, V_3, V_4, and V_5, only orbitals that are far apart are coupled, and the coupling is very small.[22] Thus, in calculating the matrix elements of r^2 and r, we expect small corrections as compared to local contributions, and we will neglect these in the calculation of the dielectric constant. Finally, we believe that our definition of localized function via Eq. (13) is universal in the sense that it is valid even in a non-tight-binding framework, i.e., that the coefficients $\alpha_{nj}(k)$ do exist regardless of the basis used. Alvarez[23] has constructed composite Wannier functions in a similar way, starting from the valence band wave functions in k space, and thus supports our above assertion. The virtue of the tight-binding representation is that such a calculation is rendered unnecessary.

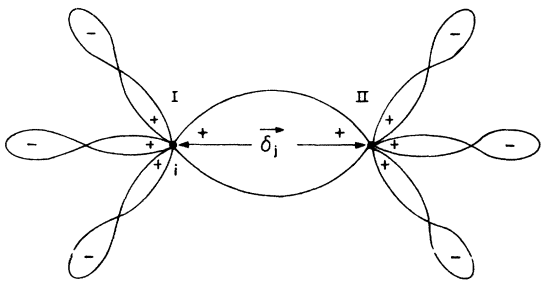

FIG. 2. Generalized Wannier function centered on bond j. $\vec{\delta}_j$ is the nearest-neighbor vector joining atoms I and II. The antibonding tails extend to the nearest-neighbors of atoms I and II.

IV. DIELECTRIC CONSTANT AND OPTICAL GAP

Having determined the localized valence-band functions, we proceed with the actual calculation of Eq. (8). As pointed out in Sec. III, this equation involves a trace over valence-band wave functions. Hence, it can be written in any general representation of the valence bands in the following way:

$$\int_0^\infty \epsilon_2(\omega)\,d\omega=\frac{8\pi^2e^2}{V\hbar}\left(\sum_{ij}\langle ij|(\hat{\epsilon}\cdot\vec{r})^2|ij\rangle-\sum_{\substack{ii'\\jj'}}|\langle ij|\hat{\epsilon}\cdot\vec{r}|i'j'\rangle|^2\right), \tag{19}$$

where i,j are quantum numbers which specify a complete set of valence-band states. For our purpose we shall use $|ij\rangle=a_j^v(\vec{r}-\vec{R}_i)$ given by Eq.

FIG. 3. Wannier function for the linear chain.

(18b). The sum over spin has now been completed, yielding a factor of 2.. The diagonal term will involve matrix elements of r^2 using the wave function shown in Fig. 2. The off-diagonal term is more complicated and will involve matrix elements of $\vec{r}$ between two Wannier functions that share an atom of type I or II. A typical configuration is shown in Fig. 5. After some tedious but straightforward algebra, we obtain the following result, to order γ^2:

$$\int_0^\infty \epsilon_2(\omega)\,d\omega = \frac{\pi m}{4\hbar}\omega_p^2(P-\gamma Q+\gamma^2 R), \tag{20}$$

where

$$P=\frac{1}{1+S}\sum_j\left[\langle r^2(j)\rangle_{\rm loc}+\langle r^2(j)\rangle_{\rm ov}-\hat{\epsilon}\cdot\vec{\delta}_j\langle r(j)\rangle_{\rm loc}+\left(\frac{\hat{\epsilon}\cdot\vec{\delta}_j}{2}\right)^2\right],$$

$$Q=\frac{2}{(1-S^2)^{1/2}}\sum_{j\neq j'}\left(\frac{S}{1+S}\hat{\epsilon}\cdot\vec{\delta}_j\langle r(j,j')\rangle_{\rm loc}-\langle r^2(j,j')\rangle_{\rm loc}\right),$$

$$R=-6P+\frac{2}{1-S}\sum_j\Big[3\langle r^2(j)\rangle_{\rm loc}-3\langle r^2(j)\rangle_{\rm ov}-3\hat{\epsilon}\cdot\vec{\delta}_j\langle r(j)\rangle_{\rm loc}+\sum_{j'\neq j}\langle r^2(j,j')\rangle_{\rm loc}$$
$$+\sum_{j\neq j'}\hat{\epsilon}\cdot\vec{\delta}_j\langle r(j,j')\rangle_{\rm loc}+\left(\frac{\hat{\epsilon}\cdot\vec{\delta}_j}{2}\right)^2(11-8S)\Big].$$

The computation, overall [i.e., Eq. (20)], is origin independent. However, once an origin is chosen, it must be used consistently throughout the calculation. On the other hand, the various integrals appearing in Eq. (20) may refer to different origins. There is no contradiction here. It is only a matter of performing translations and of grouping various terms for convenience. The quantities appearing in Eq. (20) are defined in the following, where $\psi^{\rm I}$ and $\psi^{\rm II}$ stand for the sp^3 hybrids introduced before:

$$S=\int \psi_j^{*\rm I}\left(\vec{r}_j+\frac{\vec{\delta}_j}{2}\right)\psi_j^{\rm II}\left(\vec{r}_j-\frac{\vec{\delta}_j}{2}\right)d^3r_j, \tag{21a}$$

$$\langle r^2(j)\rangle_{\rm ov}=\int \psi_j^{*\rm I}\left(\vec{r}_j+\frac{\vec{\delta}_j}{2}\right)(\hat{\epsilon}\cdot\vec{r}_j)^2 \times\psi_j^{\rm II}\left(\vec{r}_j-\frac{\vec{\delta}_j}{2}\right)d^3r_j. \tag{21b}$$

For these two overlap integrals, the origin is at the center of bond j. Referring to Fig. 4, we can see that $S=\langle 1|5\rangle$ and that $\langle r^2\rangle_{\rm ov}$ is of the type $\langle 1|(\hat{\epsilon}\cdot\vec{r})^2|5\rangle$. See discussion of S below:

$$\langle r(j)\rangle_{\rm loc}=\int \psi_j^{*\rm I}(\vec{r})(\hat{\epsilon}\cdot\vec{r})\psi_j^{\rm I}(\vec{r})\,d^3r, \tag{21c}$$

$$\langle r^2(j)\rangle_{\rm loc}=\int \psi_j^{*\rm I}(\vec{r})(\hat{\epsilon}\cdot\vec{r})^2\psi_j^{\rm I}(\vec{r})\,d^3r, \tag{21d}$$

$$\langle r(j,j')\rangle_{\rm loc}=\int \psi_j^{*\rm I}(\vec{r})(\hat{\epsilon}\cdot\vec{r})\psi_{j'}^{\rm I}(\vec{r})\,d^3r, \tag{21e}$$

$$\langle r^2(j,j')\rangle_{\rm loc}=\int \psi_j^{*\rm I}(\vec{r})(\hat{\epsilon}\cdot\vec{r})^2\psi_{j'}^{\rm I}(\vec{r})\,d^3r. \tag{21f}$$

The origin for these local integrals is at the site

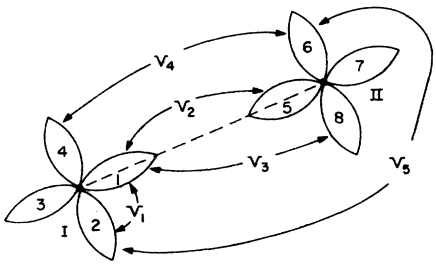

FIG. 4. Two atoms in a primitive cell with the eight hybrids surrounding them and pointing towards their nearest neighbors.

