Local-field and excitonic effects in the optical spectrum of a covalent crystal*

W. Hanke[†] and L. J. Sham[‡]

Department of Physics, University of California San Diego, La Jolla, California 92037

(Received 24 June 1975)

A method of calculating the dielectric response in the time-dependent Hartree-Fock approximation using the Wannier representation is described and applied to the optical spectrum of diamond. The local-field effect reduces the prominent peak in the optical absorption spectrum asymmetrically, more severely on the low-energy side. The electron-hole attraction gives rise to the continuum-exciton effect which tends to enhance the prominent peak in a similarly asymmetric fashion. These results are important for the identification of structures in the experimental optical spectrum with electronic interband transitions, and quite generally for calculating dielectric response in a covalent crystal.

I. INTRODUCTION

In calculating the optical spectrum from an electronic energy-band structure, two approximations are frequently made: (a) neglecting the local-field effect, and (b) taking the polarization in the random-phase approximation (RPA). In this paper we attempt to improve on these two approximations and examine the consequences of the improvements on the optical spectrum of diamond.[1]

In a crystal, a weak external electric field of a small wave vector $\vec{q}$ and frequency ω will induce microscopic fields of large wave vectors $\vec{q}+\vec{G}$, $\vec{G}$ being a reciprocal-lattice vector, which in turn contribute to the macroscopic dielectric response[2,3] of wave vector $\vec{q}$ and frequency ω. The effect of these "Bragg-diffracted" fields has been termed the local-field effect. Alternatively, it may be viewed as the indirect contribution to the nonlocal dielectric function by the electric field due to the induced polarization distributed throughout the crystal. In a nearly-free-electron solid, the local-field effect is unimportant.[3] In the tightly bound limit, it can be accounted for by the Lorentz-Lorenz relation,[4] or more generally by a multipole expansion.[3] For a covalent crystal, where the electrons are neither completely localized nor completely delocalized, Van Vechten and Martin[5] were the first to calculate the local-field effect using the plane-wave expansion. In Sec. II, we discuss a different method for calculating the local-field effect in terms of the Wannier or linear combination of atomic orbitals (LCAO) representation.

In the random-phase approximation,[6] the (irreducible) polarization is due to the uncorrelated electron-hole pair, depicted by the diagram, Fig. 1(a). We have used the time-dependent Hartree-Fock approximation as opposed to the commonly used time-dependent Hartree approximation, which is equivalent to RPA. Thus, the polarization contains terms from the self-consistent Fock term, illustrated by Fig. 1(c), which is the exchange counterpart of the terms in RPA, illustrated by Fig. 1(b). These exchange processes may also be viewed as the electron-hole attraction. They may be shown to reduce to the effective-mass approximation for excitons[7] (bound or unbound), in the Wannier limit for frequencies near the fundamental band gap. Even in the effective-mass limit, the bound and unbound excitons have strong effects on the optical spectrum.[8] In addition, there is a rich variety of excitonic effects besides those in the neighborhood of the band gap.[9] The sum of all the exchange processes is expressible in an integral equation for the polarization, which is formally solved in the Wannier representation. Superficially, the excitonic effect simply modifies the Coulomb-interaction terms of the density response in RPA. This Hartree-Fock theory of the dielectric response is discussed in Sec. II.

Section III is devoted to a description of our approximate construction of the band structure and of the local representation of the electron wave
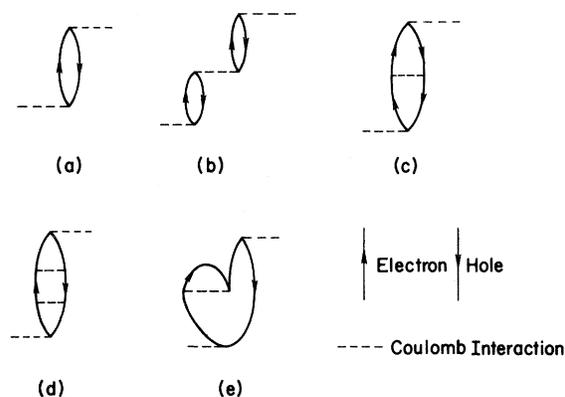
FIG. 1. Polarization diagrams.

functions of diamond for use in calculating the optical spectrum. Section IV discusses the results of the calculation of the optical absorption. We have found the local-field effect in diamond not to be negligible, but more importantly, the continuum-exciton effect to be indispensable. The combined local-field and continuum-exciton effects shift the main absorption peak by about 1.5 eV to lower frequencies compared to the RPA calculation without local-field corrections. The implications of these results for the identification of structures in the experimental optical spectrum with electronic interband transitions are discussed. In Sec. V, we present a simple model providing an explanation of our results for the optical properties of diamond, which may have wider applicability. Finally, we discuss the inapplicability of the Lorentz-Lorenz relation in covalent crystals.

II. THEORY OF DIELECTRIC RESPONSE IN INSULATORS

A. Local-field effect

When a crystal is perturbed by an external potential $V_{ext}(\vec{q}+\vec{G}, \omega)$ of wave vector $\vec{q}+\vec{G}$ and frequency ω, the total potential seen by a test charge is given, to the first order, by

$$V_{tot}(\vec{q}+\vec{G}', \omega) = \epsilon^{-1}(\vec{q}+\vec{G}', \vec{q}+\vec{G}; \omega)$$
$$\times V_{ext}(\vec{q}+\vec{G}, \omega), \qquad (2.1)$$

where ϵ^{-1} is the inverse dielectric function. $\vec{G}'$ is a reciprocal-lattice vector, signifying that the induced potential can have a wave vector, differing from that of the perturbation by a reciprocal-lattice vector. As a consequence, the macroscopic dielectric constant, which measures the macroscopic response to a macroscopic perturbation, i.e., for vanishingly small $\vec{q}$, is given by[2,3]

$$\epsilon(\omega) = \lim_{q \to 0} \frac{1}{\epsilon^{-1}(\vec{q}, \vec{q}; \omega)}. \qquad (2.2)$$

We have considered only the longitudinal dielectric constant. If we confine our attention to cubic crystals, for simplicity, then the transverse dielectric constant, relevant to the optical properties, is equal to the longitudinal dielectric constant[2,3] and is given by Eq. (2.2).

