Many-particle effects in the optical spectrum of a semiconductor

W. Hanke and L. J. Sham*
Max-Planck-Institut für Festkörperforschung, 7 Stuttgart-80, Federal Republic of Germany
(Received 25 October 1979)

A local-orbital treatment of the two-particle Green's function for the electron-hole interaction is presented which takes into account both screened electron-hole attraction and its exchange counterpart. They give rise to the excitonic effects including Frenkel and intermediate coupling regimes and to the random-phase approximation local-field effects, respectively. An alternative formulation based on the Kohn-Sham density-functional scheme is also given and numerically tested. Quantitative calculations of the absorption and modulation spectra in Si show that electron-hole interaction effects significantly modify the absorption line shape and give rise to shifts of critical-point structure up to 0.2 eV. A model analysis indicates that deviations of the one-particle spectra of column IV, III-V, and II-VI semiconductors from experiment should similarly be accounted for.

I. INTRODUCTION

One-electron band theory is known to provide a good starting point for describing the optical excitations in insulating and semiconducting crystals[1]: Light is absorbed in exciting an electron from a full to a vacant band across the energy gap—a process which may be considered as creation of an electron-hole pair. To date, many elaborate first-principles calculations of the energy bands and wave functions have been used to evaluate the optical spectra.[2] Conversely, the measured spectra (and other measured properties) have been taken within this framework of noninteracting electron-hole excitations, together with a band-structure method such as the pseudopotential[3] or the $\vec{k}\cdot\vec{p}$ method,[4] to determine the one-electron band structure.

It is by now well established[1] that the absorption spectrum determined from the noninteracting electron-hole pairs is modified by interaction effects, which consist primarily of the attraction between the electron and the hole and its exchange counterpart. In the neighborhood of the fundamental threshold (the M_0-type critical point[5]), bound exciton lines and continuum excitons drastically change the spectrum.[6] At the M_1-type critical point, the saddle-point exciton first suggested by Phillips[7] profoundly modify the Van Hove singularity structure.

Efforts towards more quantitative evaluation of the exciton effects have been of two types:

(1) The effective-mass approximation (EMA). Its use for the excitons near the threshold is common.[6] For saddle excitons, Velicky and Sak[8] and Kane[9] have constructed the effective masses appropriate to the M_1 critical point and used as the electron-hole attraction the Coulomb interaction screened by the macroscopic dielectric constant, as in the usual EMA.

(2) The Koster-Slater Method.[10] Here the electron-hole excitation is evaluated in the tight-binding approximation and its attraction is approximated by an on-site contact interaction. Velicky and Sak[8] have used this method to demonstrate the enhancement of absorption at M_1 and depression at M_2-type critical points.

Toyozawa *et al.*[11] and Hermanson,[12] also taking a short-range interaction, demonstrated more generally a "metamorphism" of the Van Hove singularities. There also exists work, where experimental evidence for deviations from the single-particle picture is described in a parametrized contact interaction exciton model.[13]

The role of the exchange to the electron-hole attraction (the so-called exciton-exchange interaction) has been reviewed by Cho.[14] It has small but important observable consequences, such as the exciton spectrum fine structure and the splitting of longitudinal and transverse excitons.[15-18] Usually this exciton exchange is studied by neglecting the $\vec{k}$ dependence of the Bloch functions which generate the exciton states in consideration.[14]

The purpose of this paper is to present a formalism for calculating the optical response taking account of the most general form of electron-hole interactions. We give a practical solution which avoids the limitations of EMA or Koster-Slater methods discussed above and use it for a quantitative calculation of the interaction effects on the optical excitations of a semiconductor.[19] In particular, the exchange term is evaluated without resorting to a $\vec{k}$-independent interaction.

The dielectric response[20] from which the observed spectrum can be calculated is given by the two-particle Green's function.[21] Our starting point is the equation of motion for this two-particle Green's function S, known as the Bethe-Salpeter equation; an integral equation of the form

$$S = S_0 + S_0 I S , \tag{1.1}$$

where S_0 is the noninteracting electron-hole pair, including, however, all the band-structure effects and the many-body effects on the *one-electron* energy and wave function, and I is the electron-hole interaction. We endeavor to solve this equation by expressing it in terms of local orbitals, thus converting it to a matrix equation. Our aim is, therefore, a generalization of the tight-binding Koster-Slater method.

In the limit of a large Wannier exciton near the fundamental threshold, it has been shown that Eq. (1.1), despite its multifarious interaction terms,[21] reduces to the EMA. The key to such a simplification is the validity of the effective-mass representation of the electron and hole energies near the fundamental gap at M_0. This allows for a simple representation of S_0 which then enables (1.1) to be converted into a Schrödinger-type equation. The importance of confining the Bloch states to a small neighborhood of $\vec{k}$ space means the dominance in I of the long-range Coulomb attraction screened by the macroscopic dielectric constant.

In the other portions of the optical spectrum, such simplifications of S_0 and I are not possible. A typical example is the E_1 peak in a whole series of semiconductors, which is contributed by the nearly parallel conduction and valence bands in the $\langle 111 \rangle$ directions.[3] Thus, the amount of phase space involved in S_0 is large.

The procedure for the solution of the Bethe-Salpeter equation in the local-orbital representation is detailed in Sec. II. The interaction term I is taken to include the screened electron-hole attraction $-V^s$ and the unscreened exchange term V. Thus, our work includes the overall band effect, the excitonic effect, and the exciton-exchange correction.

If the electron-hole attraction, $-V^s$ is omitted, the Bethe-Salpeter equation with just the unscreened exciton exchange V reduces to the usual random-phase approximation (RPA). This V term gives rise to the so-called local-field effect in crystals.[22] Recently, there have been many calculations of this local-field effect.[23-33] The first one, which attempted to include the excitonic effect in addition to the local-field effect, has been done for diamond by Hanke and Sham[26] using the local-orbital representation. In diamond, the tetrahedral bonds are sufficiently localized that only the unscreened short-range attraction between the electron and hole ("within the same bond") was taken into account in addition to the long-range local-field (or exciton exchange) term V. Such an approximation for $-V^s$ is not used in silicon, the calculation for which is described in Sec. III.

Incidentally, we note the usage of two sets of terminology for the same physical effects. From the exciton specialists's viewpoint, $-V^s$ is the electron-hole attraction and V is the exciton-exchange term. From the response function specialists's viewpoint, V is the time-dependent Hartree term (RPA), and $-V^s$ is its screened exchange counterpart. Since our calculations for silicon indicate that $-V^s$ is relatively more important than V, the usage of the exciton physicists seems more reasonable in this case.

We describe in Sec. II also a different possibility for approximately including the many-particle effects in the optical excitations.[34] It makes use of the Kohn-Sham density-functional theory for the effective one-particle equation.[35] Here the electron-hole attraction and its exchange counterpart can be derived from the self-consistently determined density change in the exchange-correlation potential and in the Hartree potential, respectively. We have tested this possibility for Si and found that the commonly used local-density approximation[35] cannot properly account for the continuum-exciton effects.

Silicon furnishes a typical example of a semiconductor for which the noninteracting electron-hole pair does not yield an adequate description of the absorption peaks. The one-electron approximation gives the oscillator strength for the E_1 peak about a factor $\frac{1}{2}$ to $\frac{3}{4}$ of the observed value and for the E_2 peak somewhat too large.[28] The same is true for many semiconductors such as Ge, GaAs, InAs, InSb, ZnS, and ZnSe, just to mention a few.[36-38] We present in Sec. IV the results of our calculation of the optical absorption and modulation spectrum of silicon. In diamond the local-field effect (i.e., exciton exchange only) reduces the intensity of the main absorption peak calculated in the one-electron approximation more severely on the low-energy side.[26] In Si, our calculations give the same trend for the E_1 peak, but the changes are smaller in accordance with other works.[28,31] This general trend of the RPA to shift absorption strength to higher energies has been found in recent work on insulators and semiconductors independent of whether a pseudopotential representation[25,28] or the local-orbital scheme which we have developed[26] was used.[31,32] It furthers the discrepancy with experiment. On the other hand, the continuum-exciton effect increases the intensity of the main peak and shifts its position to a lower energy in diamond.[26] In Si, where we include a position-dependent screening in the electron-hole attraction, the E_1 peak is shifted by about 0.2 eV, and its intensity is almost doubled by the excitonic effect. Thus, the excitonic interaction modifies the commonly used identi-

fication of critical-point structure derived from the single-particle band structure with experimental absorption data. In Sec. V, a model analysis of the continuum-exciton and the local-field effect is given for the optical properties of semiconductors in general. The quantum mechanic interrelation with the classic Lorentz-Lorenz relation is worked out which, in contrast to previous work, relies not on assumptions about self-energy corrections. In particular we derive a simple analytic dependence of the electron-hole attraction on orbital localization and lattice constant. We give arguments that the enhancement of low-energy absorption structure due to the continuum-exciton effect quite generally can explain the significant deviation of the one-particle spectra of column IV, III-V, and II-VI semiconductors from experiments.[37,38] It is expected that observed deviations from the one-particle picture in the spin-orbit splitting of absorption data, in the temperature[39] and in stress dependence,[40] can be similarly accounted for.

II. MANY-BODY EFFECTS IN THE DIELECTRIC RESPONSE

A. The Bethe-Salpeter equation

Let us start by writing down the general formulation for the dielectric response by the field-theoretical method.[21] The inverse dielectric function ϵ^{-1} is related to the density response function χ by

$$\epsilon^{-1}(1,2)=\delta(1,2)-V(1,1')\chi(1',2)\,, \tag{2.1}$$

where each numeral stands for a set of position and spin coordinates and time, and the repeated numerals 1′ are understood to indicate appropriate integration and summation. $V(1,1')$ stands for the Coulomb interaction. The density response function χ is in turn given by

$$\chi(1,2)=S(1,1;2,2)\,, \tag{2.2}$$

where $S(1,1';2,2')$ is the part of the two-particle Green's function which excludes the disconnected term $-G(1,1')G(2,2')$, where G is the one-particle Green's function. S satisfies the Bethe-Salpeter

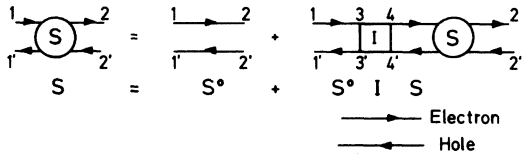

FIG. 1. The Bethe-Salpeter equation.