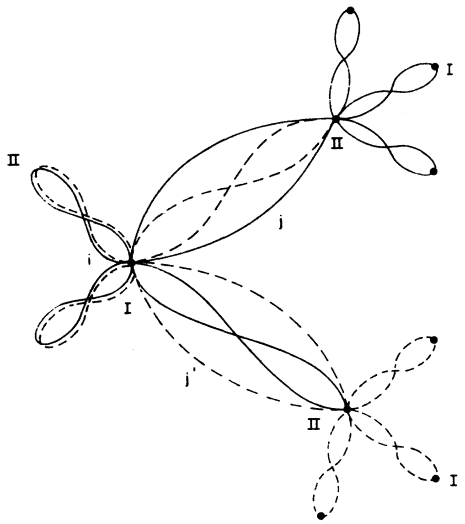

FIG. 5. The two generalized Wannier functions associated with neighboring bonds j and j'.

of atom I. Again referring to Fig. 4, we see that

$\langle r(j)\rangle_{\rm loc}$ is of the type $\langle 1|\hat{\epsilon}\cdot\vec{r}|1\rangle$,

$\langle r^2(j)\rangle_{\rm loc}$ is of the type $\langle 1|(\hat{\epsilon}\cdot\vec{r})^2|1\rangle$,

$\langle r(j,j')\rangle_{\rm loc}$ is of the type $\langle 1|(\hat{\epsilon}\cdot\vec{r})|2\rangle$,

$\langle r^2(j,j')\rangle_{\rm loc}$ is of the type $\langle 1|(\hat{\epsilon}\cdot\vec{r})^2|2\rangle$.

In obtaining Eq. (20), a number of small terms were dropped, based upon their numerical values which we calculated using the Herman-Skillman tables[24] for atomic orbitals. These small terms are discussed below. It is not the purpose of this paper to present all the details of the numerical calculations. We shall only emphasize a few major points. First of all, we note that a judicious choice of origin leads to some simplification. If we choose it at the center of bond, we see that, by symmetry, the diagonal part of the second term in Eq. (19) is identically zero. The off-diagonal part of that term we found to contain squares involving terms like $\langle 1|\hat{\epsilon}\cdot\vec{r}|1\rangle$, $\langle 1|\hat{\epsilon}\cdot\vec{r}|2\rangle$, and $\langle 1|\hat{\epsilon}\cdot\vec{r}|6\rangle$, which we neglected. The largest of these squares we found to be less than 0.2% of the dominant term. We estimate the total error in neglecting these quantities to be at most 5%. Thus, the main contribution to Eq. (19) comes from the first term of that equation, which yields P, Q, and R. The only terms dropped here are of the type $\langle 1|(\hat{\epsilon}\cdot\vec{r})^2|6\rangle$, the total contribution of which we showed to be proportional to the overlap $S'=\langle 1|6\rangle$. We computed both S and S' and found values of about 0.55 and 0.08, respectively, for the four elements considered. We believe that the Herman-Skillman orbitals are more spread out than they should be in the real crystal,[25] and that they hence overestimate these overlaps. For our numerical calculations we chose S' to be zero and S to be 0.50, a value which is consistent with that often referred to in the literature.[9]

We are now in a position to calculate $\bar{\omega}$ from Eq. (4), which, with the help of Eq. (20), can be written

$$\bar{\omega}=\frac{2\hbar/m}{P-\gamma Q+\gamma^2 R}\,. \tag{22}$$

To determine γ, two methods are possible. We can either select γ from literature values of V_1/V_2 which give a fit to valence band features (see Table I), or we can determine our own values for γ from a best fit of $\epsilon_1(0)$. The dielectric constant is then calculated from Eq. (6b) using the values for D given by Van Vechten.[13]

Our final results with literature values for γ are presented in Table II, while those using best fit values for γ are presented in Table III. In Table II we also indicate the energy gaps, E_h, calculated by Phillips.[7] These tables are discussed in Sec. VI.

V. CORE EFFECT AND THE EVALUATION OF D

In this section we show how the presence of high-lying core states (usually d states) affects our calculation and how the correction factor D, introduced in Sec. II, arises from our formalism. We then propose a scheme for the evaluation of D from experimental data for those semiconductors, such as Ge and Sn, which have high-lying d states. Finally, we compare our calculated values with those quoted by Van Vechten,[13] which he obtains from a simple empirical prescription for D.

After having examined the $\epsilon_2(\omega)$ data for various

TABLE I. Literature values of V_1/V_2 and the corresponding antibonding parameters γ. For the $S\neq 0$ case, γ was calculated using the expression derived in Appendix A.

	$(V_1/V_2)_{S=0}$				$\gamma_{S=0}$				$\gamma_{S\neq 0}$			
	C	Si	Ge	α-Sn	C	Si	Ge	α-Sn	C	Si	Ge	α-Sn
Pandey and Phillips[a]		0.27	0.39			0.07	0.10			0.12	0.17	
Chadi and Cohen[b]	0.22	0.29	0.39		0.05	0.07	0.10		0.09	0.13	0.17	
Harrison[c]	0.20	0.59	0.74	0.91	0.05	0.15	0.18	0.23	0.09	0.25	0.32	0.39
Harrison and Ciraci[d]	0.28	0.64	0.74	0.76	0.07	0.16	0.19	0.19	0.12	0.28	0.32	0.33
Weaire and Thorpe[e]			0.37				0.09				0.16	
Hirabayashi[f]	0.28	0.41	0.54	0.47	0.07	0.10	0.13	0.12	0.12	0.18	0.23	0.20

[a] We calculated V_1 and V_2 from the parameters of Ref. 20 with the help of the relations given by Hirabayashi (Eqs. 3 of Ref. 18) and the integrals given by Slater and Koster [Eqs. (12) of Ref. 22].