In a calculation of the dielectric constant, one usually starts with the "polarizability", which measures the induced electron density n_{ind} due to

the total screened potential,[10]

$$n_{ind}(\vec{q}+\vec{G}', \omega) = \chi(\vec{q}+\vec{G}', \vec{q}+\vec{G}; \omega)$$
$$\times V_{tot}(\vec{q}+\vec{G}, \omega). \qquad (2.3)$$

Since the total potential is the sum of the external potential and the induced potential due to n_{ind}, the dielectric function is given by

$$\epsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega)$$
$$= \delta_{\vec{G},\vec{G}'} - v(\vec{q}+\vec{G})\chi(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega), \qquad (2.4)$$

where $v(q) = 4\pi e^2/\Omega_0 q^2$ is the Coulomb potential, Ω_0 being the volume of the unit cell. To find the macroscopic dielectric constant from Eq. (2.2) involves an inversion of the dielectric matrix (2.4). The contribution of the nondiagonal elements of the dielectric matrix to the macroscopic constant is known as the local-field effect.[3] If one were to neglect the $\vec{G} \neq \vec{G}'$ elements, the macroscopic dielectric constant would simply be given by

$$\bar{\epsilon}(\omega) = \lim_{q \to 0} \epsilon(\vec{q}, \vec{q}; \omega). \qquad (2.5)$$

This much used approximation is only justified for systems with constant electronic density, i.e., for nearly-free-electron systems. For crystals with more or less tightly bound electrons the local-field effect, or formally, the inversion of the dielectric matrix, has to be considered.

B. Inversion of dielectric matrix in Wannier representation

We have developed a method of inverting the dielectric matrix by expressing the electron wave functions in terms of Wannier functions or LCAO's.[10-13] We briefly review the method, introducing some notations needed for the subsequent discourse. The single-electron Bloch wave ψ_{nk} of band n and wave vector $\vec{k}$ is expressible in terms of a set of Wannier functions[14,15] ϕ_ν:

$$\psi_{nk}(\vec{r}) = N^{-1/2} \sum_{\nu l} c_{n\nu}(\vec{k}) e^{i\vec{k}\cdot\vec{R}_l} \phi_\nu(\vec{r}-\vec{R}_l), \qquad (2.6)$$

where N is the number of lattice vectors $\vec{R}_l$. The number of Wannier functions is equal to the number of bands in contact with each other but isolated from the other bands.

Quite generally, the polarizability function, defined in Eq. (2.3), has the form, in terms of the Bloch waves, of[16]

$$\chi(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega) = \sum_{n_1 \cdots n_4 kk'} \langle \psi_{n_2k} | e^{-i(\vec{q}+\vec{G})\cdot\vec{r}} | \psi_{n_1k+q} \rangle \chi(n_1k+q, n_2k, n_3k', n_4k'+q; \vec{q}, \omega)$$
$$\times \langle \psi_{n_4k'+q} | e^{i(\vec{q}+\vec{G}')\cdot\vec{r}} | \psi_{n_3k'} \rangle. \qquad (2.7)$$

When the Bloch waves are expressed in terms of the Wannier functions, the polarizability has the "separa-

ble" form

$$\chi(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega) = \sum_{ss'} A_s(\vec{q}+\vec{G})N_{ss'}(\vec{q}, \omega)A_{s'}^*(\vec{q}+\vec{G}'), \tag{2.8}$$

where

$$A_s(\vec{q}+\vec{G}) = \int d^3r\, \phi_\nu^*(\vec{r})e^{-i(\vec{q}+\vec{G})\cdot\vec{r}}\phi_\mu(\vec{r}-\vec{R}_l) \tag{2.9}$$

and

$$N_{ss'}(\vec{q}, \omega) = \sum_{n_1 \cdots n_4 kk'} c_{n_2\nu}^*(\vec{k})c_{n_1\mu}(\vec{k}+\vec{q})e^{i(\vec{k}+\vec{q})\cdot\vec{R}_l}\, \chi(n_1 k+q, n_2 k, n_3 k', n_4 k'+q; \vec{q}, \omega)$$

$$\times e^{-i(\vec{k}'+\vec{q})\cdot\vec{R}_{l'}}c_{n_4\mu'}^*(\vec{k}'+\vec{q})c_{n_3\nu'}(\vec{k}'). \tag{2.10}$$

The index s stands for the set of indices l, ν, and μ, occurring in Eq. (2.9). We may interpret A_s as a form factor for a generalized "charge-density wave", and $N_{ss'}$ as the polarizability of the charge-density wave s induced by the wave s'.

From Eq. (2.4) for the dielectric function, it follows that the inverse is given by the equation

$$\epsilon^{-1}(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega) = \delta_{\vec{G},\vec{G}'} + \sum_{\vec{G}''} v(\vec{q}+\vec{G})$$

$$\times \chi(\vec{q}+\vec{G}, \vec{q}+\vec{G}''; \omega)\epsilon^{-1}(\vec{q}+\vec{G}'', \vec{q}+\vec{G}'; \omega). \tag{2.11}$$

The separable form of the polarizability, Eq. (2.8), and therefore of the kernel of Eq. (2.11), enables this equation to be solved,[10-13]

$$\epsilon^{-1}(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega) = \delta_{\vec{G},\vec{G}'} + v(\vec{q}+\vec{G})$$

$$\times \sum_{ss'} A_s(\vec{q}+\vec{G})S_{ss'}(\vec{q}, \omega)A_{s'}^*(\vec{q}+\vec{G}'), \tag{2.12}$$

where the screening matrix S is given by

$$S = N(1-VN)^{-1}. \tag{2.13}$$

N is defined by Eq. (2.10) and the Coulomb interaction between the charge-density waves by

$$V_{ss'}(\vec{q}) = \sum_{\vec{G}} A_s^*(\vec{q}+\vec{G})v(\vec{q}+\vec{G})A_{s'}(\vec{q}+\vec{G}')$$

$$= \sum_m e^{-i\vec{q}\cdot\vec{R}_m}\int d^3r \int d^3r'\, \phi_\mu^*(\vec{r}-\vec{R}_l-\vec{R}_m)$$

$$\times \phi_\nu(\vec{r}-\vec{R}_m)v(\vec{r}-\vec{r}')\phi_{\nu'}^*(\vec{r}')\phi_{\mu'}(\vec{r}'-\vec{R}_{l'}). \tag{2.14}$$

If the Bloch waves (or, strictly speaking, the pseudo wave functions) are expanded in terms of plane waves, the dielectric matrix $\epsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega)$ can be inverted directly.[5] By contrast, the method described here is more practical if the Bloch waves are more tight binding than nearly-free-electronlike. Then, the dimension of the matrix which has to be inverted in Eq. (2.13) is conveniently limited.[10,11]

C. Hartree-Fock approximation for polarizability

There remains the problem of calculating the polarizability function χ. In RPA or the time-dependent Hartree approximation, it is given by the noninteracting electron-hole pair[6]

$$\chi(n_1 k+q, n_2 k, n_3 k', n_4 k'+q; \vec{q}, \omega)$$

$$\simeq \delta_{kk'}\, \delta_{n_1 n_4}\delta_{n_2 n_3}\chi_0(n_1 k+q, n_2 k; \omega)$$

$$\equiv \frac{\delta_{kk'}\delta_{n_1 n_4}\delta_{n_2 n_3}2N^{-1}(f_{n_1 k+q} - f_{n_2 k})}{E_{n_1 k+q} - E_{n_2 k} - \omega - i0}, \tag{2.15}$$

where E_{nk} and f_{nk} denote the single-electron energy and occupation number for the Bloch state $n\vec{k}$. The factor of 2 comes from the spin degeneracy. The corresponding RPA term for the polarizability in the Wannier representation, denoted by $N_{ss'}^0(\vec{q}, \omega)$ is given by substituting Eq. (2.15) in Eq. (2.10).