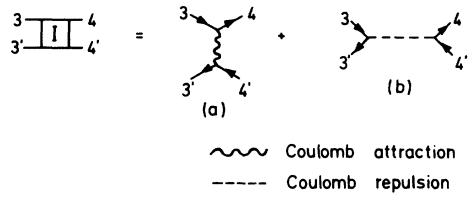

FIG. 2. The irreducible electron-hole interaction I: (a) electron-hole attraction; (b) unscreened exchange.

equation[21] (Fig. 1)

$$S(1,1';2,2')=S_0(1,1';2,2') +S_0(1,1';3,3')I(3,3';4,4') \times S(4,4';2,2')\,, \tag{2.3}$$

where

$$S_0(1,1';2,2')=G(1',2')G(2,1)\,. \tag{2.4}$$

The irreducible electron-hole interaction I will be taken here to be the sum of a screened electron-hole attraction plus an unscreened exchange (Fig. 2),

$$I(3,3';4,4')=-\delta(3,4)\delta(3',4')V^s(3,3') +\delta(3,3')\delta(4,4')V(3,4)\,. \tag{2.5}$$

The niceties of vertex corrections,[21] etc., are neglected here. It is obvious[21] that the exchange term must not be screened (e.g., by including the term in Fig. 3(a). Otherwise it will not be irreducible; i.e., we will be double counting in Eq. (2.3). More complicated exchange-type terms such as Fig. 3(b) are allowed but are not included here. If we further restrict the screening in V^s to be static, then I is time independent. The times of $1,1'$ in $S(1,1';2,2')$ are the same and so are those of 2 and 2′. The time difference between $1,1'$ and $2,2'$ is Fourier transformed to a frequency ω. From now on, the coordinates 1, 2, etc., will exclude the time dependence, and instead the frequency dependence ω of S and S^0 is understood.

B. The local-orbital representation

The formalism of Hanke and Sham[26] is used. We give here a slightly different derivation of the solution of the Bethe-Salpeter equation in the local-orbital representation. If we neglect the

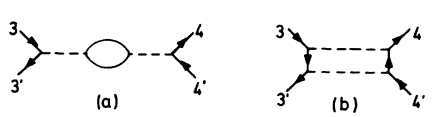

FIG. 3. Further interactions between an electron and hole.

lifetime effect of the one-particle propagation, the one-particle Green's function can be expressed in terms of Bloch waves $\psi_K(1)$, with energies E_K, where K includes both the wave vector $\vec{k}$ and band index n:

$$G(1,2;E)=\sum_K \frac{\psi_K(1)\psi_K^*(2)}{E-E_K}. \tag{2.6}$$

In the local representation, the Bloch wave is expressed in terms of the local orbitals $\phi_L(1)$:

$$\psi_K(1)=\sum_L T_{LK}\phi_L(1), \tag{2.7}$$

where

$$\phi_L(1)=\phi_\nu(\vec{r}_1-\vec{R}_l), \tag{2.8}$$

$\vec{R}_l$ being a lattice vector and ν an orbital index. Thus

$$T_{LK}=c_{n\nu}(\vec{k})\,e^{i\vec{k}\cdot\vec{R}_l} \tag{2.9}$$

in terms of the coefficients given in Ref. 26, Eq. (2.6). For simplicity, the number of unit cells in the crystal will be omitted from all expressions. Any set of localized orbitals ϕ_L can be chosen, provided they give a good description of the wave functions ψ_K. In our application to Si we take a linear combination of atomiclike orbits (LCAO), but we can equally well use a combination of muffin-tin orbitals, for example, the linearized muffin-tin orbitals (LMTO's) introduced by Andersen.[41] Substituting Eq. (2.7) in Eq. (2.6), we obtain the Green's function in local representation:

$$G(1,2;E)=\sum_{LL'}\phi_L(1)G_{LL'}(E)\phi_{L'}^*(2), \tag{2.10}$$

with

$$G_{LL'}(E)=\sum_K T_{LK}T_{L'K}^*/(E-E_K). \tag{2.11}$$

From Eq. (2.4), we have

$$S_0(1,1';2,2';\omega)=\sum_{L_1L_1'L_2L_2'}\phi_{L_1'}(1')\phi_{L_1}^*(1)S^0_{L_1L_1',L_2L_2'}(\omega)\times\phi_{L_2}(2)\phi_{L_2'}^*(2'), \tag{2.12}$$

where

$$S^0_{L_1L_1',L_2L_2'}(\omega)=\sum_{K,K'}T_{L_1K}^*T_{L_1'K'}\frac{f_{K'}-f_K}{E_{K'}-E_K-\omega}T_{L_2K}T_{L_2'K'}^*, \tag{2.13}$$

f_K being the occupation number of the K Bloch state. $S_{L_1L_1',L_2L_2'}(\omega)$ and $I_{L_3L_3',L_4L_4'}$, are defined by equations similar to Eq. (2.12). Now express $S^0_{L_1L_1',L_2L_2'}(\omega)$ in terms of new variables l_1, $l_1'-l_1$, l_2, $l_2'-l_2$ and Fourier transform with respect to l_1 and l_2 yielding wave-vector dependences $\vec{q}_1$, $\vec{q}_2$. By lattice translational symmetry, $\vec{q}_1=\vec{q}_2$. Therefore, we have

$$S^0_{L_1L_1',L_2L_2'}(\omega)=\sum_{\vec{q}}e^{i\vec{q}\cdot(\vec{R}_{l_1}-\vec{R}_{l_2})}N^0_{\lambda_1\lambda_2}(\vec{q};\omega), \tag{2.14}$$

where λ_1 is short for $l_1'-l_1$, ν_1, ν_1', with N^0 given by

$$N^0_{\lambda_1\lambda_2}(\vec{q};\omega)=\sum_{nn'\vec{k}}c^*_{n\nu_1}(\vec{k})c_{n'\nu_1'}(\vec{k}+\vec{q})\,e^{i(\vec{k}+\vec{q})\cdot\vec{R}_l}\frac{f_{n'\vec{k}+\vec{q}}-f_{n\vec{k}}}{E_{n'\vec{k}+\vec{q}}-E_{n\vec{k}}-\omega-i0}e^{-i(\vec{k}+\vec{q})\cdot\vec{R}_{l'}}c^*_{n'\nu_2'}(\vec{k}+\vec{q})c_{n\nu_2}(\vec{k}). \tag{2.15}$$

Similarly, the electron-hole interaction can be expressed as

$$I_{L_3L_3',L_4L_4'}=\sum_{\vec{q}}e^{i\vec{q}\cdot(\vec{R}_{l_3}-\vec{R}_{l_4'})}J_{\lambda_3\lambda_4}(\vec{q};\omega).$$

From Eq. (2.5),

$$J_{\lambda\lambda'}=-\tfrac{1}{2}V^s_{\lambda\lambda'}+V_{\lambda\lambda'}, \tag{2.16}$$

with V given by Eq. (2.14) of Ref. 26, or

$$V_{\lambda\lambda'}(\vec{q})=\int d^3r\int d^3r' A_\lambda^*(\vec{q};\vec{r})v(\vec{r}-\vec{r}')A_{\lambda'}(\vec{q};\vec{r}')$$
$$=\sum_m e^{-i\vec{q}\cdot\vec{R}_m}\int d^3r\,d^3r'\phi_\mu^*(\vec{r}-\vec{R}_l-\vec{R}_m)\phi_\nu(\vec{r}-\vec{R}_m)v(\vec{r}-\vec{r}')\phi_{\nu'}^*(\vec{r}')\phi_{\mu'}(\vec{r}'-\vec{R}_{l'}). \tag{2.17}$$

$A_\lambda(\vec{q},\vec{r})$ is a charge-density wave of wave vector $\vec{q}$ excited by the disturbance $\sim e^{i\vec{q}\cdot\vec{r}}$; i.e.,

$$A_\lambda(\vec{q},\vec{r})=\sum_m e^{i\vec{q}\cdot\vec{R}_m}\phi_\nu^*(\vec{r}-\vec{R}_m)\phi_\mu(\vec{r}-\vec{R}_l-\vec{R}_m), \tag{2.18}$$

where $\lambda=\nu,\mu,\vec{R}_l$. $\phi_\mu(\vec{r}-\vec{R}_l)\phi_\nu(\vec{r})$ is the charge density of an electron-hole pair with "dipole moment" $\vec{R}_l$.

The fact that in a periodic crystal one needs to consider the interaction of "dipoles" excited in a screening process having structure on an atomic scale reflects the physical intuition of the microscopic or local-field effects. V in Eq. (2.17) corresponds to a transition where an electron-hole pair at site

$\vec{R}=0$ induces a new electron-hole pair on a site $\vec{R}_m$ and destroys itself, thus the excitation can move around the crystal. The dipole-dipole interaction between Frenkel excitons,[42] the splitting of longitudinal and transverse excitons,[15] and the interference of continuous x-ray edges in simple metals have been recognized as an aspect of this exciton exchange.[14] However, in most investigations the exciton exchange has been studied by neglecting the $\vec{k}$ dependence of the Bloch functions which generate the exciton states in consideration.[14] Thus, V in Eq. (2.17) represents a general form of electron-hole exchange.