[b] Reference 21.

[c] Reference 8.

[d] Reference 9.

[e] Reference 5.

[f] The values for C are those of Herman, for Si and Ge those of Bassani and Yoshimine, and for α-Sn those of Bassani and Liu. Values quoted by Hirabayashi in Ref. 18.

TABLE II. Theoretical and experimental values of the static dielectric constant $\epsilon_1(0)$ and the average gap $\hbar\bar{\omega}$. Literature values for γ were used. Energies are in eV.

Element	γ	D[a]	$\hbar\bar{\omega}$	Experimental[b] peak in $\epsilon_2(\omega)$	E_h[c]	$\epsilon_1(0)$ theor.	$\epsilon_1(0)$[a] expt.	P_a(%)
C	0.09[d]	1.0	13.7	12.2	13.5	6.2	5.7	4.5
Si	0.13[d]	1.0	5.5	4.4	4.8	10.2	12.0	9
Ge	0.17[d,e]	1.25	4.6	4.3	4.3	15.4[f]	16.0	15
α-Sn	0.20[g]	1.46	3.2	3.5	3.1	24.1	24.0	19

[a]Reference 13.
[b]References 7 and 9.
[c]References 6 and 7.
[d]Reference 21.
[e]See footnote a in Table I.
[f]With our calculated value of $D=1.21$, we obtain the numerical result $\epsilon_1(0)=14.9$.
[g]Since data by the authors of Refs. 20 and 21 did not exist for α-Sn, we chose to use the value quoted by Hirabayashi (Ref. 18), obtained by Bassani and Liu from an orthogonal-plane-wave band-structure calculation.

semiconductors[14] (including the III–V compounds), we make the following observation. Larger values of D, as given by Van Vechten, correspond to larger widths of the $\epsilon_2(\omega)$ curves (see Fig. 6). This observation indicates a correlation between the effect of the high-lying core states, and the broadness of the $\epsilon_2(\omega)$ curves. To understand this phenomena, let us compare the $\epsilon_2(\omega)$ curves for Si (Fig. 1) and Ge (Fig 7). Since the scales are different, the Ge curve looks, misleadingly, excessively broad. It is clear, however, that transitions at both lower and higher energies are enhanced in Ge as opposed to Si. The reason for this enhancement is attributed to differences in band structure of both elements, the most striking being the $\Gamma_{2'}$, L_1, and Λ_1 states which drop by a few eV on going from Si to Ge. The transitions associated with these states account for the enhancement of the lower part of the $\epsilon_2(\omega)$ curve for Ge. The X_1 states remain almost unchanged, accounting for the main peaks in Ge and Si, which are at about the same energy. Other states are also modified, yielding a slightly broader curve at high energies. This behavior is explained by the existence of core d states in Ge. Because of their s-like character, conduction states such as $\Gamma_{2'}$ and L_1 are able to penetrate the d shell which lowers their energy considerably (for more details, see Ref. 26). We emphasize at this point that the core effect we are talking about is a potential effect, and has nothing to do with real core excitations which we discussed previously [see Eq. (2c) and the discussion following it]. Because of this broadness of the $\epsilon_2(\omega)$ curves, our moment method, as described in Sec. II, breaks down. More specifically, the third term of Eq. (5) which is, in fact, a measure of the broadness of $\epsilon_2(\omega)$, becomes large. For the case of silicon,

TABLE III. Numerical results using best fit values for γ. Energies are in eV.

Element	γ	$\hbar\bar{\omega}$	Experimental[a] peak in $\epsilon_2(\omega)$	P_a(%)
C	0	14.3	12.2	0
Si	0.16	5.0	4.4	13
Ge	0.18	4.5	4.3	16
α-Sn	0.20	3.2	3.5	19

[a]References 7 and 9.

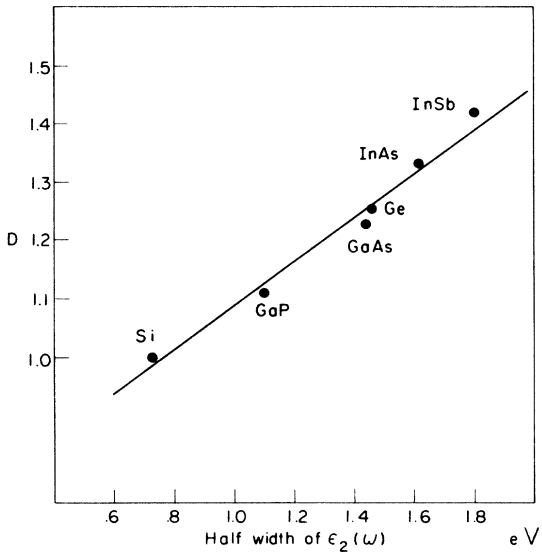

FIG. 6. Plot of the core correction factor D as a function of the half-width at $(1/e)$ times the maximum of the $\epsilon_2(\omega)$ curves for various semiconductors. The choice of semiconductors depended only on availability of the data.

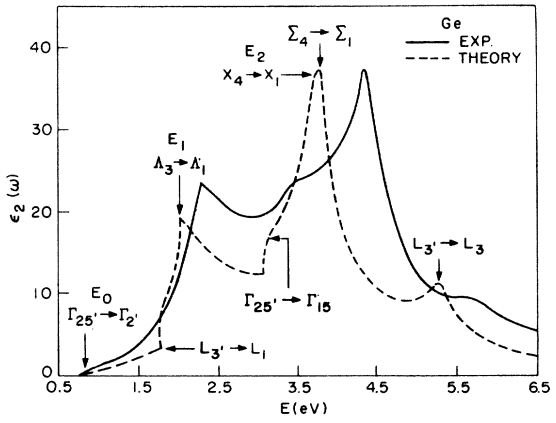

FIG. 7. Spectral structure of $\epsilon_2(\omega)$ in Ge. [D. Brust, J. C. Phillips, and F. Bassani, Phys. Rev. Lett. 9, 94 (1962); and D. Brust (Ref. 12).]

this quantity is 10%, whereas for germanium it is 27% of the total contribution to $\epsilon_1(0)$, as indicated in Sec. II. We can thus no longer neglect it. Fortunately, we do not have to calculate it explicitly, since we can approach the problem in a different manner.