We work in the time-dependent Hartree-Fock approximation and include the exchange effects. For the polarizability χ, we have to add to the bubble diagram (RPA) of Fig. 1(a) a series of ladder diagrams, the first two terms of which are represented by Figs. 1(c) and 1(d). We assume that pure self-energy corrections such as Fig. 1(e) are included in the electron energy E_{nk}. The sum of all these ladder terms can be expressed in terms of an integral equation,[7] as shown in Fig. 2, giving

$$\chi(n_1 k+q, n_2 k, n_5 k'', n_6 k''+q; \vec{q}, \omega) = \chi_0(n_1 k+q, n_2 k; \omega)\delta_{n_1 n_6}\delta_{n_2 n_5}\delta_{kk''} - \tfrac{1}{2}\chi_0(n_1 k+q, n_2 k; \omega)$$

$$\times \sum_{n_3 n_4 k'} v(n_1 k+q, n_2 k, n_3 k', n_4 k'+q)\chi(n_4 k'+q, n_3 k', n_5 k'', n_6 k''+q; \vec{q}, \omega), \tag{2.16}$$

where the exchange Coulomb integral is

$$v(n_1k + q, n_2k, n_3k', n_4k' + q) = \int d^3r \int d^3r' \; \psi^*_{n_1k+q}(\vec{r})\psi_{n_2k}(\vec{r}')v(\vec{r} - \vec{r}')\psi^*_{n_3k'}(\vec{r}')\psi_{n_4k'+q}(\vec{r}). \tag{2.17}$$

By using Eqs. (2.6) and (2.10), the integral equation (2.16) for the polarizability can be put in the Wannier representation

$$N_{ss'}(\vec{q}, \omega) = N^0_{ss'}(\vec{q}, \omega)$$

$$- \tfrac{1}{2}\sum_{s_1 s_2} N^0_{ss_1}(\vec{q}, \omega)V^x_{s_1 s_2}(\vec{q})N_{s_2 s'}(\vec{q}, \omega), \tag{2.18}$$

where N^0 is the RPA expression and V^x is the exchange Coulomb interaction

$$V^x_{ss'}(\vec{q}) = \sum_m e^{-i\vec{q}\cdot\vec{R}_m} \int d^3r \int d^3r' \; \phi^*_\mu(\vec{r} - \vec{R}_l - \vec{R}_m)$$

$$\times \phi_\nu(\vec{r}' - \vec{R}_m)v(\vec{r} - \vec{r}')\phi^*_{\nu'}(\vec{r}')\phi_{\mu'}(\vec{r} - \vec{R}_{l'}). \tag{2.19}$$

By comparing this expression with Eq. (2.14), it is clear that $V^x_{ss'}$ is the exchange correspondence of $V_{ss'}$.

Formally inverting the matrix in Eq. (2.18) yields

$$N = N^0(1 + \tfrac{1}{2}V^xN^0)^{-1}, \tag{2.20}$$

and by Eq. (2.13),

$$S = N^0[1 - (V - \tfrac{1}{2}V^x)N^0]^{-1}. \tag{2.21}$$

The effect of the electron-hole attraction is to modify the matrix of the Coulomb-interaction energy between the charge-density waves. For studying the properties of the exciton, this represents a generalization of the contact-interaction approximation.[17]

The Coulomb interaction in the exchange integral is subject to further screening. In the effective-mass limit, it is appropriately screened by the macroscopic dielectric constant.[7] However the way in which we shall make use of the exchange term, the local terms are more important than the long-range terms. Thus, it is more appropriate to use the unscreened Coulomb interaction.

A different method of approximately including the exchange and correlation effects in the dielectric function of a crystal has been suggested.[10,18,19] In Eq. (2.4) for the dielectric function, the polarizability is taken in RPA, and $v(\vec{q}+\vec{G})$ is modified by a factor which has been taken as approximately accounting for the exchange and correlation effects of the dielectric function in the homogeneous electron gas.[20,21] Such an approximation applied to a crystal is not as firmly based on the method we presented above. For semiconductors with small direct gaps, the Hub-

bard-type approximation yields incorrect bound-exciton contributions to the dielectric constant.[22] We have tested the Hubbard approximation[20] for the optical spectrum of diamond and found it to be unsatisfactory (see Sec. IV).

D. Long-wavelength limit

To obtain the optical constant, we have to take the $q \to 0$ limit. Now, the inverse dielectric matrix involves divergent Coulomb factors $v(q)$ and it is convenient to isolate them. Following Ref. 2, we define $\hat{\chi}(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega)$ as the sum of all polarization processes not involving the long-range part of the Coulomb interaction $v(q)$. Call it the susceptibility function. It is given in terms of the polarizability $\tilde{\chi}$ by

$$\hat{\chi}(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega) = \tilde{\chi}(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega)$$

$$+ \sum_{G''\neq 0} \tilde{\chi}(\vec{q}+\vec{G}, \vec{q}+\vec{G}''; \omega)$$

$$\times v(\vec{q}+\vec{G}'')\hat{\chi}(\vec{q}+\vec{G}'', \vec{q}+\vec{G}'; \omega). \tag{2.22}$$

Furthermore, the inverse dielectric function has a term

$$\epsilon^{-1}(\vec{q}, \vec{q}; \omega) = [1 - v(q)\hat{\chi}(\vec{q}, \vec{q}; \omega)]^{-1}. \tag{2.23}$$

It has been shown[23] that, for insulating crystals, $\hat{\chi}(\vec{q}, \vec{q}; \omega)$ is proportional to q^2 for small q. By Eq. (2.22), for cubic crystals, as $q \to 0$,

$$\hat{\chi}(\vec{q}, \vec{q}; \omega) = \hat{\chi}^2(\omega)q^2 + O(q^4). \tag{2.24}$$

Hence, by Eq. (2.2), the macroscopic dielectric constant is

$$\epsilon(\omega) = 1 - \hat{\chi}^2(\omega)4\pi e^2/\Omega_0. \tag{2.25}$$

This expression holds for all orders of perturbation theory in the Coulomb interactions.