The electron-hole attraction is described by

$$-\tfrac{1}{2}V^s_{\lambda\lambda'}=-\tfrac{1}{2}\sum_m e^{-i\vec{q}\cdot\vec{R}_m}\int d^3r\int d^3r'\phi^*_\mu(\vec{r}-\vec{R}_l-\vec{R}_m)\phi_\nu(\vec{r}'-\vec{R}_m)v_s(\vec{r},\vec{r}')\phi^*_\nu(\vec{r}')\phi_{\mu'}(\vec{r}-\vec{R}_{l'})\,. \tag{2.19}$$

If all orbitals generating $-V^s$ are well localized, then the $\vec{R}_m$'s entering the summation in Eq. (2.19) as well as $\vec{R}_l$ and $\vec{R}_{l'}$ are limited to the same shells of first neighbors. In particular, if $\vec{R}_l=\vec{R}_{l'}$, we have the short-range attraction between an electron-hole pair of dipole moment $\vec{R}_l$, which determines the central-cell corrections in the excitonic picture. The factor $\frac{1}{2}$ in front of V^s is used to compensate for the double spin-state summation in N^0. With $S_{\lambda_1\lambda_2}(\vec{q};\omega)$ defined from S as in Eq. (2.14), Eq. (2.3) becomes a matrix equation and can be inverted to yield

$$S=N^0[1-(V-\tfrac{1}{2}V^s)N^0]^{-1}\,. \tag{2.20}$$

This procedure is practical only if the matrices are not too large. The specific size in silicon will be considered in Sec. III.

The matrix inversion of Eq. (2.20) becomes impractical when (i) the exciton radius (R_l) becomes large, as near the fundamental gap, in which case, one has to resort to EMA or (ii) the wave functions are free electron-like, in which case, the plane-wave representation should be used, as will be discussed in Sec. V.

The nonlocal electron-hole interaction $v_s(\vec{r},\vec{r}')$ in Eq. (2.19) has been discussed in the most general form by Sham and Rice.[21] In Ref. 26 we considered a short-range approximation to it, which is the electron-hole interaction in the "same bond."

From now on we can exactly proceed as in Ref. 26 and obtain in the long-wavelength limit the dielectric constant

$$\epsilon(\omega)=1-4\pi e^2\Omega_0^{-1}\sum_{\lambda\lambda'}f^\alpha_\lambda S_{\lambda\lambda'}(\omega)f^{\alpha,*}_{\lambda'}, \tag{2.21}$$

where

$$f^\alpha_\lambda=\int d^3r\,\phi^*_\nu(\vec{r})r_\alpha\phi_\mu(\vec{r}-\vec{R}_l)\,, \tag{2.22}$$

with $S(\omega)=S(q\to 0;\omega)$, α denoting a principal axis of the cubic crystal and Ω_0 the unit cell volume.

C. Local-density approximation for the electron-hole interaction

Before concluding this section we would like to discuss a different method of approximately including exchange and correlation effects in the dielectric function of a crystal.[34] The basic idea is to start from the Kohn-Sham equation for the Bloch state $\psi_{n\vec{k}}$ in the self-consistent field φ,[35]

$$[-(\hbar^2/2m)\nabla^2+\varphi(\boldsymbol{n}(\vec{r}))]\psi_{n\vec{k}}(\vec{r})=\boldsymbol{E}_{n\vec{k}}\psi_{n\vec{k}}(\vec{r})\,, \tag{2.23}$$

where $\varphi(\boldsymbol{n}(\vec{r}))=v_{\rm ext}(\vec{r})+v_{\rm H}(\vec{r})+v_{\rm xc}(\boldsymbol{n}(\vec{r}))$. Here $v_{\rm ext}$ is some external potential, for example the potential of the ions, $v_{\rm H}$ is the Hartree, and $v_{\rm xc}$ the exchange-correlation potential being a functional of the density $\boldsymbol{n}(\vec{r})$. If the electron system is subjected to an external perturbation $\delta v=v_{\rm ext}$, which we take proportional to $e^{i\vec{q}\cdot\vec{r}}$, then the Schrödinger-type equation (2.23), due to density change $\delta\boldsymbol{n}=\boldsymbol{n}_{\rm ind}$, contains the potential change

$$\delta\varphi=\delta v+\delta v_{\rm H}+\delta v_{\rm xc}\,, \tag{2.24}$$

where

$$\delta v_{\rm xc}=[\delta v_{\rm xc}(\boldsymbol{n}(\vec{r}))/\delta\boldsymbol{n}]\delta\boldsymbol{n}(\vec{r})\,. \tag{2.25}$$

This potential change in turn gives rise to a change $\delta\psi_{n\vec{k}}$ of the wave function, and thus generates the usual self-consistent circle of density response, described by Eq. (2.3), with the effective interaction $I(3,3';4,4')$ $=\delta(3,3')\delta(4,4')J(3,4)$ and

$$J(\vec{r}_1,\vec{r}_2)=v(\vec{r}_1-\vec{r}_2)+\frac{\delta v_{\rm xc}(\boldsymbol{n}(\vec{r}_1))}{\delta\boldsymbol{n}}\delta(\vec{r}_1-\vec{r}_2)\,. \tag{2.26}$$

This simply means that in Eq. (2.16) we use for V^s

$$-\tfrac{1}{2}V^s_{\lambda\lambda'}(\vec{q})=\Omega_0\int d^3r\,A^*_\lambda(\vec{q},\vec{r})\times\frac{\delta v_{\rm xc}(\boldsymbol{n}(\vec{r}))}{\delta\boldsymbol{n}}A_{\lambda'}(\vec{q},\vec{r})\,, \tag{2.27}$$

with A_λ defined in Eq. (2.18).

In the spirit of the "local-density" approximation[35] $v_{\rm xc}$ and its density derivative can be ex-

tracted from homogeneous electron gas results for the exchange-correlation energy of an electron. In contrast to the commonly used Hubbard-type approximation, where the Coulomb interaction v is simply modified by a factor which is taken as approximately accounting for the exchange and correlation effects, the result (2.27) takes into account the crystal-structure effects in the induced density. However, the crucial approximation in the above scheme is the "local-density" approximation for the electron-hole attraction V^s. We have tested the local-density scheme for the optical spectrum of Si. Since the Kohn-Sham theory is ω independent, an appropriate test is provided by the static $\epsilon(\omega=0)$. We have found that the local-density approximation for V^s cannot properly account for the continuum-exciton effect and results in too low a value of $\epsilon(\omega=0)$ (see discussion in Sec. IV).

III. LOCAL REPRESENTATION FOR A COVALENT SEMICONDUCTOR: Si

A. Energy bands and wave functions

In constructing energies and wave functions for a covalent crystal and in calculating dielectric response in the local representation, we closely follow the method discussed in our previous paper.[26] The wave functions and energy bands are considered in a silicon structure of cube edge a, arising from s and p atomic orbitals. Corresponding to the eight s and p spatial orbitals per unit cell, there are eight bands which split into two disconnected band complexes, the valence- and the conduction-band complex. Wannier or localized functions corresponding to this structure have been discussed and used by Slater and Koster as a basis for their interpolation scheme.[43] A method of directly constructing the valence-band Wannier functions of Si by minimizing the bond energy in the presence of pseudopotential-like crystal potentials has been reported by the Kanes.[44] They expanded bonding and antibonding orbitals not just in sp^3 hybrids but also in a linear combination of d-type orbitals. However, a large part of the d character of the wave functions can be attributed to the overlap of p states on nearest-neighbor atoms,[45] and thus is effectively contained in our Slater-Koster fit of energies. Thus, we stay with the sp^3-hybrid basis localized in the four tetrahedral directions

$$h_{\vec{\nu}}(\vec{r}) = (4\pi^{1/2})^{-1}[R_s(r) + 3^{1/2}(\vec{\nu}\cdot\vec{r}/r)R_r(r)], \quad (3.1)$$

where $\vec{\nu}$ represents the tetrahedral vector [111], [1$\bar{1}\bar{1}$], [$\bar{1}$1$\bar{1}$], and [$\bar{1}\bar{1}$1]. Two hybridized orbitals of nearest-neighbor atoms pointing in the same direction are added and subtracted to form bonding and antibonding orbitals:

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

where $N_{\pm}$ is the normalization and $b = a/4$, with a the lattice parameter. The $\phi_{\vec{\nu}\pm}$ then play the role of the Wannier functions of the band complex employed in Eq. (2.7).

The valence bands are expanded in bonding orbitals and the conduction bands in antibonding orbitals, including overlap integrals of the one-electron Hamiltonian H up to the third-nearest neighbors. The overlap integrals were fitted to the orthogonalized-plane-wave (OPW) energy bands calculated by Ortenburger and Rudge.[47] The results of this third-nearest-neighbor LCAO model are shown in Fig. 4 with the symmetry-point values given in Table I. The adjustment of energies to optical experiment is justified, since the many-particle corrections to the one-electron critical points are small ($\lesssim$0.2 eV, see below) and within the error range of our approximate band-structure determination.

Our LCAO band structure is also in good overall agreement for the lowest conduction and highest valence bands with the recent pseudopotential results of Chelikowsky and Cohen.[48] The energy differences $E(L_1)$-$E(L_3')$ and $E(X_1)$-$E(X_4)$ are 3.49 and 4.57 eV in our calculation compared to 3.37 and 4.21 eV in the pseudopotential work.[48]

We then have for every $\vec{k}$ point in the Brillouin zone two (4 × 4) secular equations, one for the valence and one for the conduction bands. The

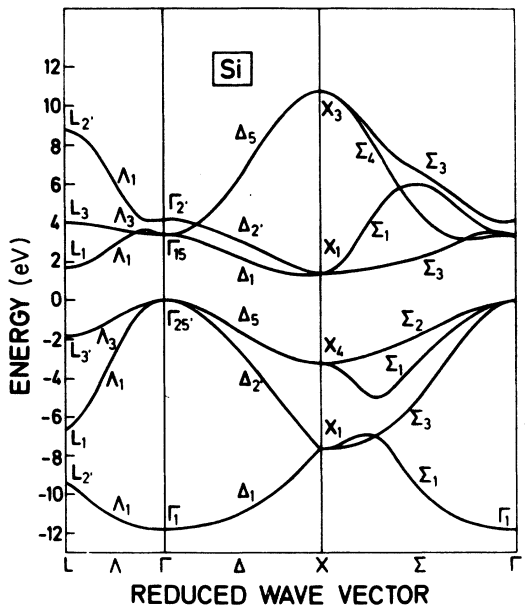

FIG. 4. Third-neighbor LCAO interpolation of the band structure of Si.