The third term of Eq. (5) represents a weighted deviation of ω from an average energy $\overline{\omega}$ defined in Eq. (4). Because of the $1/\omega$ factor in the integrand which weights low energies, this term becomes large if $\epsilon_2(\omega)$ is broad in that region. Likewise, it becomes large at high energies if $\epsilon_2(\omega)$ does not rapidly converge to zero, since in this range, the integrand is proportional to ω. We hope to improve this situation by truncating our integrals at some frequency ω_c. As a result, the high-energy tail of $\epsilon_2(\omega)$ is cut off and the third term of Eq. (5) decreases. Moreover, introducing the cut-off frequency ω_c has the effect of reducing $\overline{\omega}$ to some new value $\overline{\omega}_{eff}$, as shown below. Consequently, the third term of Eq. (5) will represent the deviation of ω from this new lower value $\overline{\omega}_{eff}$, and should be much smaller than before. We will then set a lower bound on $\epsilon_1(0)$ by dropping this term, and we will show that an appropriate choice of ω_c will maximize it. The correction factor D will emerge from this process. With the help of experimental data once again, we were able to perform the integration in Eq. (1) numerically up to any finite frequency ω_c. We found that if ω_c is greater than or equal to a value roughly twice the peak of $\epsilon_2(\omega)$, the error in $\epsilon_1(0)$ is at most 5%. Physically, this corresponds to neglecting transitions between valence bands and high conduction bands. Thus, if we introduce a cut-off frequency in Eq. (1), the resulting approximation to $\epsilon_1(0)$ is still good as long as we bear in mind the condition on ω_c. We shall henceforth replace the infinite upper limits of all integrals by ω_c. Proceeding as in Sec. II, we find the analog of $\overline{\omega}$, namely,

$$\overline{\omega}_{eff}(\omega_c) = \int_0^{\omega_c} \omega\epsilon_2(\omega)\,d\omega \Big/ \int_0^{\omega_c} \epsilon_2(\omega)\,d\omega = \overline{\omega}(\omega_c)\,\frac{\eta_{eff}(\omega_c)}{4}\,, \tag{23a}$$

where

$$\overline{\omega}(\omega_c) = \frac{\pi}{2}\,\omega_{pv}^2 \Big/ \int_0^{\omega_c} \epsilon_2(\omega)\,d\omega\,, \tag{23b}$$

and where we have used

$$g(\omega_c) = \int_0^{\omega_c} \omega\epsilon_2(\omega)\,d\omega = \frac{\pi}{2}\,\omega_{pv}^2 \left(\frac{\eta_{eff}(\omega_c)}{4}\right)\,. \tag{24}$$

Equation (24) is the usual sum rule, modified by the presence of high-lying core levels. ω_{pv} is the valence-electron plasma frequency and $\eta_{eff}(\omega_c)$ is a number describing how many electrons effectively contribute to the sum rule[13] (see Appendix B). Equation (24) is another result of the broadness of the $\epsilon_2(\omega)$ curves. For diamond and silicon, η_{eff} saturates at the value of 4. Equation (23a) illustrates how the introduction of ω_c modifies our average gap. We can now obtain our lower bound on $\epsilon_1(0)$ from Eqs. (1) and (3) by dropping the third term in Eq. (3) and writing

$$\epsilon_1(0) - 1 \geq \frac{2}{\pi} \int_0^{\omega_c} \epsilon_2(\omega)\,d\omega \Big/ \overline{\omega}_{eff} = \frac{2}{\pi}\,\frac{f^2(\omega_c)}{g(\omega_c)}\,, \tag{25}$$

where $f(\omega_c) = \int_0^{\omega_c} \epsilon_2(\omega)\,d\omega$ and $g(\omega_c)$ was defined in (24). We must now determine the appropriate ω_{0c} to maximize Eq. (25), namely, the solution of the equation

$$\frac{d}{d\omega_c}\left(\frac{f^2(\omega_c)}{g(\omega_c)}\right)\Bigg|_{\omega_c=\omega_{0c}} = 0\,. \tag{26}$$

This relation yields

$$\omega_{0c} = \frac{2g(\omega_{0c})}{f(\omega_{0c})} = 2\overline{\omega}(\omega_{0c})\,\frac{\eta_{eff}(\omega_{0c})}{4}\,, \tag{27}$$

where we have used Eq. (23a). The curve $\eta_{\rm eff}(\omega_c)$ [i.e., $g(\omega_c)$] is shown in Fig. 8 for various semiconductors. We must disregard the knees in these curves, which arise from the core part of $\epsilon_2(\omega)$ that we have omitted from the start. The $f(\omega_c)$ curves have the same general shape except that they rise faster in the low-energy range and saturate faster in the high range. This result is due to the weighting factor ω in $g(\omega_c)$. From Eq. (27) it appears that ω_{0c} will satisfy the condition we established before [for Ge, we find $\omega_{0c} \simeq$ 8 eV, i.e., roughly twice the peak of $\epsilon_2(\omega)$]. Moreover, since $g(\omega_c)$ falls faster than $f(\omega_c)$ in that range, it follows

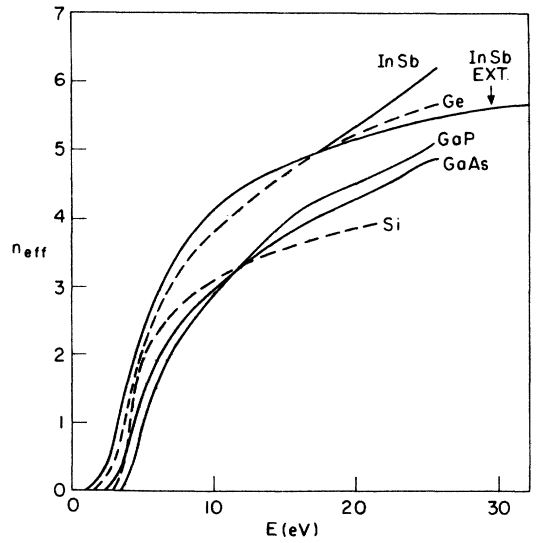

FIG. 8. η_{eff} vs energy (ω_c) for group-IV and -III-V semiconductors. [H. R. Philipp and H. Ehrenreich (Ref. 26)].

clearly that ω_{0c} cannot correspond to a minimum of f^2/g but must be a maximum, which is what we set out to find. Combining Eqs. (23)–(25), we obtain

$$\epsilon_1(0) = 1 + (\omega_{pv}^2/\bar{\omega}^2)D \ , \tag{28}$$

where $D = 4/\eta_{\text{eff}}(\omega_{0c})$, the form we quoted in Sec. II [Eq. (6b)].