The inversion procedure described in Sec. II C
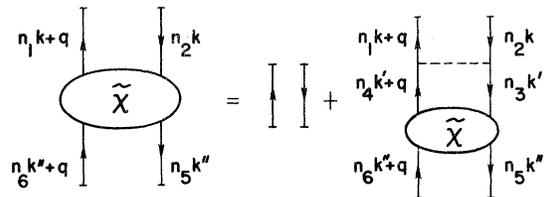
FIG. 2. Integral equation for the ladder sum of exchange processes.

can be used to solve Eq. (2.22) for the susceptibility $\hat{\chi}$. In the Wannier representation, we obtain the dielectric constant

$$\epsilon(\omega) = 1 - 4\pi e^2 \Omega_0^{-1} \sum_{ss'} f_s^\alpha S_{ss'}(\omega) f_{s'}^{\alpha *}, \qquad (2.26)$$

where

$$f_s^\alpha = \int d^3r \, \phi_\nu^*(\vec{r}) r_\alpha \phi_\mu(\vec{r} - \vec{R}_l), \qquad (2.27)$$

α denoting a principal axis of the cubic crystal. S is given by Eq. (2.21), with

$$N_{ss'}^0(\omega) = 2N^{-1} \sum_{n_1 \cdots n_4 k} c_{n_2\nu}^*(\vec{k}) c_{n_1\mu}(\vec{k})$$

$$\times e^{i\vec{k}\cdot\vec{R}_l} \frac{f_{n_1 k} - f_{n_2 k}}{E_{n_1 k} - E_{n_2 k} - \omega - i0}$$

$$\times e^{-i\vec{k}\cdot\vec{R}_{l'}} c_{n_4\mu'}^*(\vec{k}) c_{n_3\nu'}(\vec{k}), \qquad (2.28)$$

$$V_{ss'} = \sum_{G' \neq 0} A_s^*(\vec{G}') v(G') A_{s'}(\vec{G}') \qquad (2.29a)$$

$$= \sum_m \int d^3r \int d^3r' \, \phi_\mu^*(\vec{r} - \vec{R}_l - \vec{R}_m) \phi_\nu(\vec{r} - \vec{R}_m)$$

$$\times v(\vec{r} - \vec{r}') \phi_{\nu'}^*(\vec{r}') \phi_{\mu'}(\vec{r}' - \vec{R}_{l'}), \qquad (2.29b)$$

and

$$V_{ss'}^x = \sum_m \int d^3r \int d^3r' \, \phi_\mu^*(\vec{r} - \vec{R}_l - \vec{R}_m) \phi_\nu(\vec{r}' - \vec{R}_m)$$

$$\times v(\vec{r} - \vec{r}') \phi_{\nu'}^*(\vec{r}') \phi_{\mu'}(\vec{r} - \vec{R}_{l'}). \qquad (2.30)$$

E. Current representation

So far we have treated the dielectric constant in terms of the density response. We could equally well have used the current response.[3,10] Instead of the matrix elements of the density $e^{-i(\vec{q}+\vec{G})\cdot\vec{r}}$, we have those of the current

$$\vec{j}(\vec{q}) = (\vec{p} e^{-i\vec{q}\cdot\vec{r}} + e^{-i\vec{q}\cdot\vec{r}} \vec{p})/2m, \qquad (2.31)$$

$\vec{p}$ being the electron momentum. The formulas are the same as before, except that

$$A_s(\vec{q}+\vec{G}) = (\vec{q}+\vec{G}) \cdot \int d^3r \, \phi_\nu^*(\vec{r}) \vec{j}(\vec{q}+\vec{G}) \phi_\mu(\vec{r} - \vec{R}_l), \qquad (2.32)$$

$$f_s^\alpha = m^{-1} \int d^3r \, \phi_\nu^*(\vec{r}) \nabla_\alpha \phi_\mu(\vec{r} - \vec{R}_l), \qquad (2.33)$$

and

$$\tilde{\chi}_0(n_1 k + q, n_2 k; \omega) = 2N^{-1}(f_{n_1 k + q} - f_{n_2 k})/$$

$$[(E_{n_1 k + q} - E_{n_2 k} - \omega - i0)$$

$$\times (E_{n_1 k + q} - E_{n_2 k})^2]. \qquad (2.34)$$

The current form of dielectric response has the

advantage that the short-wave-vector properties, as in Eq. (2.24), are more evident, but the density is simpler to interpret physically, as we shall see in Sec. III.

III. LOCAL REPRESENTATION FOR DIAMOND

A. Band-structure interpolation

Since Wannier functions are not yet available, we have used Hall's method[24] to construct approximate local orbitals for the diamond lattice, which emphasize the covalent nature. The s-like orbital $R_s(r)$ and the p-like orbitals $\hat{r}R_p(r)$ are combined to form the hybridized orbitals according to Pauling's prescription,[25]

$$\chi_{\vec{\nu}}^+(\vec{r}) = (4\sqrt{\pi})^{-1}[R_s(r) + \sqrt{3}(\vec{\nu}\cdot\vec{r}/r)R_p(r)], \qquad (3.1)$$

$\vec{\nu}$ being one of the tetrahedral vectors $(1, 1, 1)$, $(1, \bar{1}, \bar{1})$, $(\bar{1}, 1, \bar{1})$, $(\bar{1}, \bar{1}, 1)$. The pair of hybridized orbitals from neighboring atoms directed along the line joining their positions are added and subtracted to form bonding and antibonding orbitals, respectively,

$$\phi_{\vec{\nu}\pm}^+(\vec{r}) = N_\pm^{-1}[\chi_{\vec{\nu}}^+(\vec{r}) \pm \chi_{\vec{\nu}}^+(\vec{r} - b\vec{\nu})], \qquad (3.2)$$

$b = \frac{1}{4}a$, a being the lattice constant, and $N_\pm$ the normalization constants (see Fig. 3). These orbitals form the Bloch waves by Eq. (2.6).