TABLE I. Symmetry-point values of third-nearest-neighbor Slater-Koster model for Si.

Point	Energy (eV)
(A) Conduction bands	
E (Γ_{15})	3.37
E (L_3)	4.01
E (X_3)	10.74
E ($L_{2'}$)	4.19
E ($L_{2'}$)	8.72
E (L_1)	1.65
E (X_1)	1.37
(B) Valence bands	
E ($\Gamma_{25'}$)	0.0
E (X_4)	−3.2
E ($L_{2'}$)	−9.48
E (Γ_1)	−11.78
E ($L_{3'}$)	−1.84
E (X_1)	−7.65
E (L_1)	−6.68

corresponding eigenvalues give the energies $E_{n\vec{k}}$ and the eigenvectors the coefficients $c_\nu(n\vec{k})$ of Eq. (2.9). These are the input quantities required for calculating the $N^0(q \to 0; \omega)$ matrix of Eq. (2.15). The $\vec{k}$ sums in this matrix are computed using Gilat-Raubenheimer routines, separately for the real and imaginary parts of N. The sums are performed only over the irreducible Brillouin zone. The contributions from the eigenvectors c and the exponentials in Eq. (2.15), which come from other parts of the zone, are transformed into the irreducible zone using the symmetry of the lattice. The irreducible part is then divided into small cubes with a maximum box width of 0.153 a.u. Reducing the box width by a factor of 20 introduces an average difference of ~10% in $\bar{\epsilon}(\omega)$ of Si. We do not consider this difference worth the heavy computational effort involved with the finer mesh and use the coarse grid in our actual calculation, presented in the next section.

We have used the local orbitals in the Gaussian representation.[26] It allows for an analytic integration of all the multicenter integrals involved in the electron-hole interaction (in particular the screened attraction $-V^s$, see Appendix) and makes possible a transparent interpretation of our results. To this the atomic s and p orbitals are expanded as

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

$$xR_r(r) = \sum_j x b_j e^{-\beta_j r^2}. \tag{3.4}$$

One way of determining the Gaussian parameters is to fit the density $n(\vec{r})$, constructed from the bonding orbitals of Eq. (3.2), to the density of the band-structure calculation underlying our energy interpolation. Since Ref. 47 gives no density we use the "experimental" density profile[49] reported for Si in the pseudopotential work of Chelikowski and Cohen.[48]

We have devised another way of determining the Gaussian parameters which, in contrast to the above density adjustment and to the bond energy minimization procedure,[44] does include also excited state properties.[26] It is based on the equation of continuity or on the resulting equality between density and current response. This condition is expressed in the equation

$$\langle n\vec{k}|\, e^{i\vec{q}\cdot\vec{r}}\,|\, n'\vec{k}+\vec{q}\rangle (E_{n'\vec{k}+\vec{q}} - E_{n\vec{k}}) = \langle n\vec{k}|\, \vec{q}\cdot\vec{j}(\vec{q})\,|\, n'\vec{k}+\vec{q}\rangle, \tag{3.5}$$

where $\vec{j}(\vec{q})$ denotes the current operator. Armed with the knowledge of eigenvalues $E_{n\vec{k}}$ and eigenvectors $c_\nu(n\vec{k})$, we determine the Gaussian coefficients in Eqs. (3.3) and (3.4) such that, indeed, both current and density representation give the same answer for the dielectric constant without local-field corrections [$\bar{\epsilon}(\omega)$ in Sec. IV]. Utilizing some arbitrariness in the local orbitals[50] this method effectively constructs "ultralocalized" orbitals which are consistent with our third-nearest neighbor model for the interpolated band structure.[26] Being interested not only in ground-state properties but in transitions to excited states, we consider the current conservation as an important criterion for internal self-consistency in any dielectric-response calculation. In our actual computation of the optical properties of Si we use a set of Gaussian parameters which is optimized with respect to the current-conservation criterion but at the same time gives also a reasonable fit to the charge density.[49] These parameters are listed in Table II.

The overlap between sp^3 hybrids located on neighboring Si sites which is derived from this parameter set is 0.7 in close agreement with the atomic value of 0.69 calculated with Hartree-Fock orbitals. The atomiclike decay of orbitals

TABLE II. Coefficients of $3s$ and $3p$ Gaussians.

Parameter	Combined current-conservation and density fit (a.u.)
a_1	3.2252
a_2	−2.8788
α_1	0.3
α_2	0.4
b_1	0.1344
b_2	0.0287
β_1	0.09
β_2	0.1915

within nearest-neighbor distance is also consistent with the actual Wannier-function results.[44]

Using our Gaussian wave functions we found that the contributions of second-nearest neighbor interactions between bonding and antibonding orbitals to the dipole elements f_λ and to the densities entering Coulomb and exchange matrices are of the order of 5–10%. Thus, in analogy with our previous diamond calculation,[26] only nearest-neighbor overlaps are retained in computing the dielectric function $\epsilon(\omega)$ of Eq. (2.21) for Si. The matrices N, V, V^s, and S are then of dimension 28. It should be noted that the electron-hole attraction $-V^s$ in this approximation includes interactions between electron and hole up to the second-nearest neighbors in the Si lattice (nearest neighbors in the bonding, antibonding lattice). The screening in $-V^s$ is approximated by the local screening function $\epsilon^{-1}(\vec{q})$ of Ref. 51.

B. The interaction matrices V and V^s

As discussed in Sec. II the exchange interaction matrix V is a direct measure for the importance of the local fields, whereas the matrix $-V^s$ determines the influence of the electron-hole attraction or excitonic effect on the dielectric function. The combined matrices $(V - V^s)$ give the interaction energies of the charge distributions $\phi_\nu^* \phi_\mu$ and $\phi_{\nu'}^* \phi_{\mu'}$ which have been excited by the external perturbation.

In the Gaussian representation analytic results can be derived for both interaction matrices. This is demonstrated in particular for the screened electron-hole attraction in the Appendix. Now a great deal of physical insight and also practical ways of calculating the dielectric function including these many-body effects can be gained by introducing a cluster expansion of the Coulomb interactions in $(V - \frac{1}{2}V^s)$.

Consider the bonding and antibonding states ν and μ which are expressed as linear combinations of hybridized orbitals h_ν and h_μ, as before. Again, we assume that the charge density $\phi_\nu^*(\vec{r})\phi_\mu(\vec{r} - \vec{R}_l)$ can adequately be described in a nearest-neighbor model of the overlap between bonding and antibonding states. Then the density $\phi_\nu(\vec{r})\phi_\mu(\vec{r} - \vec{R}_l)$ is entirely determined by the basic interactions which are shown in Fig. 5. These fundamental charge distributions also constitute the density waves $A_\lambda(\vec{q}, \vec{r})$ of Eq. (2.18) and the dipole matrix elements f_λ of Eq. (2.22). The distribution (e) may be considered as some kind of a microscopic measure of the bond charge of the crystal, while the overlap $h_\nu^*(\vec{r})h_\nu(\vec{r})$ depicted in Fig. 5(d) corresponds to the atomic charge of the constituent Si atoms. Each of these charge distributions is composed out of monopole (s-s), dipole (s-p), and quadrupole (p-p) contributions which are in the Gaussian representation of Eqs. (3.3) and (3.4) of the form, respectively:

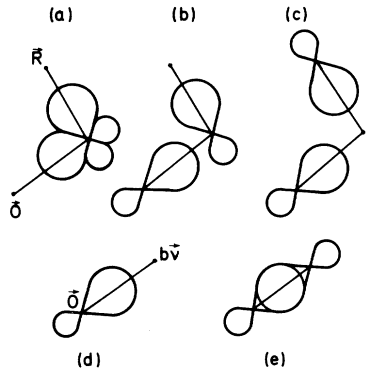

FIG. 5. Overlap terms of the hybridized orbitals.

$$e^{-\alpha_i(\vec{r}-\vec{A})^2} e^{-\alpha_j(\vec{r}-\vec{B})^2}, \tag{3.6}$$

$$\vec{\nu}(\vec{r} - \vec{A}) e^{-\beta_i(\vec{r}-\vec{A})^2} e^{-\alpha_j(\vec{r}-\vec{B})^2}, \tag{3.7}$$

$$[\vec{\nu} \cdot (\vec{r} - \vec{A})][\vec{\mu} \cdot (\vec{r} - \vec{B})] e^{-\beta_i(\vec{r}-\vec{A})^2} e^{-\beta_j(\vec{r}-\vec{B})^2}. \tag{3.8}$$

These charge distributions are located within a unit cell Ω_0 and interact in the matrix V of Eq. (2.17) via the Coulomb interaction $v(\vec{r} - \vec{r}')$ with corresponding charge distributions in all the unit cells with centers at sites $\vec{R}^m$ of the fcc sublattice. Thus, they give rise to a multipole expansion of the density interaction energy V.