We are now left with the problem of solving Eq. (27) for $\eta_{\text{eff}}(\omega_{0c})$. Assuming that $f(\omega_{0c})$ has effectively reached its saturation value, $f(\infty)$, Eq. (27) becomes easy to solve graphically since $\bar{\omega}$ is then a constant [see Eq. (23b)] which we can calculate from Eq. (22). With this assumption, we find $\eta_{\text{eff}} = 3.3$ for Ge and 3.0 for InSb. The corresponding values for D are 1.21 and 1.34, respectively, which are close to the values of Van Vechten, namely, 1.25 and 1.42. With the new value $\bar{\omega}_{\text{eff}}$, we find that the third term of Eq. (5) which was previously 27% has now fallen to about 10%. If on the other hand $f(\omega_c)$ does not saturate much faster than $g(\omega_c)$, as in the cases of diamond and Si, the above procedure fails, and Eq. (27) would have to be solved rigorously. Fortunately, the whole analysis of this section is unnecessary for these semiconductors, in which the d states are absent. Our numerical results, as well as the similarity between Eq. (28) and the Phillips–Van Vechten form, strongly suggest the use of the values for D quoted by Van Vechten.

VI. COMPARISON WITH EXPERIMENT AND OTHER MODELS

From Tables I, II, and III we observe that, although the values of γ are small, they show a definite trend in descending the column-IV semiconductors. The generalized Wannier function becomes more delocalized as we move from diamond to Sn, in keeping with the fact that these elements become more metallic as one approaches Sn. Our predicted values of $\epsilon_1(0)$, using literature values of γ from Table I, lie within about 10% of the experimental values. In addition, the mean optical gaps we have calculated are close to the experimental peaks in $\epsilon_2(\omega)$ and agree well with the gaps obtained by Phillips. Harrison's values for V_1/V_2 for the four elemental semiconductors do not give us as good results for our predicted $\epsilon_1(0)$, as can be seen by comparing Tables I and III. We do not use the Harrison values in our computations for the following reason. We are interested in the values of V_1/V_2 which give the best fit for the valence-band wave function. Harrison *et al.* determined V_2 either from its bond-length dependence (first set of values in Table I), or from the absorption peak of $\epsilon_2(\omega)$ (second set of values in Table I), and V_1 from atomic values. Pandey and Phillips and Chadi and Cohen, on the other hand, determined their parameters precisely from a fit to the valence band.

Assuming no bonding-antibonding mixing, we obtain energy gaps of 14.2, 6.1, 5.7, and 4.45 eV, respectively, for C, Si, Ge, and Sn. These values give results for the dielectric constant that are excellent for diamond, but about 30%, 35%, and 45% too low (respectively) for the other three elemental semiconductors. Turning to the work of Harrison and Pantelides[27] for a moment, we quote the values of their multiplicative scale factor. For C, Si, Ge, and Sn they are, respectively, 1.06, 1.22, 1.43, and 1.74. Had we neglected bonding-antibonding coupling and used the gaps indicated above, we would have had to introduce a similar factor in our expression for $\epsilon_1(0)$ in order to obtain good correspondence between theory and experiment. In this case, the values we would have needed for that correction factor are, respectively, 1.00, 1.22, 1.26, and 1.39. Harrison and Pantelides explain this scale factor as a parameter which describes the metallic trends of the matrix elements appearing in the expression for $\epsilon_2(\omega)$. This is precisely the character we have attributed to our γ. This discussion suggests that Harrison's scale factor is necessitated, at least in part, by neglect of antibonding mixing in the valence band.

It is interesting to examine how a small antibonding correction accounts for the relatively large discrepancy noted above. The explanation is twofold. First, because the Wannier function has six antibonding tails, there exists many states between which the electrons can hop. This gives a large value for R, almost 20 times as large as P, and

more than 10 times as large as Q (see Table VI). Secondly, $\epsilon_1(0)$ is proportional to $(P-\gamma Q+\gamma^2 R)^2$ which further enhances the contribution of the γ-dependent parts. The reason why R is by far the largest of the three coefficients can be seen by noting that the expression for R contains many diagonal matrix elements of r^2 between orbitals of the antibonding "tails," evaluated at distances of the order of a bond length or more from the center of bond origin. In P and Q, on the other hand, the contribution comes from overlap integrals evaluated at shorter distances from the origin. Physically, this tells us that these "outer" states, the more delocalized ones, play a large role in the behavior of the more metallic semiconductors.

We must mention that since the initiation of this work, an independent attempt has been made by Decarpigny and Lannoo[28] to calculate $\epsilon_1(0)$ from a molecular model, also using a moment method. The procedure they follow is close to ours but they evaluate their matrix elements using bonding combinations of Slater-type orbitals, the exponents for the free-atom functions being those of Clementi *et al.* Their results for $\epsilon_1(0)$ are too low by 30%, 53%, 55%, and 60%, respectively, for C, Si, Ge, and Sn, a discrepancy they explain by local-field effects. For the III-V semiconductors, even after including local-field effects, their results for $\epsilon_1(0)$ are still too low by 12% for GaAs to 46% for AlP. There is documentation both for[29] and against[30] including a local-field correction, the most recent evidence pointing against it. We shall not debate this issue here but merely point out that since the work of Decarpigny and Lannoo is so similar to ours, our results are an overwhelming indication that the local field effect is at most a small correction to this calculation of $\epsilon_1(0)$, certainly not a 30%–60% contribution. If we compare their energy gaps to our predictions when we neglect bonding-antibonding coupling, we find that the ratio of their gaps to ours is almost constant, but larger than unity. The significance of this is twofold. First, their neglect of antibonding mixing accounts for the fact that their agreement with experiment is less good as they descend the Periodic Table (where antibonding mixing becomes more important—see discussion earlier in this section). Secondly, the wave functions they used yield consistently a smaller quantity for the integral $\int \epsilon_2(\omega)\,d\omega$ than ours do, and hence a larger gap. This difference is possibly due to the more simplified nature of the Slater-type orbitals as compared to the actual atomic wave functions which we use. We note further that we repeated our computations using the Gaussian wave functions of sp^3 character constructed by Kane for the case of Si. The resulting values for $\hbar\bar{\omega}$ and $\epsilon_1(0)$ were better than those we obtained from Herman-Skillman orbitals.