Linear combinations of the four bonding orbitals are used to construct the valence bands and linear combinations of the four antibonding orbitals to construct the conduction bands. Cohan et al.[26] have found the effects of bonding and antibonding cross terms on the band structure of diamond to be quite small. We neglect the nonorthogonality of the orbitals on different sites. As is well known, the orthogonality correction may be taken into account by a unitary transformation of the one-electron Hamiltonian, giving rise to a modified band struc-
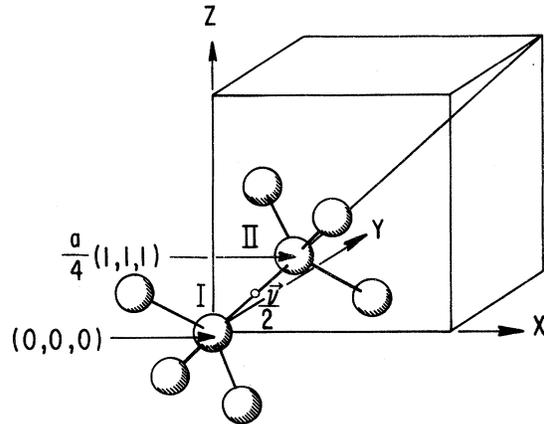
FIG. 3. Unit cell of diamond.

ture.[27] Using our diamond parameters, we found the energy difference of the dominant optical transition ($X_4 - X_1$) and the shape of the lowest conduction (Δ_1, $\Delta_{2'}$) and highest valence bands (Δ_5) only slightly modified by orthogonal corrections.

We include the overlap integrals of the one-electron Hamiltonian with respect to bonding and antibonding orbitals up to the third-nearest neighbors and treat them as parameters to fit the energy bands calculated by Painter, Ellis, and Lubinsky[28] from first principles. This procedure yields two sets of 4×4 secular equations, one for the conduction bands and one for the valence bands, determining, for each $\vec{k}$, the energy and coefficients $c_{nv}(\vec{k})$ of the orbitals by Eq. (2.6). The results of this third-nearest-neighbor energy-band fit are shown in Fig. 4 in comparison with the bands of Painter et al. The overlap parameters are given in Table I.

Although for the purpose of the interpolation of the energy bands the use of bonding and antibonding orbitals is completely equivalent to using the atomic orbitals,[27] i.e., orbitals located on the atomic sites, we have found the former very much more convenient for evaluating the dielectric properties.

B. Gaussian representation for atomic orbitals

The atomic orbitals are expressed in terms of Gaussian functions,[29]

$$R_s(r) = \sum_i a_i e^{-\alpha_i r^2} \qquad (3.3)$$

and
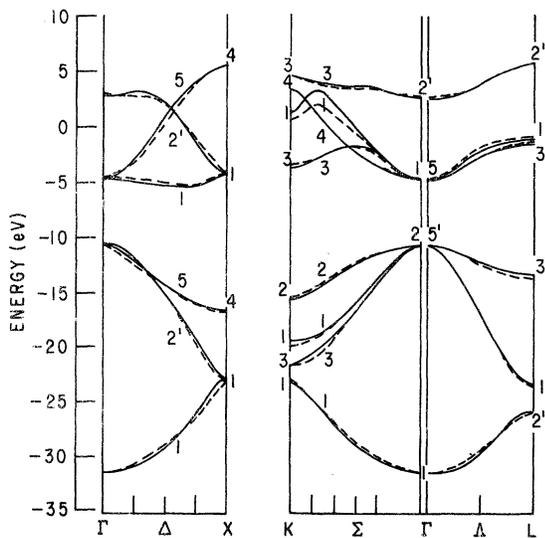
FIG. 4. Band structure of diamond. Solid line, Painter et al., Ref. 28; dashed line, third-neighbor LCAO interpolation.

$$xR_p(r) = \sum_j c_j(e^{-\beta_j(\vec{r}-\vec{d}_{jx})^2} - e^{-\beta_j(\vec{r}+\vec{d}_{jx})^2}), \qquad (3.4)$$

where

$$\vec{d}_{jx} = (1, 0, 0)\kappa/\beta_j^{1/2}. \qquad (3.5)$$

The p-type orbitals are the so-called Gaussian lobe orbitals.[29] The displacement vector $\vec{d}$ is designed to reproduce the angular dependence of the wave function. In diamond this angular dependence is well reproduced for values of κ less than 0.1.[30] We set κ equal to 0.02. The Gaussian orbitals are convenient for LCAO band calculations.[30,31] We have also found them invaluable in the dielectric-response calculation, as we shall show later.

In principle, one can fit the atomic orbitals to the Hartree-Fock orbitals of an isolated atom[32] and proceed with an LCAO calculation, as was done by Chaney et al.[31] To keep our dielectric calculation within reasonable limits, we have truncated the Hamiltonian to the third-nearest neighbors, and obtained its elements by interpolation to a calculated band structure.[27] The third-nearest

TABLE I. Overlap parameters of the third-nearest-neighbor energy-band model.

Parameter	Bonding (eV)	Antibonding (eV)
Nearest neighbors parallel:		
$(a\|H\|a)^a$	16.3156	35.3719
antiparallel:		
$(a\|H\|b)$	−1.4844	−0.0490
Second-nearest neighbors parallel:		
$(a\|H\|a')$	0.8203	−0.6724
antiparallel:		
$(a\|H\|b')$	−0.5844	0.5479
Third-nearest neighbors parallel:		
$(a\|H\|a'')$	−0.3172	0.2318
antiparallel:		
$(a\|H\|b'')_1$	−0.0622	−0.1172
$(a\|H\|b'')_2$	0.0318	−0.0339
$(a\|H\|b'')_3$	−0.4	−0.3
$(a\|H\|b'')_4$	−0.055	−0.0989
$(a\|H\|b'')_5$	0.1750	0.1495

^{a} $(a\|H\|a)$ denotes the matrix element of the Hamiltonian with respect to orbitals a, etc.

neighbor is not far enough for convergence.[31] Thus, our wave functions, based on isolated-atom orbitals, would be rather poor. Instead, we take advantage of some arbitrariness in the local orbitals[33] and vary the parameters in Eqs. (3.3) and (3.4) to improve the wave functions. As discussed in Sec. II E, the dielectric constant can be calculated in terms of either the density response or the current response. To achieve a certain measure of internal self-consistency, we take two Gaussians for each atomic orbital and adjust the parameters in the Gaussians such that both the density and the current representation give the same dielectric constant $\bar{\epsilon}_{RPA}(\omega)$ in RPA without the local-field correction. The resulting Gaussian parameters are listed in Table II. For comparison, we also list the parameters of a fit to the $2s$ and $2p$ Hartree-Fock (HF) orbitals of the carbon atom.[32]

In evaluating Eqs. (2.26)–(2.30) for the dielectric constant, we include up to the nearest neighbor overlaps between the bonding and antibonding orbitals, having found that the contributions of the second-nearest neighbors of the bonds to f_s^{α} and A_s are about (5–10)%. The dimension of the matrices N^0, V, and S is then 28×28. In terms of the directed hybridized orbitals centered at the atomic site (Fig. 5), the overlaps included are given in Fig. 6.