The monopole-monopole interaction is given by

$$V_{\lambda\lambda'}^{mm} = \sum_m \sum_{ijkl,\vec{A}\vec{B}\vec{C}\vec{D}} M(ijkl) \iint d^3r\, d^3r'\, e^{-\alpha_i(\vec{r}-\vec{A}-\vec{R}_m)} e^{-\beta_j(\vec{r}-\vec{B}-\vec{R}_m)} v(\vec{r} - \vec{r}') e^{-\alpha_k(\vec{r}'-\vec{C})^2} e^{-\beta_l(\vec{r}'-\vec{D})^2}, \tag{3.9a}$$

where $\vec{A}$ and $\vec{B}$ stand for all possible s-type Gaussian centers occurring in $\phi_\nu^*(\vec{r})$ and $\phi_\mu(\vec{r} - \vec{R}_l)$. The analytic expression for the four-center integral in Eq. (3.9a) is listed in the Appendix. There it is also shown how the higher-order multipole interactions in V can be derived from the monopole interaction by repeated differentiations with respect to the Gaussian centers. Using (A2) we have, for example, for the dipole-dipole contribution

$$V_{\lambda\lambda'}^{dd} = \sum_m \sum_{ijkl,\vec{A}\vec{B}\vec{C}\vec{D}} D(ijkl) \sum_{xx'} \vec{\nu}_x \vec{\nu}'_{x'} \frac{\partial^2}{\partial A_x \partial A_{x'}} \int d^3r \int d^3r'\, e^{-\alpha_i(\vec{r}-\vec{A}-\vec{R}_m)^2} e^{-\beta_j(\vec{r}-\vec{B}-\vec{R}_m)^2} v(\vec{r} - \vec{r}') e^{-\alpha_k(\vec{r}'-\vec{C})^2} \times e^{-\beta_l(\vec{r}'-\vec{D})^2}. \tag{3.9b}$$

Alternatively, the full interaction matrix can be expressed in the monopole form of Eq. (3.9a) by expanding, for example, a p_x-type Gaussian in two Gaussian lobe orbitals of s-type symmetry

$$xR_p(r) = \sum_j c_j (e^{-\beta_j(\vec{r}-\vec{d}_{jx})^2} - e^{-\beta_j(\vec{r}+\vec{d}_{jx})^2}), \tag{3.10}$$

where $\vec{d}_{jx}$ is a small vector simulating the angular dependence of the wave function.[26]

A further simplification of the interaction in Eqs. (3.9) can be introduced by making use of the well-known property of Gaussians that their product is again a Gaussian,

$$e^{-\alpha(\vec{r}-\vec{A})^2} e^{-\beta(\vec{r}-\vec{B})^2} = \kappa e^{-(\alpha+\beta)(\vec{r}-\vec{C})^2}, \tag{3.11}$$

where

$$\kappa = \exp\left(-\frac{\alpha\beta}{\alpha+\beta}(\vec{A}-\vec{B})^2\right), \tag{3.12a}$$

$$\vec{C} = (\alpha\vec{A}+\beta\vec{B})/(\alpha+\beta). \tag{3.12b}$$

As a consequence, the charge densities $\phi_\mu^*(\vec{r})\phi_\nu(\vec{r}-\vec{R}_l)$ can themselves be described by a sum over single s-type Gaussians distributed over a unit cell and centered in the plane spanned by the vectors $\vec{\nu}$ and $\vec{\mu}$.

It is easy to verify from the basic interactions shown in Fig. 5 that the charge densities are given by cluster of Gaussians with centers 1 to 6 as illustrated in Fig. 6.[26] Each cluster may then be approximated by only one Gaussian centered at atomic sites or bond centers respectively, such that the total charge in each cluster is conserved.

We have checked the accuracy of such an approximation in our explicit calculation of the RPA dielectric function $\epsilon_{\mathrm{RPA}}(\omega)$ of Si. The average error introduced in this one-Gaussian site RPA representation for the optical absorption is of the order of 15% for all ω's.

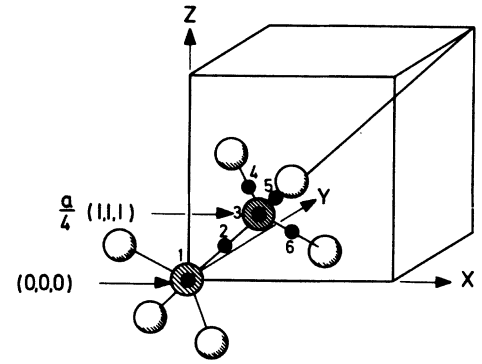

FIG. 6. Distribution of equivalent charge densities in the Gaussian representation.

Using Eqs. (3.9) and (3.11) it is then straightforward to show that the interaction matrix V can be written in the form

$$V_{\lambda\lambda'} = \sum_{\kappa\kappa'} W_\lambda^\kappa U_{\kappa\kappa'} W_{\lambda'}^{\kappa'*}, \tag{3.13}$$

with

$$U_{\kappa\kappa'} = \sum_m \int d^3r \int d^3r'\, e^{-\kappa(\vec{r}-\vec{C}_\kappa-\vec{R}_m)^2} \times v(\vec{r}-\vec{r}')\, e^{-\kappa'(\vec{r}'-\vec{C}_{\kappa'})^2}. \tag{3.14}$$

$U_{\kappa\kappa'}$ is the density interaction matrix in the site representation and the $\vec{C}_\kappa$ ($\kappa = 1, 6$) denote the centers of Gaussian distributions in Fig. 6. The matrix W_λ^κ contains the coefficients $M(ijkl)$ and κ of Eqs. (3.9) and (3.12).

Let us take for the moment the electron-hole attraction as contained in the effective polarizability matrix N, with $N = N^0(1 + \frac{1}{2}V^s N^0)^{-1}$. Owing to the separable form of the density-interaction matrix V in Eq. (3.13), we can then write for the screening matrix S in a short-hand matrix notation

$$S = N(1-VN)^{-1} = N + NWU(1-\hat{N}U)^{-1}W^\dagger N, \tag{3.15}$$

with

$$\hat{N} = W^\dagger N W. \tag{3.16}$$

We see that the inversion problem of the dielectric matrix has been reduced from that of a 28-dimensional one in the Wannier representation to that of a 6-dimensional matrix in the site representation. Introducing the result (3.15) into Eq. (2.21) for the dielectric constant gives

$$\epsilon(\omega) = 1 - \frac{4\pi e^2}{\Omega_0}\sum_{\lambda\lambda'} f_\lambda^\alpha N_{\lambda\lambda'} f_{\lambda'}^{\alpha*} - \frac{4\pi e^2}{\Omega_0}\sum_{\lambda\lambda'} f_\lambda^\alpha [NWU(1-\hat{N}U)^{-1}W^\dagger N]_{\lambda\lambda'} f_{\lambda'}^{\alpha*} = \bar{\epsilon}(\omega) + \Delta\epsilon(\omega), \tag{3.17}$$

where $\bar{\epsilon}(\omega)$ is the dielectric constant without local-field corrections and $\Delta\epsilon(\omega)$ gives explicitly the local-field contribution to the dielectric constant. Thus, the site representation allows in a natural way to separate the dielectric response in a part which fluctuates on the microscopic scale of the charges $\vec{C}_\kappa$ involved (and in fact is *caused* by these localization centers), and in a part $[\bar{\epsilon}(\omega)]$ which fluctuates with the wavelength of the applied external perturbation.

Following the same arguments as given above for the interaction matrix V, the site represen-

tation can be utilized to write the matrix elements f_λ^α in Eq. (2.21) as

$$f_\lambda^\alpha \cong \sum_\kappa F_\kappa^\alpha W_\lambda^{\kappa *} . \tag{3.18}$$

Again W_λ^κ contains the space-group operations necessary to transform the Gaussian charge distributions in $\phi_\nu^*(\vec{r})\phi_\mu(\vec{r}-\vec{R}_l)$ for general $\vec{\nu}$, $\vec{\mu}$, and $\vec{R}_l$ back to the distributions situated at positions $\vec{C}_\kappa$ in Fig. 6. Then the dielectric constant is given by

$$\epsilon(\omega) = 1 - \frac{4\pi e^2}{\Omega_0} \sum_{\kappa\kappa'} F_\kappa^\alpha [\hat{N}(1-U\hat{N})^{-1}]_{\kappa\kappa'} F_{\kappa'}^{\alpha *} . \tag{3.19}$$

The practical value of this result lies in the fact that it allows for a simplified inclusion of crystal-structure or local-field effects in the dielectric response and makes model predictions possible (see Sec. V A).

In Eqs. (3.17) and (3.19) electron-hole attraction is incorporated in the effective polarization $N = N^0(1 + \frac{1}{2}V^s N^0)^{-1}$. However, the electron hole attraction can approximately also be directly included in the above interaction scheme by making use of the short-range properties of the interaction: We have found in our calculation of Si (see Sec. IV) that the dominant effect of the screened electron-hole interaction is given by (a) the almost unscreened term in Eq. (2.19) with $\mu = \nu = \mu' = \nu'$, (b) the short-range term $\mu = \mu'$, $\nu = \nu'$ but $\nu \neq \mu$. These Frenkel-type short-range contributions, which are about a factor 5 larger than the next larger terms, mean the dominance of terms which occur also in the interaction matrix V. Thus, by reducing the terms of these types in the density interaction V of Eq. (3.13) by terms (a) to (b) the effect of electron-hole attraction can approximately be incorporated into the above scheme.

IV. OPTICAL SPECTRUM OF Si

We present here the calculated optical absorption spectrum, the imaginary part $\mathrm{Im}\epsilon(\omega)$ of $\epsilon(\omega)$, for three different approximations: In $\bar{\epsilon}(\omega)$ both local-field and excitonic effects are neglected. The calculation denoted by $\epsilon_{\mathrm{RPA}}(\omega)$ takes the local-field effects into account but leaves electron-hole attraction out. Finally, $\epsilon_{xc}(\omega)$ denotes the result for the optical spectrum where both local-field and excitonic effects are taken into account. The three different approximations for the dielectric constant are then used to construct the corresponding optical modulation spectrum $R^{-1}(\omega)dR(\omega)/d(\omega)$, where $R(\omega)$ is the reflectivity.