To summarize, we have established a model-independent method to calculate $\epsilon_1(0)$ directly for semiconductors, with an error of the order of 10%. Our method was not based on a specific model for the electron Hamiltonian, rather on the sharpness of the $\epsilon_2(\omega)$ curve. Consequently, we expanded $\epsilon_1(0)$ via the Kramers-Kronig relation in terms of the moments of $\epsilon_2(\omega)$, and found that the expansion converged to within 10% if we kept only the *zeroth and the first moments*. The first moment was determined via the f sum rule and was found to be proportional to the valence electron plasma frequency, as is well known. The zeroth moment, on the other hand, could be calculated if the valence- and conduction-band *wave functions* were known. Having realized that conduction bands were inadequately described by analytically simple models or expressions, we chose to eliminate the conduction bands entirely via closure. The result was an expression for $\epsilon_1(0)$ and $\hbar\bar{\omega}$ in terms of localized functions describing the valence bands. For diamond and Si, the problem was thus reduced to finding these valence-band wave functions, whereas for Ge and Sn, we also had to account for the effect of high-lying core d states which interact with the valence bands. This new feature led us to introduce the factor D which we calculated for Ge. Unfortunately, no data was available for Sn, so in our final numerical results as shown in Table II, we used Van Vechten's values for both Ge and Sn for consistency.

The second part of the work consisted of determining appropriate localized functions for the valence bands. We started by using tight-binding sums of purely bonding combinations of sp^3 hybrids centered on adjacent atoms and found numerical results for $\epsilon_1(0)$ that were good for diamond but far too low for Si, Ge, and Sn. We then went back a few steps and accounted for the small but important coupling between bonding and antibonding states, thus obtaining the generalized Wannier functions. These functions were essentially bonding in character but contained, in addition, antibonding combinations of sp^3 hybrids introduced via a power-series expansion in the quantity $\gamma = V_1/4V_2$. Closed forms for $\epsilon_1(0)$ and $\hbar\bar{\omega}$ were then obtained in terms of the atomic orbitals, and the Herman-Skillman values for the latter yielded our numerical results. It must be pointed out that due to the relative arbitrarinesss in our choice of γ for α-Sn (see Table I) and to some ambiguity in the corresponding literature value for $\epsilon_1(0)$, our results for this element are not to be taken too seriously. They only serve to show the right trend.

In conclusion, we have been successful in describing diamond following the above scheme with

no parameters at all, and the other group-IV semiconductors with only one. Various possibilities for extending this work are clear. A treatment of the III-V semiconductors is feasible with only a few modifications and is now in process. Nontetrahedral structures, ranging from elements such as Te and Se to compounds like the rock salts, can also be treated within our framework. However, the atomic orbitals to be used in the more ionic cases are still to be determined, since it is clear that the neutral ones are no longer appropriate. Finally, a study of amorphous materials should also be possible, as well as the problem of magnetic susceptibility. The two limitations of our theory, of course, must be borne in mind. First, our method is perturbative. The bonding-antibonding coupling must be small, which restricts us to materials that are more insulating than conducting. Secondly, the curves describing the imaginary part of the frequency-dependent susceptibilities (either electric or magnetic) must be relatively well peaked in order for our moment method to be a good description.

ACKNOWLEDGMENTS

The authors would like to thank Professor M. Kastner and Professor J. D. Joannopoulos for many helpful discussions, as well as Dr. H. R. Philipp and Dr. H. Ehrenreich for making their data available to us.

APPENDIX A: CALCULATION OF ANTIBONDING PARAMETER

With the notation established in Sec. III, the Hall-Weaire Hamiltonian,[4] neglecting overlap, may be written

$$H = V_2 \sum_{i=1}^{N} \sum_{j=1}^{4} [\,|\psi_j^{\mathrm{I}}(\vec{r}-\vec{R}_i)\rangle\langle\psi_j^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_j)| + |\psi_j^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_j)\rangle\langle\psi_j^{\mathrm{I}}(\vec{r}-\vec{R}_i)|\,]$$
$$+ V_1 \sum_{i=1}^{N} \sum_{j\neq j'} [\,|\psi_j^{\mathrm{I}}(\vec{r}-\vec{R}_i)\rangle\langle\psi_{j'}^{\mathrm{I}}(\vec{r}-\vec{R}_i)| + |\psi_j^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_1)\rangle\langle\psi_{j'}^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_1)|\,], \tag{A1}$$

where $\vec{\delta}_1$ is the vector joining atom I to atom II in the primitive cell. In the representation defined by Eqs. (10), the above Hamiltonian takes the form of Eq. (11). Its matrix elements, defined in Eq. (17c), are given by the following expressions:

$$H_{jj'}^{aa} = -V_2\delta_{jj'} + \tfrac{1}{2}V_1(1+\theta_j\theta_{j'}^*)(1-\delta_{jj'}), \tag{A2}$$

$$H_{jj'}^{bb} = +V_2\delta_{jj'} + \tfrac{1}{2}V_1(1+\theta_j\theta_{j'}^*)(1-\delta_{jj'}), \tag{A3}$$

$$H_{jj'}^{ab} = H_{jj'}^{ba} = \tfrac{1}{2}V_1(1-\theta_j\theta_{j'}^*)(1-\delta_{jj'}), \tag{A4}$$

where

$$\theta_j = \exp(-i\vec{k}\cdot\vec{\delta}_j). \tag{A5}$$

The four vectors $\vec{\delta}_j$ are given by

$$\vec{\delta}_1 = \frac{a}{4}\begin{bmatrix}-1\\-1\\-1\end{bmatrix}, \quad \vec{\delta}_2 = \frac{a}{4}\begin{bmatrix}-1\\1\\1\end{bmatrix},$$

$$\vec{\delta}_3 = \frac{a}{4}\begin{bmatrix}1\\-1\\1\end{bmatrix}, \quad \vec{\delta}_4 = \frac{a}{4}\begin{bmatrix}1\\1\\-1\end{bmatrix},$$

where a is the lattice constant. Diagonalizing H^{bb} and H^{aa}, we obtain the eigenvalues

$$E_1^b = V_2 - V_1, \quad E_2^b = V_2 - V_1,$$
$$E_3^b = V_2 + V_1(1-\beta), \quad E_4^b = V_2 + V_1(1+\beta), \tag{A6}$$

and a similar set for the antibonding levels with $V_2 \to -V_2$, where

$$4(\beta^2 - 1) = (\gamma_x\gamma_y + \gamma_x\gamma_z + \gamma_y\gamma_z + \gamma_x^*\gamma_z + \gamma_x^*\gamma_y + \gamma_y\gamma_z^*) + \text{c.c.}$$

and

$$\gamma_j = \exp(-\tfrac{1}{2}ik_j a).$$

The energy bands shown in (A6) are not actually needed for our computation. They become harder to calculate as one includes more parameters in the model. They only serve here to illustrate that the bonding levels are separated from the antibonding levels by an energy difference of approximately $\Delta = 2V_2$, which justifies treating the bonding-antibonding coupling as a perturbation. From the secular equation we can, in principle, also find expressions for the coefficients $\alpha_{jn}^{a,b}$ of the eigenfunctions in Eqs. (12). We emphasize again, however, that explicit expressions are not needed.