C. Density-density interaction matrix

The product of two Gaussians centered at different sites is another Gaussian,[34]

$$e^{-\alpha(\vec{r}-\vec{A})^2} e^{-\beta(\vec{r}-\vec{B})^2} = K e^{-\gamma(\vec{r}-\vec{C})^2} , \qquad (3.6)$$

with

$$\gamma = \alpha + \beta, \quad \vec{C} = (\alpha\vec{A} + \beta\vec{B})/(\alpha + \beta) ,$$

and

$$K = \exp[-(\vec{A} - \vec{B})^2 \alpha\beta/(\alpha + \beta)] .$$

This property not only facilitates the evaluation of the multicenter integrals, which can be carried out analytically,[34] but also provides us with a simple physical interpretation. The "charge density" $\phi_\nu^*(\vec{r})\phi_\mu(\vec{r} - \vec{R}_l)$ is a sum of Gaussian orbitals, and as a result, the Coulomb energy $V_{ss'}$ and the exchange energy $V_{ss'}^x$ are just the sum of Coulomb energies between pairs of charges with Gaussian distributions. For example, if only the s orbitals are considered, the Coulomb matrix $V_{ss'}$ is due to the pair interaction of charges centered at the atomic sites and the bond sites (midway between two nearest-neighbor atomic sites) as shown in Fig. 7. If the p orbitals are included, each of the charges 1–6 in Fig. 7 becomes a cluster of charges centered at closely spaced sites. The equivalent charges produced by the s and p orbitals are approximately a combination of monopoles, dipoles, and quadrupoles.

The symmetry of the diamond lattice is used to simplify the matrices V and N^0. In the nearest-neighbor overlap model each of the 28×28 matrices has only 26 independent elements.

For V^x the dominant contribution comes from those exchange processes within the same bond; i.e., in the sum of Eq. (2.30), the terms kept are either

$$\vec{R}_m = \vec{R}_{l'} \quad \text{and} \quad \nu = \mu' , \qquad (3.7a)$$

or

$$\vec{R}_m = -\vec{R}_l \quad \text{and} \quad \mu = \nu' . \qquad (3.7b)$$

There are also terms of these types in the direct Coulomb energy $V_{ss'}$.

D. Polarizability matrix

The k sums in Eq. (2.28) for the real and imaginary parts of N^0 are calculated separately by the Gilat-Raubenheimer method.[35] The sum over the full Brillouin zone is first mapped onto a sum over

TABLE II. Coefficients of $2s$ and $2p$ Gaussians.

Parameter	Current-conservation fit (a.u.)	Fit to HF atomic orbitals (a.u.)
a_1	1.3570	0.4685
a_2	0.0887	0.4605
α_1	0.502	0.502
α_2	0.213	0.155
c_1	8.5691	15.4981
c_2	7.1294	13.0291
β_1	0.25	1.55
β_2	0.445	0.337

the irreducible Brillouin zone, which has $\frac{1}{48}$ of the volume of the full Brillouin zone. The irreducible part is then divided into small cubes, the largest of which has a width of 0.2335 a.u. The energies inside a cube are obtained by linear interpolation, but the matrix elements of N^0 are assumed to be constant within a cube.

IV. OPTICAL SPECTRUM OF DIAMOND

Using the results of Sec. III, we calculate the real and imaginary parts of the dielectric constant in three different approximations: (a) denoted by $\bar{\epsilon}_{RPA}(\omega)$ with S in Eq. (2.26) replaced by N^0, thus neglecting both the local-field and the excitonic (or exchange) effects; (b) denoted by $\epsilon_{RPA}(\omega)$, with the local-field correction but without the exchange term, V^x; and (c) denoted by $\epsilon(\omega)$, including both the local-field and the full Hartree-Fock terms. The results for the imaginary part of $\epsilon(\omega)$ are plotted in Fig. 8 and are compared with the experimental curve.[36]

The most interesting feature is the remarkable change due to the local-field and the excitonic effects. Consider the main peak close to 12 eV. The local-field correction within RPA reduces the strength of the peak, considerably more on the low-energy side, thus giving the appearance of moving the peak to a higher energy by about 0.5 eV. The exchange correction again enhances the strength of the peak with the bias on the low-energy side, thus apparently moving the peak to a lower energy by about 1.5 eV. Because the enhancement occurs above the fundamental band gap, the exchange term gives the continuum-exciton correction. When the exchange term is calculated in the Hubbard approximation (Sec. II C), the peak is moved back only a little, not even back to the position of $\bar{\epsilon}$

Our calculation shows a discernible feature at about 8 eV, which is found in Ref. 37. The peak at 14 eV is probably overestimated because there our approximate band structure gives too large a joint density of states.

The f sum calculated for our $\epsilon_2(\omega)$ up to the maximum bandwidth of our model is 1515 compared with the exact theoretical value of $\frac{1}{2}(\pi\omega_p^2)$ = 1550. Our f sum is clearly larger than the measured value in the experimental range. The static dielectric constant $\epsilon(0)$ for the time-dependent Hartree-Fock approximation is 6.0 compared with the RPA value without local field of 5.1 and with the experimental value of 5.85.[38]

The rather good agreement of the calculated main peak in the Hartree-Fock approximation and the experiment may be fortuitous. Lubinsky et al.[37] have calculated $\bar{\epsilon}_{RPA}(\omega)$ from their band structure[27] and wave functions, within RPA neglecting the local-field effect, and have gotten much better agreement with experiment than our $\bar{\epsilon}_{RPA}(\omega)$. The only remaining discrepancy in their calculation is the smaller calculated value on the lower-energy side of the main peak. Their work does not void our findings of the local-field and exciton effects.

Van Vechten and Martin[5] have calculated the RPA dielectric matrix in a plane-wave expansion for the electron wave functions including reciprocal-lattice vectors $\vec{G}$ and $\vec{G}'$ through the set (2,2,2) for diamond. The resulting 59×59 matrix $\epsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega)$ was inverted directly. The effect of the RPA local field on the main peak is in qualitative agreement with our calculation. However, Van Vechten and Martin ascribe the remaining discrepancy with experiment to the dynamical correlations. In terms of the many-body perturbation theory, the most important correction to RPA is

FIG. 5. Hybridized orbital at an atomic site.
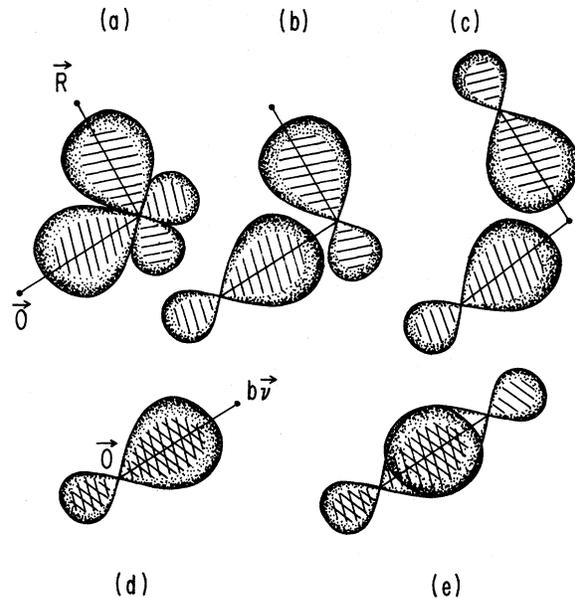
FIG. 6. Overlap terms of the hybridized orbitals included in the dielectric calculation.

the exchange effect, which we attempt to take into account. The correlation effect in the proper polarization part should be smaller than the exchange term for an insulating crystal with a large band gap.