The results for the optical absorption are plotted in Fig. 7 and compared with experiment.[52] The optical absorption spectrum is quite sensitive to many-particle effects. Consider the E_1 and E_2 main absorption peaks. The one-electron calculation [$S = N^0$ in Eq. (2.20)], including neither excitonic nor local-field effects, just indicates structure around the E_1 position, and gives the oscillator strength for the E_1 peak only about $\frac{1}{2}$ of the observed value. For energies above the E_2 peak the oscillator strength is significantly above experiment. The same general discrepancy of optical absorption calculated in the one-electron approximation was found for Si using quite different band-structure and wave-function descriptions, for example in the empirical pseudopotential calculation,[28] where the principal energy gaps are fitted to experiment. Quite generally, in

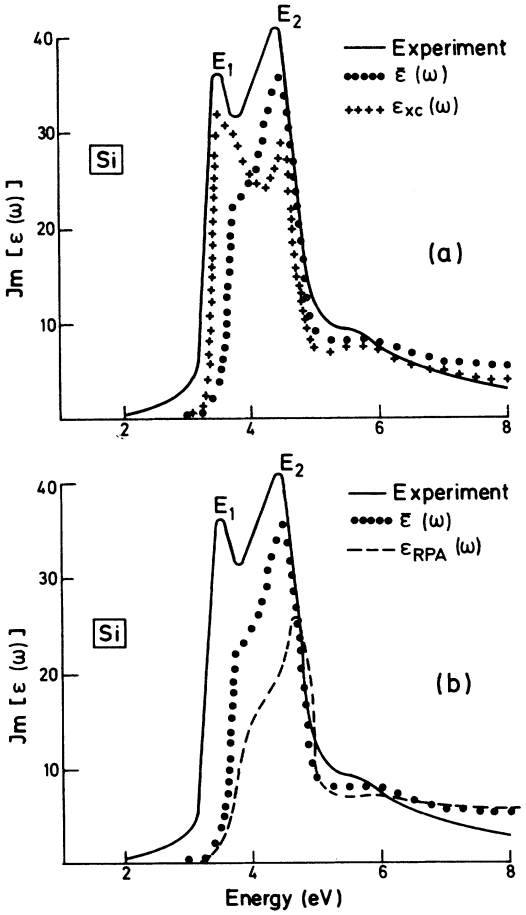

FIG. 7. Imaginary part of the dielectric constant versus energy; experimental data from Ref. 52; $\bar{\epsilon}$ gives the single-particle calculation, ϵ_{RPA} the calculation within RPA with local-field correction, and ϵ_{xc} the calculation including screened electron-hole attraction and local-field correction.

many column IV, II-V, and II-VI semiconductors,[36,37] the one-electron dielectric function gives the oscillator strength for the E_1 peak about $\frac{1}{2}$ or $\frac{3}{4}$ of the observed value and that for the E_2 peak somewhat too large.

We have calculated orthogonality corrections on $\epsilon(\omega)$. The bonding and antibonding orbitals in Eq. (3.2) contain already the main atomic overlap between the sp^3 hybrids of nearest-neighbor atoms (≈ 0.7 in Si). Thus, constructing orthogonalized orbitals, one can make use of the well-known overlap expansion[46] and truncate the series after quadratic terms in the overlap between orbitals, which are typically of the order of magnitude 10^{-2} in Si. In the main absorption region (3-10 eV) there is a small enhancement of oscillator strength with maximum rise ($\approx 10\%$) around the E_2 compared to the "nonorthogonal" E_2 region. On the average the changes introduced by these corrections are of the order of 5% and thus the further neglect of orthogonality corrections is justified.

Local-field effects within the RPA [$S = N^0(1 - VN^0)^{-1}$, ϵ_{RPA} in Fig. 7(b)], produce an additional shift of about 0.1 eV for the E_1 structure and of about 0.15 eV for the E_2 peak to higher energies. These Coulomb-repulsion effects shift oscillator strength toward higher energies, contrary to what is required to reconcile theory with experiment. Again these findings on the local-field effect are of general nature: Using our local-orbital scheme we have shown for diamond[26] that the repulsive Coulomb interaction V has the effect of widening the interband transitions, thus shifting oscillator strength to higher energies. This RPA result and trend is also found in the pseudopotential calculation for diamond by Van Vechten and Martin.[25] In Si the changes introduced in the RPA are smaller and therefore the local-field effect is smaller, as is to be expected from the more delocalized character of wave functions when compared to the insulator diamond. This is in qualitative agreement with other work.[28,31] In recent calculations of the optical spectra of thallous halides Schäfer, Schreiber, and Treusch[32] using our local-orbital method have established similar findings for the local-field effect in TlCl. The general trend of the RPA local-field effect to shift absorption strength to higher energies is reflected in a decrease of the static dielectric function. The value for our calculated static dielectric constant $\bar{\epsilon}(0)$ without local-field effects is 9.85, compared to 10.1 of the pseudopotential calculation of Louie *et al.*[28]; with local-field effects it decreases to 8.0 in our work and to 9.0 in the pseudopotential work. The experimental value is 11.7. Again the general trend of the RPA furthers the discrepancy with experiment.

On the other hand, the electron-hole attraction screened by a momentum-dependent dielectric function, effectively tends to lower transition energy thereby shifting absorption strength to lower energies. In Si, as is shown by the curve $\epsilon_{xc}(\omega)$ in Fig. 7(a) which contains both local-field and excitonic effects, the E_1 peak is shifted by about 0.2 eV compared to the peak position determined in the one-electron calculation $\bar{\epsilon}(\omega)$, where the peak structure coincides with the interband transitions of the underlying band structure (Fig. 4). The intensity of the E_1 peak is almost doubled by the excitonic effect. Concomitantly, the intensity of the E_2 peak and also of absorption at higher energies is reduced. Thus, the excitonic interaction modifies the commonly used identification of critical-point structure derived from the single-particle band structure with experiment. The macroscopic dielectric constant is raised to 10.4 by including the continuum-exciton effect. A similar rise in $\epsilon(\omega = 0)$ due to excitonic effects is reported in the work of Schäfer *et al.* for TlCl.[32]

In real space, the long-range electron-hole attraction in Eq. (2.19) is effectively screened by the macroscopic dielectric constant $\epsilon_0 = \epsilon(\omega = 0)$, whereas the short-range terms are essentially unaffected by screening. As already mentioned, the dominant contribution to the short-range terms in V^s comes from processes with either (a) $\mu = \nu = \mu' = \nu'$ or (b) $\mu = \mu'$, $\nu = \nu'$. There are also terms of these types in the direct Coulomb energy $V_{ss'}$. The short-range approximation, without screening which we used for diamond,[26] gives also the overall exciton effect in Si and is thus justified. The argument[31] that the exchange terms should be uniformly reduced by the macroscopic dielectric constant is incorrect. The two dominant matrix elements (a) and (b) of V^s, where (a) takes the value -0.0304 a.u. and (b) the value -0.0266 a.u. are a factor of about 5 larger than all remaining matrix elements. They are also a factor of about 2 larger than the corresponding two Coulomb matrix element (V_{ss} with $\nu = \mu = \nu' = \mu'$ and $\vec{R}_l = \vec{R}_{l'} = 0$ in atomic units has the value 0.014). This indicates that in fact the electron-hole attraction or the continuum-exciton effect is more important in Si than the local-field effect. Furthermore, the fact that at least two interaction matrix elements are needed in V^s and V has as a consequence that neither the electron-hole interaction in a covalent semiconductor[8,13] nor the local-field effect[13] can be treated within the contact approximation.

When the electron-hole attraction is calculated utilizing the density-functional scheme in the local-density approximation (Sec. II C), the E_1

structure of the RPA calculation ϵ_{RPA} is moved back only a little, not even back to the position of $\bar{\epsilon}$. The extracted value of $\epsilon(\omega=0)$ is 8.9. Thus, this local-density scheme cannot properly account for the continuum-exciton effect. The use of the local-density approximation in the response function is exact in the limit of a long-wavelength perturbation of a homogeneous gas. Its use in a covalent crystal apparently does not reproduce well the short-wavelength Bragg terms in the electron-hole attraction.

The continuum-exciton effect dominant around E_1 position is caused not only by transitions in $\langle 111 \rangle$ directions near the L point but also significantly by transitions near the Γ point. Thus, quite extended regions in the Brillouin zone and several critical points contribute. Therefore it is somewhat unsatisfactory to treat this saddle-point-type excitonic effect in the effective-mass approximation as was successfully done for insulators.[9]

We have used the above results for the real and imaginary parts of $\bar{\epsilon}(\omega)$, $\epsilon_{RPA}(\omega)$, and $\epsilon_{xc}(\omega)$ to calculate the reflectivity $R(\omega)$ and from that, by numerical differentiation the wavelength modulation spectrum $R^{-1}(\omega)dR(\omega)/d\omega$ in Si. With our coarse grid evaluation of N^0, we cannot hope to reproduce all the details of the Van Hove singularities in our calculated modulation spectrum. We simply wish to examine the overall line shape connected with the E_1 and E_2 peaks and the interaction effects thereupon.

The results of our calculation of the modulation spectrum are shown in Fig. 8(a) and compared with experiment and an empirical pseudopotential one-electron calculation in Fig. 8(b). The empirical pseudopotential has its form factors adjusted to achieve agreement with principal optical transitions of the experiment. More recent pseudopotential calculations of the modulation spectrum[48] are in general agreement with the pseudopotential results in Fig. 8(b), particularly around the 3.5-eV regime which is of most interest to us here. There exist also very accurate recent electroreflectance measurements of $R^{-1}dR/d\omega$ in Si,[54] which, for the accuracy we have in mind in our comparison with experiment, are in accordance with the experimental data in Fig. 8(b).