We are now in a position to obtain the final form of our generalized Wannier function via Eq. (18a) and with the help of the matrix elements in (A4). The parts of these matrix elements which are independent of k yield the antibonding tails around atoms I whereas the k-dependent parts are essentially phase factors responsible for a shift in site, thus yielding antibonding tails around atoms II. The resulting expression is given by Eq. (18b) and

Fig. 2, with $\gamma_{s=0}=\frac{1}{4}V_1/V_2$.

We shall now repeat the calculation, including overlap. The matrix elements of the Hamiltonian, in this case, are defined to be

$$M_2=\langle\psi_j^{\mathrm{I}}(\vec{r}-\vec{R}_i)|H|\psi_j^{\mathrm{II}}(\vec{r}-R_i-\vec{\delta}_j)\rangle$$
$$=\langle\psi_j^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_j)|H|\psi_j^{\mathrm{I}}(\vec{r}-\vec{R}_i)\rangle\,,$$
$$M_1=\langle\psi_j^{\mathrm{I}}(\vec{r}-\vec{R}_i)|H|\psi_{j'}^{\mathrm{I}}(\vec{r}-\vec{R}_i)\rangle$$
$$=\langle\psi_j^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_1)|H|\psi_{j'}^{\mathrm{II}}(\vec{r}-\vec{R}_i-\vec{\delta}_1)\rangle\quad(j\neq j')\,,$$

instead of V_2 and V_1. Note also that the particular form of Eq. (A1) is no longer valid because it requires orthogonality of all the orbitals. We can start, however, with a Hamiltonian in matrix form such as that given in Ref. 4, p. 2518 in which we make the substitution $V_2\to M_2$ and $V_1\to M_1$. We can then find the matrix elements of the Hamiltonian in the representation defined by Eqs. (10). We obtain the analogs of Eqs. (A2)–(A5), namely,

$$H_{jj'}^{aa}=-\frac{M_2}{1-S}\,\delta_{jj'}+\frac{M_1}{2(1-S)}$$
$$\times(1+\theta_j\theta_{j'}^{*})(1-\delta_{jj'})\,,\qquad(A7)$$

$$H_{jj'}^{bb}=\frac{M_2}{1+S}\,\delta_{jj'}+\frac{M_1}{2(1+S)}$$
$$\times(1+\theta_j\theta_{j'}^{*})(1-\delta_{jj'})\,,\qquad(A8)$$

$$H_{jj'}^{ab}=H_{jj'}^{ba}=\frac{M_1}{2(1-S^2)^{1/2}}(1-\theta_j\theta_{j'}^{*})(1-\delta_{jj'})\,.\qquad(A9)$$

The energy bands are identical to those given in Eq. (A6) with the following substitutions:
For the bonding case,

$$V_2\to M_2/(1+S)\quad\text{and}\quad V_1\to M_1/(1+S)\,,$$

for the antibonding case,

$$V_2\to -M_2/(1-S)\quad\text{and}\quad V_1\to M_1/(1-S)\,.$$

The rest of the procedure is unchanged and we can find our generalized Wannier function exactly as before. Note that the average energy separation between bonding and antibonding bands is now approximately $\Delta=2M_2/(1-S^2)$. The final result is given by Eq. (18b) with $\gamma_{s\neq0}=\frac{1}{4}(M_1/M_2)(1-S^2)^{1/2}$.

Finally, we must relate M_2 to V_2 and M_1 to V_1 since most literature values are given for the zero overlap case and we need the $S\neq0$ values in our computations. Equating the average energy separation Δ in both cases yields immediately $V_2=M_2/(1-S^2)$. As for V_1, we obtain the corresponding relation by equating the width of the valence band (at $k=0$) for $S=0$ and $S\neq0$. We obtain $V_1=M_1/(1+S)$. With the help of these relations, we find $\gamma_{s\neq0}$ in terms of V_1 and V_2:

$$\gamma_{s\neq0}=\tfrac{1}{4}(V_1/V_2)[(1+S)/(1-S)]^{1/2}\,.\qquad(A10)$$

APPENDIX B

In writing Eq. (8) we omitted core terms which we can write explicitly as

$$\int_0^\infty \epsilon_2^{\text{valence}}(\omega)\,d\omega\Big|_{\text{core}}=-\frac{4\pi^2e^2}{V\hbar}$$
$$\times\sum_{\substack{v\,c\\kk'}}|\langle vk|\hat{\epsilon}\cdot\vec{r}|\sigma k'\rangle|^2\,.\qquad(B1)$$

Our aim here is to show that these terms are small compared to the remaining ones in Eq. (8). We shall, in fact, obtain an upper bound for the above expression by considering the well-known f sum rule in solids. For some valence band v, this sum rule[31,32] can be written, in isotropic materials,

$$\frac{2}{m}\sum_{\delta\neq v}\frac{|\langle vk|\hat{\epsilon}\cdot\vec{p}|\delta k\rangle|^2}{E_\delta(k)-E_v(k)}$$
$$=1-m\,\frac{\partial^2E_v(k)}{\hbar^2\partial k^2}$$
$$=\frac{2m}{\hbar^2}\sum_{k'}\sum_{\delta\neq v}[E_\delta(k)-E_v(k)]\,|\langle vk|\hat{\epsilon}\cdot\vec{r}|\delta k'\rangle|^2\,.\qquad(B2)$$

In the last expression we have used the relation[32] between momentum and position matrix elements

$$\langle nk|\hat{\epsilon}\cdot\vec{r}|n'k'\rangle=-i\vec{\nabla}_k\delta_{\vec{k},\vec{k}'}\delta_{nn'}+\frac{i\hbar\delta_{\vec{k},\vec{k}'}(1-\delta_{nn'})}{m[E_{n'}(k)-E_n(k)]}$$
$$\times\langle nk|\hat{\epsilon}\cdot\vec{p}|n'k'\rangle\,.\qquad(B3)$$