Louie et al.[39] have recently calculated the local-field effect within RPA in silicon in the same way as Van Vechten and Martin did for diamond. Again the local-field effect lowers the strength on the low-energy side of the prominent peak, but not as much as in diamond. This is expected since the wave functions in Si are more extended than in diamond. In Si, there also clearly remains the discrepancy of the calculated RPA optical absorption with experiment, being too weak on the lower-energy side of the prominent peak and too strong above the peak for a range of several eV. This can again be explained by the exchange term creating the continuum-exciton effect.

V. DISCUSSION

Let us try to understand the results of our calculation of the local-field and excitonic effects in a simple and fairly general way. To avoid cumbersome matrix notation, we treat the matrix N^0 and others like it as scalars. If we write the inverse of N^0,

$$(N^0)^{-1}(\omega) = W_1(\omega) + i W_2(\omega), \qquad (5.1)$$

where W_1 and W_2 are real, then,

$$N^0(\omega) = (W_1 - i W_2)/(W_1^2 + W_2^2) \qquad (5.2)$$

and

$$S = [(W_1 - V + \tfrac{1}{2}V^x) - i W_2]/[(W_1 - V + \tfrac{1}{2}V^x)^2 + W_2^2]. \qquad (5.3)$$

If the imaginary part of $N^0(\omega)$ is approximately

a Lorentzian centered about a zero of $W_1(\omega)$, $\sim \omega - \Delta$, from Eq. (2.28), then the imaginary part of $S(\omega)$ is a Lorentzian with its peak moved in energy by $V - \tfrac{1}{2}V^x$. The change in height of the peak depends on the change of $W_2(\omega)$. In RPA, the peak position is changed by just V, which is positive, being the Coulomb-interaction energy between the charge-density waves. The electron-hole attraction V^x moves the Lorentzian peak to a lower energy.

In many materials, when the imaginary part of $\overline{\epsilon}_{RPA}$ calculated without the local-field and excitonic effects is compared with experiment, the low-energy side of the prominent absorption peak above the band gap is too small.[40] The local-field effect within RPA would increase this discrepancy with experiment. The continuum-exciton effect, as formulated in Sec. IIC, would account for it.

It is interesting to note that the Lorentz-Lorenz relation can be put in the same form as our scheme above with

$$\epsilon = 1 + 4\pi\alpha/(1 - \tfrac{4}{3}\pi\alpha). \qquad (5.4)$$

If the atomic polarizability α is written

$$-3/4\pi\alpha = W_1(\omega) + i W_2(\omega), \qquad (5.5)$$

in analogy with Eq. (5.1), then the dielectric constant is

$$\epsilon = 1 - 3/(W_1 + 1 + i W_2). \qquad (5.6)$$

The effect of the Lorentz-Lorenz relation is also to move a peak in the polarizability to a lower energy.

A dipolar expansion of the local-field effect within RPA yields a very similar formula to Lorentz-Lorenz with the addition of a self-interaction

FIG. 7. Distribution of equivalent charge densities in the Gaussian representation.
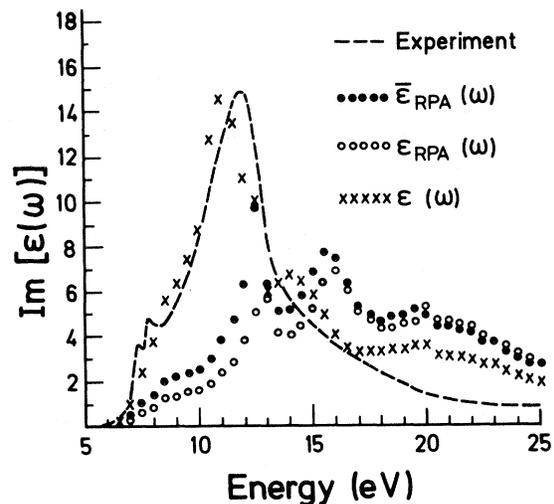
FIG. 8. Imaginary part of the dielectric constant. See the text for a detailed explanation of the legends.

term of a dipole.[3] This formula will move a Lo-
rentzian peak to a higher energy. To obtain the
classic Lorentz-Lorenz relation, it is necessary
to exclude the self-interaction of a dipole by re-
moving from the time-dependent Hartree approxi-
mation the self-interaction term of an electron
on the ion.[41,42] This overwhelms the local-field
effect in RPA, resulting in Eq. (5.6).

Finally, we wish to comment that the Lorentz-
Lorenz relation is not applicable to the covalent
crystals, even though it has the same qualitative
feature as our formulation in terms of the Wan-
nier functions. Sinha, Gupta, and Price[42] have
derived the Lorentz-Lorenz formula by assuming
the following separable form for the polarizability:

$$\chi(\vec{q}+\vec{G}, \vec{q}+\vec{G}'; \omega) = \sum_{\alpha\beta\kappa\kappa'} (\vec{q}+\vec{G})_\alpha f_\kappa(\vec{q}+\vec{G})$$

$$\times a_{\kappa\alpha,\kappa'\beta}(\vec{q}, \omega) f^*_{\kappa'}(\vec{q}+\vec{G}')(\vec{q}+\vec{G}')_\beta .$$
$$(5.7)$$

The index κ need not have the same meaning as
the Wannier representation index s. However, let
us consider the question whether a first-principles
calculation in the Wannier representation can yield
the particular separable form, Eq. (5.7), which
leads to a Lorentz-Lorenz relation. In terms of
Wannier functions, the form factor of the charge-
density wave is, from Eq. (2.32), of the form

$$A_s(\vec{q}+\vec{G}) = \sum_\alpha (\vec{q}+\vec{G})_\alpha F_{s\alpha}(\vec{q}+\vec{G}). \qquad (5.8)$$