Consider the one-electron calculation in Figs. 8(a) and 8(b). Both the local-orbital calculation based on $\bar{\epsilon}$ and the pseudopotential model give a typical, by about a factor of 2, too small negative modulation strength around 3.5 eV. This strength is connected with the slope of $\mathrm{Im}\epsilon(\omega)$ between E_1 peak and the minimum between E_1 and E_2 peaks. Furthermore, our one-electron modulation spectrum is too low below 3.5 eV and significantly

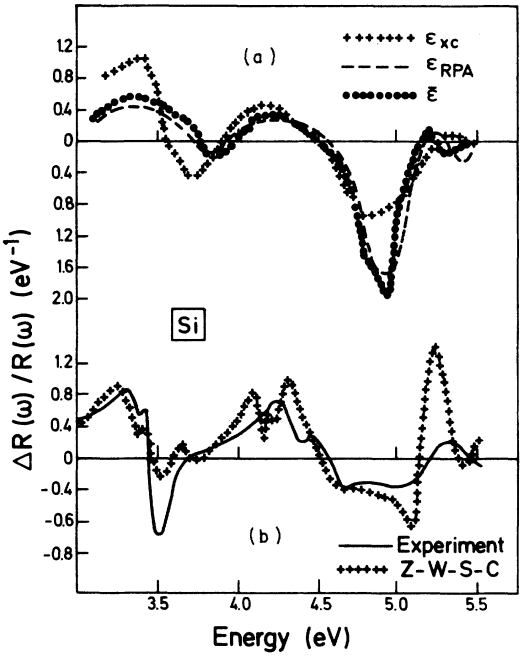

FIG. 8. A comparison of theoretical and experimental modulated reflectivity for Si. For legend, see also Fig. 1. Experimental and single-particle pseudopotential data ($Z-W-S-C$) are taken from Ref. 53.

overshoots the experimental negative result around 5 eV. The pseudopotential calculation on the other hand overestimates drastically the positive value around 5.5 eV. The critical-point structure of the pseudopotential result is in much better agreement with experiment, not surprisingly, since the principal gaps are fitted to experiment. Quite generally, *whatever* energy and wave-function description is used in $\bar{\epsilon}$, it does not reproduce the experimental low-energy absorption line shape (see also Refs. 28 and 31). Inclusion of RPA local-field effects in Fig. 8(a) makes both positive and negative values of $\Delta R(\omega)/R(\omega)$ around the E_1 critical point smaller, thus diminishing the structure and enhancing discrepancy with experiment as in the optical absorption. The electron-hole attraction significantly improves agreement with experiment: It enhances both positive and negative slopes in $R(\omega)$ around the E_1 peak; it also shifts the oscillator strength and critical peak structure, again in particular the E_1 peak, which is shifted by about 0.2 eV to lower energies as discussed for the imaginary part of $\epsilon_{xc}(\omega)$. The excitonic effect also reduces significantly the large discrepancy of both one-electron and RPA calculations between 4.5 and 5 eV.

There exists a large variety of experimental evidence for the existence of exciton effects at critical points. Just to mention a few, Shaklee,

Rowe, and Cardona[55] gave evidence for the contribution of the electron-hole attraction to the observed optical spectrum of InSb. From measurements of the thermoreflectance[39] of a broad spectrum of semiconductors including Si, GaAs, InAs, ZnS, ZnSe, and CdS, it was found necessary to invoke the presence of excitons at all the interband edges, in order to achieve reasonable agreement between theory and experiment. This is also the case in very recent work on the pressure dependence of the electroreflectance spectrum in Ge and GaAs,[40] which closely shows the excitonic contributions to the interband transitions.

Louie *et al.*[28] reported a plasmon-peak lowering and improved agreement with measured energy-loss spectra when just RPA local-field effects are included. In our calculation electron-hole attraction gives again an enhancement of the loss resulting in a plasmon peak strength between $\bar{\epsilon}$ and ϵ_{RPA} calculation.

V. TIGHT-BINDING AND PLANE-WAVE MODELS

In order to get a simple understanding of our findings on the many-particle corrections and to generalize the results to other semiconductors we consider two models.

A. The tight-binding limit

In the special case of a strongly ionic crystal, in which the electrons are so tightly bound that the approximation of nonoverlapping dipole distributions centered at lattice sites is valid, the inversion of the dielectric matrix has been shown to lead to the classical Lorentz-Lorenz model of local-field corrections.[56-59] In all these discussions the crucial point was the *ad hoc* assumption that the dipolar self-interaction is completely removed by exchange and correlation. On the basis of the microscopic theory of the screened electron-hole interaction developed in Sec. II, we are in a position to quantitatively investigate this main assumption of the Lorentz-Lorenz model. We shall demonstrate for an s-p tight-binding model of a covalent crystal that the screened electron-hole interaction $-V^s$ reduces to only about $\frac{1}{2}$ of the self-interaction of the dipolar s and p charges. We derive the explicit dependence of this self-interaction correction on orbital localization, and show that for sufficient localization it gives rise to the continuum-exciton effect. Thus, our purpose here is *not* to add to the already numerous derivations of the Lorentz-Lorenz model but instead give microscopic prescriptions for its validity, or in other words for the size of the self-interaction term. This will provide us with a deeper understanding of the physical content of the continuum-exciton effect.

Let us start by assuming that only the interactions which are due to the overlap of hybridized orbitals $h_{\vec{\nu}}$ and $h_{\vec{\mu}}$ on the same lattice site [(a) and (d) in Fig. 5] are important. It should be noted that the dominant "bond-charge" overlap (e) in Fig. 5 between nearest-neighbor h's is canceled in the main overlap $\phi^*_{\vec{\mu}}(\vec{r})\phi_{\vec{\nu}}(\vec{r})$ of bonding and antibonding orbitals pointing in the same direction. Furthermore, let us take into account only the dipole (s-p) contribution Eq. (3.7) to these charge distributions. Then, after an elementary integration, the dipole form factor $f_\lambda^{\alpha=x}$ can be written as

$$f_\lambda^{\alpha=x}=\{3^{1/2}ab[\pi/(2\alpha)]^{3/2}(16\pi)^{-1}-\tfrac{1}{4}a_0\}W_\lambda^x\,, \tag{5.1}$$

with

$$W_\lambda^x=(\vec{\nu}+\vec{\mu})_x\,, \tag{5.2}$$

where λ stands for $(\vec{\nu},\vec{\mu},\vec{R}_l)$. To simplify the notation with respect to summations over Gaussian indices i and j, the basic orbitals $R_s(r)$ and $R_p(r)$ of Eqs. (3.3) and (3.4) are here expanded in only one Gaussian with the same exponential coefficient α, and $a_i=a$, $b_i=b$.

Using the site representation Eq. (3.9b) and Eqs. (3.11) and (A20) the dipole-dipole interaction matrix $V_{\lambda\lambda'}$ can be written in the factorized form

$$V_{\lambda\lambda'}=W_\lambda^x U_{xx'}W_{\lambda'}^{x'}\,, \tag{5.3}$$

with

$$U_{xx'}=C_1\sum_m \frac{\partial^2}{\partial C_{\lambda,x}\partial C_{\lambda',x'}} \times\int d^3r\int d^3r'\, e^{-2\alpha(\vec{r}-\vec{C}_\lambda-\vec{R}_m)^2} \times v(\vec{r}-\vec{r}')\,e^{-2\alpha(\vec{r}'-\vec{C}_{\lambda'})^2} \tag{5.4}$$

$$=C_2\frac{\partial^2}{\partial C_{\lambda,x}\partial C_{\lambda',x'}}\sum_m \Gamma_0[\alpha(\vec{C}_\lambda-\vec{C}_{\lambda'}+\vec{R}_m)^2] \tag{5.5}$$

$$=C_3T_{xx'}\,, \tag{5.6}$$

where C_1, C_2, and C_3 are constant and $\Gamma_0(t)$ is related to the error function in Eq. (A17). $\vec{C}$ is the center of the s-p overlap which is given by silicon lattice sites

$$\vec{C}_\lambda=\begin{cases}b\vec{\nu}, & \vec{R}_l\neq 0\\ 0, & \vec{R}_l=0\end{cases} \tag{5.7}$$

where R_l is the distance involved in the overlap $\phi^*_{\vec{\nu}}(\vec{r})\phi_{\vec{\mu}}(\vec{r}-\vec{R}_l)$. $T_{xx'}$ is a tensor in the Cartesian

coordinates x and x' and can easily be shown to be invariant under the tetrahedral group of operations. Therefore, T is diagonal in the Cartesian coordinates x and x' (Ref. 61):

$$T_{xx'} = T\delta_{xx'}. \tag{5.8}$$

In principle, the tensor $T_{xx'}$ is still dependent on the indices λ and λ'. As we shall see immediately, this is due to the self-interaction of the charge distribution centered at $\vec{R}_m = 0$ which is contained in $T_{xx'}$ for $C_\lambda - C_{\lambda'} = 0$. This self-interaction can be explicitly calculated by using the definition (A17) of $\Gamma_0(t)$:

$$\Gamma_0(t) = \tfrac{1}{2}(\pi/t)^{1/2}\,\mathrm{erf}(t^{1/2}), \tag{5.9}$$

with the result

$$T_{xx}^{\mathrm{self}} = 2\alpha/3. \tag{5.10}$$

We note that the self-interaction depends on localization and, as is to be expected, becomes stronger if the Gaussians are more localized.

Let us follow for the moment the derivations,[56-59] and suppose that the effect of exchange and correlation correction to the Coulomb interaction $v(\vec{r}-\vec{r}')$, in the density interaction matrix V is to remove the self-interaction of the dipolar s-p charge distribution but to leave it as pure Coulombic between different sites. Then, noting that $\mathrm{erf}(t^{1/2})$ rapidly tends to the limit 1 for larger t [for the Si nearest-neighbor distance $C_\lambda - C_{\lambda'} = b\vec{\nu}$ and a typical Gaussian exponent $\alpha = 0.15$ of atomiclike decay: $\mathrm{erf}(t^{1/2}) = 0.99$], we can replace $T_{xx'}$ in Eqs. (5.5) and (5.6) to a very good degree of accuracy by

$$T_{xx'} = \tfrac{1}{2}(\pi/\alpha)^{1/2}\frac{\partial^2}{\partial C_{\lambda,x}\partial C_{\lambda',x'}}\sum_m \frac{1}{|\vec{C}_\lambda - \vec{C}_{\lambda'} - \vec{R}_m|}. \tag{5.11}$$

Thus, $T_{xx'}$ is, apart from the factor $\frac{1}{2}(\pi/\alpha)^{1/2}$, the $q \to 0$ limit of the ordinary Coulomb-coupling coefficient between point dipoles on sublattices λ and λ' with the dipole self-interaction removed. For cubic and tetrahedral lattices this coefficient is independent of C_λ and $C_{\lambda'}$ and reduces to the standard result[61]

$$T_{xx'} = -\tfrac{4}{3}\pi\Omega_0^{-1}\delta_{xx'}. \tag{5.12}$$

With equations (2.21), (3.15), (5.1), and (5.3) we finally get for the dielectric constant the generalized Lorentz-Lorenz formula

$$\epsilon(\omega) = [1 + \tfrac{8}{3}\pi P(\omega)]/[1 - \tfrac{4}{3}\pi P(\omega)], \tag{5.13}$$

where

$$P(\omega) = \frac{C_2}{2}\Omega_0^{-1}(\pi/\alpha)^{1/2}\sum_{\lambda\lambda'} W_\lambda^x N_{\lambda\lambda'}(\omega) W_{\lambda'}^{x*}. \tag{5.14}$$

$P(\omega)$ may be considered as a measure for the polarizability in the unit cell of volume Ω_0.