In integral form, the sum rule is equivalent to the first moment of $\epsilon_2(\omega)$ which, with the help of Eq. (2b), can be written

$$\int_0^\infty\omega\epsilon_2^{\text{valence}}(\omega)\,d\omega=\frac{4\pi^2e^2}{V\hbar^2}\sum_{\substack{kk'\\vc}}[E_c(k)-E_v(k)]$$
$$\times|\langle vk|\hat{\epsilon}\cdot\vec{r}|ck'\rangle|^2\,.\qquad(B4)$$

Writing the sum over conduction bands (c) as the sum over all bands $\delta\neq v$ minus the sum over core states (σ), Eq. (B4) splits into two terms. The first can be evaluated with the help of the sum rule (B2), yielding the valence-electron plasma frequency. The second gives a positive correction to this term and thus describes the effective number of electrons that contribute to the sum rule. This is the origin of the core correction factor D as explained by Van Vechten[13] and others.[26] We can express these results in the following expressions:

$$\int_0^\infty\omega\epsilon_2^{\text{valence}}(\omega)\,d\omega=\frac{\pi}{2}\,\omega_{pv}^2D\,,$$

where

$$\omega_{pv}^2 = (4\pi e^2/m)(N_{\text{val}}/V), \tag{B5}$$

$$D = 1 + N_{\text{core}}/N_{\text{val}},$$

and

$$N_{\text{core}} = \frac{2m}{\hbar^2} \sum_{\substack{kk' \\ v\sigma}} (E_v - E_\sigma) |\langle vk|\hat{\epsilon}\cdot\vec{r}|\sigma k'\rangle|^2. \tag{B6}$$

Since the core states lie below the valence bands, Eq. (B6) consists of a sum of positive terms. We can thus obtain an upper bound for expression (B1) via Eq. (B6) by letting $E_v - E_\sigma = \Delta E_{\min}$, the smallest energy difference between valence and core states. Combining Eqs. (20), (22), (B1), (B5), and (B6) we obtain our final result

$$\int \epsilon_2^{\text{val}}(\omega)\, d\omega \Big|_{\text{core}} \Big/ \int \epsilon_2^{\text{val}}(\omega)\, d\omega \leq (D-1)\frac{\hbar\bar{\omega}}{\Delta E_{\min}}.$$

For diamond and Si for which $D = 1$, this term is equal to zero. For Ge, the high-lying core d states are about 30 eV below the valence states. With our value for $\hbar\bar{\omega}$ and Van Vechten's value for D (see Table II), we obtain an upper bound of 4%.

---

*Work supported by the U. S. Air Force Office of Scientific Research, Air Force Systems Command under Contract/Grant No. AFOSR-76-2894.

†On leave of absence from Instituto de Fisica "Gleb Wataghin," Universidade Estadual de Campinas, Campinas, Sp, Brasil. Fellowship from FAPESP (Brasil).

[1]L. Pauling, *The Nature of the Chemical Bond* (Cornell U.P., Ithaca, N.Y., 1960).
[2]G. G. Hall, Philos. Mag. 43, 338 (1952).
[3]G. G. Hall, Philos. Mag. 3, 429 (1958).
[4]D. Weaire and M. F. Thorpe, Phys. Rev. B 4, 2508 (1971).
[5]M. F. Thorpe and D. Weaire, Phys. Rev. Lett. 27, 1581 (1971).
[6]J. C. Phillips, Rev. Mod. Phys. 42, 317 (1970).
[7]J. C. Phillips, *Bonds and Bands in Semiconductors* (Academic, New York, 1973).
[8]W. A. Harrison, Phys. Rev. B 8, 4487 (1973).
[9]W. A. Harrison and S. Ciraci, Phys. Rev. B 10, 1516 (1974).
[10]Calculation by R. N. Nucho and J. D. Joannopoulos using computer programs of J. D. Joannopoulos.
[11]D. R. Penn, Phys. Rev. 128, 2093 (1962).
[12]D. Brust, Phys. Rev. 134, A1337 (1964), Eq. (1), and with the help of Eq. (B3) of this paper.
[13]J. A. Van Vechten, Phys. Rev. 182, 891 (1969).
[14]Data taken jointly by H. Ehrenreich and H. R. Philipp.
[15]See Sec. V and Ref. 26 for a complete description.
[16]E. O. Kane, Phys. Rev. B 13, 3478 (1976).
[17]V. P. Sukhatme and P. A. Wolff, Phys. Rev. Lett. 35, 1369 (1975).
[18]K. Hirabayashi, J. Phys. Soc. Jpn. 27, 1475 (1969).
[19]A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1964), Vol. I, p. 287.
[20]K. C. Pandey and J. C. Phillips, Phys. Rev. B 13, 750 (1976).
[21]D. J. Chadi and M. L. Cohen, Phys. Status Solidi B 68, 405 (1975).
[22]With the help of the second-neighbor parameters of Pandey and Phillips (Ref. 20) and the integrals given by J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954), we estimated that the couplings $V_{i\geq 6}/V_2$ for Ge were at most 0.05 (compare with $V_1/V_2 \simeq 0.4$).
[23]C. V. de Alvarez and M. L. Cohen (private communication).
[24]F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice-Hall, Englewood Cliffs, N.J., 1963).
[25]J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1963), Vol. I.
[26]Reference 6, 13, and H. R. Philipp and H. Ehrenreich, Phys. Rev. 129, 1550 (1963).
[27]W. A. Harrison and S. T. Pantelides, Phys. Rev. B 14, 691 (1976).
[28]J. N. Decarpigny and M. Lannoo, Phys. Rev. B 14, 538 (1976).
[29]Chr. Flytzanis and J. Ducuing, Phys. Rev. 178, 1218 (1969); N. Wiser, *ibid.* 129, 62 (1963); A. R. Lubinsky, D. E. Ellis, and G. S. Painter, Phys. Rev. B 6, 3950 (1972); S. K. Sinha, R. P. Gupta, and D. L. Price, *ibid.* 9, 2564 (1974).
[30]J. A. Van Vechten and R. M. Martin, Phys. Rev. Lett. 28, 446 (1972); M. Kastner, Phys. Rev. B 7, 5237 (1973); W. Hanke and L. J. Sham, Phys. Rev. Lett. 33, 582 (1974); S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, *ibid.* 34, 155 (1975).
[31]P. M. Platzman and P. A. Wolff, *Waves and Interactions in Solid State Plasmas* (Academic, New York, 1973), p. 112.
[32]J. M. Luttinger and W. Kohn, Phys. Rev. 97, 869 (1955), Appendix B.