To arrive at the ansatz, Eq. (5.7), we further need

$$F_{s\alpha}(\vec{q}+\vec{G}) = W_{s\alpha} f_s(\vec{q}+\vec{G}), \qquad (5.9)$$

where $W_{s\alpha}$ is independent of $\vec{G}$. This is possible if,
for instance, the valence-band Wannier function is
p-like, the conduction-band function s-like, and
both are tightly bound so that intersite overlap can
be neglected. For the covalent crystals such as
diamond, with the Wannier functions having the
symmetries of the bonding and antibonding orbitals
as discussed in Sec. III, such a form as Eq. (5.9) is
not a satisfactory approximation. Let us conclude
by mentioning that the polarizability α in the Lo-
rentz-Lorenz relation (5.4) has to be independent
of density, a condition which can be examined ex-
perimentally by measuring the pressure depen-
dence of the refractive index. This density inde-
pendence of α was found far from fulfilled in dia-
mond and even less fulfilled in other covalent
crystals like Si.[38]

*Work supported in part by the NSF under Grant No.
 GH-38345.
†Supported in part by the Deutsche Forschungsgemein-
 schaft and on leave during 1972–74 from the Max-
 Planck Institut für Festkörperforschung, Stuttgart,
 Germany.
‡On leave during 1974–75 at IBM Research Center,
 Yorktown Heights, N. Y. 10598.
[1] A brief account of our work is given in W. Hanke and
 L. J. Sham, Phys. Rev. Lett. **33**, 582 (1974).
[2] V. Ambegaokar and W. Kohn, Phys. Rev. **117**, 423
 (1960).
[3] S. L. Adler, Phys. Rev. **126**, 413 (1962); N. Wiser,
 ibid. **129**, 62 (1963).
[4] The Lorentz-Lorenz relation has been used to study
 the local-field effect in the optical spectrum of the
 ionic crystal CsCl, by T. K. Bergstresser and G. W.
 Rubloff, Phys. Rev. Lett. **30**, 794 (1973); and has been
 used by S. R. Nagel and T. A. Whitten, Phys. Rev.
 B **11**, 1623 (1975) to discuss local-field corrections
 in inelastic electron scattering.
[5] J. A. Van Vechten and R. M. Martin, Phys. Rev. Lett.
 28, 446 (1972).
[6] H. Ehrenreich and M. H. Cohen, Phys. Rev. **115**, 786
 (1959).
[7] L. J. Sham and T. M. Rice, Phys. Rev. **144**, 708 (1966).

[8] R. J. Elliott, Phys. Rev. **108**, 1384 (1957).
[9] J. C. Phillips, Solid State Phys. **18**, 55 (1966).
[10] For a review of this and other related concepts, see,
 for example, L. J. Sham, in Dynamical Properties of
 Solids, edited by G. K. Horton and A. A. Maradudin
 (North-Holland, Amsterdam, 1974), Vol. 1, Chap. 5.
[11] W. Hanke, in Phonons, edited by M. A. Nusimovici
 (Flammarion, Paris, 1971), p. 294; Proceedings of
 the International Conference on Inelastic Scattering
 of Neutrons, Grenoble, 1972 (IAEA, Vienna, 1972),
 p. 3; and Phys. Rev. B **8**, 4585 (1973); and **8**, 4591
 (1973).
[12] L. J. Sham, Phys. Rev. B **6**, 3584 (1972).
[13] R. M. Pick, in Phonons, edited by M. A. Nusimovici
 (Flammarion, Paris, 1971), p. 20.
[14] J. Des Cloizeaux, Phys. Rev. **129**, 554 (1963); **135**,
 A685 (1964); **135**, A698 (1964).
[15] W. Kohn, Phys. Rev. B **7**, 4388 (1973).
[16] L. J. Sham, Phys. Rev. **188**, 1431 (1969).
[17] G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167
 (1954); Y. Toyozawa, M. Inoue, T. Inui, M. Okazaki,
 and E. Hanamura, J. Phys. Soc. Jpn. **22**, 1337 (1967).
[18] L. J. Sham, Proc. R. Soc. A **283**, 33 (1965).
[19] D. L. Price, S. K. Sinha, and R. P. Gupta, Phys. Rev.
 B **9**, 2573 (1974).
[20] J. Hubbard, Proc. R. Soc. A **243**, 336 (1958).

[21] P. Vashishta and K. S. Singwi, Phys. Rev. B **6**, 875 (1972).

[22] L. J. Sham and T. V. Ramakrishnan, in *Physics of Semimetals and Narrow Gap Semiconductors*, edited by D. L. Carter and R. T. Bate (Pergamon, New York, 1971), p. 219.

[23] W. Kohn, Phys. Rev. **110**, 857 (1958); L. J. Sham, *ibid*. **150**, 720 (1966).

[24] G. G. Hall, Philos. Mag. **43**, 338 (1952).

[25] L. Pauling, J. Am. Chem. Soc. **53**, 1367 (1931).

[26] N. V. Cohan, D. Pugh, and R. H. Tredgold, Proc. Phys. Soc. Lond. **82**, 65 (1963).

[27] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

[28] G. S. Painter, D. E. Ellis, and A. R. Lubinsky, Phys. Rev. B **4**, 3610 (1971).

[29] J. L. Whitten, J. Chem. Phys. **44**, 359 (1966).

[30] R. N. Euwema, D. L. Wilhite, and G. T. Surratt, Phys. Rev. B **7**, 818 (1973).

[31] R. C. Chaney, C. C. Lin, and E. E. Lafon, Phys. Rev.

B **3**, 459 (1971).

[32] A. Jucy, J. Phys. USSR **11**, 49 (1947).

[33] P. W. Anderson, Phys. Rev. Lett. **21**, 13 (1968).

[34] I. Shavitt, in *Methods in Computational Physics*, edited by B. Alder, S. Feshbach, and M. Rotenberg (Academic, New York, 1963), p. 1.

[35] G. Gilat and L. J. Raubenheimer, Phys. Rev. **144**, 390 (1966).

[36] R. A. Roberts and W. C. Walker, Phys. Rev. **161**, 730 (1967).

[37] A. R. Lubinsky, D. E. Ellis, and G. S. Painter, Phys. Rev. B **6**, 3950 (1972).

[38] M. Kastner, Phys. Rev. B **7**, 5237 (1973).

[39] S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. Lett. **34**, 155 (1975).

[40] J. L. Freeouf, Phys. Rev. B **7**, 3810 (1973).

[41] Y. Onodera, Prog. Theor. Phys. **49**, 37 (1973).

[42] S. K. Sinha, R. P. Gupta, and D. L. Price, Phys. Rev. B **9**, 2564 (1974).