After establishing the classical result let us now study in somewhat more detail the influence of the short-range electron-hole interaction on the dielectric constant.

Consider the one-electron dielectric function without local-field corrections $\bar{\epsilon}(\omega)$ which, in the nonoverlapping s-p model, is given by the Drude formula

$$\bar{\epsilon}(\omega) = 1 + 4\pi P(\omega). \tag{5.15}$$

Using Eqs. (5.13) and (5.15) we can write the Lorentz-Lorenz formula in terms of real and imaginary parts (subscripts 1 and 2) of $\bar{\epsilon}(\omega)$:

$$\epsilon_1(\omega) = 1 + \frac{\bar{\epsilon}_1 - 1 - \frac{1}{3}[(\bar{\epsilon}_1 - 1)^2 + \bar{\epsilon}_2^2]}{[1 - \frac{1}{3}(\bar{\epsilon}_1 - 1)]^2 + \frac{1}{9}\bar{\epsilon}_2^2} \tag{5.16}$$

and

$$\epsilon_2(\omega) = \frac{\bar{\epsilon}_2}{[1 - \frac{1}{3}(\bar{\epsilon}_1 - 1)]^2 + \frac{1}{9}\bar{\epsilon}_2^2}. \tag{5.17}$$

We note that the local-field corrections in this Lorentz-Lorenz limit enhance $\epsilon_2(\omega)$ above $\bar{\epsilon}_2(\omega)$, provided

$$[1 - \tfrac{1}{3}(\bar{\epsilon}_1 - 1)]^2 < 1 - \tfrac{1}{9}\bar{\epsilon}_2^2, \tag{5.18}$$

a condition, which is usually fulfilled at energies just above the threshold, whereas for higher energies, the opposite is true (see the results in Sec. IV). Thus, structures in $\epsilon_2(\omega)$ are shifted to lower energies and the Lorentz-Lorenz model can partly resolve the common disagreement between one-electron theoretical and experimental absorption for lower energies.

In the next step we consider the behavior of $\epsilon_2(\omega)$ if the self-interaction is not excluded from the dipole interaction matrix V. This then gives the RPA local-field effect. Within the above point dipole model [$\mathrm{erf}(t^{1/2}) \to 1$] we have from Eq. (5.11)

$$T_{xx'}^{\lambda\lambda'} = \tfrac{1}{2}\left(\frac{\pi}{\alpha}\right)^{1/2}\tfrac{4}{3}\pi\Omega_0^{-1}\left[-1 + \frac{\Omega_0\alpha^{3/2}}{\pi^{3/2}}\delta_{\vec{C}_\lambda\vec{C}_{\lambda'}}\right]\delta_{xx'}. \tag{5.19}$$

Owing to the self-interaction $\sim\delta_{\vec{C}_\lambda\vec{C}_{\lambda'}}$, the tensor T depends on the indices λ and λ' and therefore, it is only approximately possible to derive a Lorentz-Lorenz-type formula for the dielectric constant. But the average

$$T_{xx'} = \left(\sum_{\lambda'} T_{xx'}^{\lambda\lambda'} \Big/ \sum_{\lambda'}\right)\delta_{xx'}, \tag{5.20}$$

where $\sum_{\lambda'}$ denotes the total number of possible indices, λ' (28 in our example Si) can be shown to give a very accurate replacement for the actual $T_{xx'}^{\lambda\lambda'}$. With this we can proceed as in deriving the Lorentz-Lorenz result and have

$$\epsilon(\omega) = 1 + 4\pi P(\omega)/[1 - \tfrac{4}{3}\pi I P(\omega)] \tag{5.21}$$

and thus for the imaginary part

$$\epsilon_2(\omega) = \frac{\bar{\epsilon}_2}{[1 + \frac{1}{3} I(\bar{\epsilon} - 1)]^2 + I^2 \frac{1}{9} \bar{\epsilon}_2^2}, \tag{5.22}$$

where

$$I = I_{\mathrm{RPA}} = -1 + \frac{\Omega_0 \alpha^{3/2}}{\pi^{3/2}} \frac{1}{28} \sum_{\lambda'} \delta_{\vec{c}_\lambda \vec{c}_{\lambda'}}, \tag{5.23}$$

is, apart from an unimportant factor, the RPA contribution to the interaction I of Eq. (2.3). In Eq. (5.23) it measures the deviation from the Lorenz local field due to inclusion of the self-interaction. The averaging factor is $\frac{4}{7}$ for $\lambda = (\vec{\nu}, \vec{\mu}, \vec{R}_l = 0)$ and $\frac{3}{7}$ for $\lambda = (\vec{\nu}, \vec{\mu}, \vec{R}_l \neq 0)$ and in the following is set equal to $\frac{1}{2}$. So the enhancement in $\epsilon_2(\omega)$ over $\bar{\epsilon}_2(\omega)$ occurs if

$$[1 + \tfrac{1}{3} I(\bar{\epsilon}_1 - 1)]^2 < 1 - \tfrac{1}{9} I^2 \bar{\epsilon}_2^2. \tag{5.24}$$

Taking an atomiclike[44] Gaussian exponent $\alpha = 0.135$, $\Omega_0 = 270.25$ a.u. we see that $I_{\mathrm{RPA}} = 0.2$ becomes positive and the Coulomb interaction is repulsive. Thus, in the RPA model, where the dipole self-interaction is not excluded, local-field effects usually enhance $\epsilon_2(\omega)$ above $\bar{\epsilon}_2(\omega)$ for higher energies ω, where $\bar{\epsilon}_1$ becomes negative, whereas for lower energies $\epsilon_{\mathrm{RPA},2}(\omega)$ is decreased below $\bar{\epsilon}_2(\omega)$. From our Green's-function treatment of Sec. II it is, however, clear that it is not the *full* self-interaction of a dipole charge distributions which determines the relative shift of $\epsilon_2(\omega)$, with respect to $\bar{\epsilon}_2(\omega)$.

Taking the Frenkel-type "electron-hole interaction in the same bond" into account it is obvious that the electron-hole attraction in the s-p tight-binding model amounts to subtracting from the RPA local-field factor I_{RPA} in Eq. (5.23) just $\frac{1}{2}$ of the screened self-interaction

$$I^s \cong \frac{\Omega_0 \alpha^{3/2}}{4\pi^{3/2}} \epsilon_s^{-1}; \tag{5.25}$$

ϵ_s^{-1} stands for the screening reduction which is of the order of 1.

Combining the electron-hole correction Eq. (5.25), with the RPA local-field result in Eq. (5.23) and using $\epsilon_s^{-1} = 1$, $\alpha = 0.135$, and the Si parameters we have $I = I_{\mathrm{RPA}} - I^s = -0.4$ and thus the Coulomb plus electron-hole interaction $(V - V^s)$ becomes attractive. As a consequence, we have the continuum-exciton effect enhancing $\epsilon_2(\omega)$ above $\bar{\epsilon}_2(\omega)$ for lower energies ω.

From Eqs. (5.21) and (5.22) it is straightforward to show that if the imaginary part of the polarization $P(\omega) \sim N^0(\omega)$ is approximately a Lorentzian centered about the critical-point energy, then $\epsilon_2(\omega)$ is a Lorentzian with its peak moved in energy by $V - \frac{1}{2}V^s$, which is given by $\frac{4}{3}\pi I$. The attractive nature of $V - V^s$ obviously depends on localization. The simple analytic dependence of I on the volume Ω_0 may be used to approximately study the local-field and excitonic effects under pressure.

B. The plane-wave model

Let us finally try to understand our findings on the many-particle corrections in a simple but fairly general way by invoking a two-plane-wave model.

The origins of the E_1 and E_2 peaks in a semiconductor are given in the one-electron approximation by the model of Cardona and Pollak.[4] The E_2 peak is calculated by a two-plane-wave model which produces the gap on the whole Jones-Zone surface. The E_1 peak is calculated by the model of the two parallel bands along $\langle 111 \rangle$ directions excluding a region around the Γ point. For simplicity, let us neglect the spin-orbit splitting of the valence band. The sum of these two contributions yields N^0, now in plane-wave basis rather than the local-orbital basis. The excitonic effect due to the screened Coulomb interaction is qualitatively of the form (2.20) with V set at zero, since for extended wave functions the local-field effect is negligible. Because the E_2 peak has an inverse square-root singularity, by the Kramers-Kronig relation the real part of $(1 + \frac{1}{2}V^s N^0)$ in the same energy range is always larger than unity. Therefore, the exciton effect reduces the strength of the E_2 absorption peak as is also shown in the local-orbital result in Fig. 7. Since the model E_1 peak has a step-function discontinuity, Kramers-Kronig analysis yields the real part of $(1 + \frac{1}{2}V^s N^0)$ in the E_1 energy range always reduced below unity. Hence, the excitonic effect increases the strength of the E_1 peak.

VI. CONCLUSIONS

The basic aim of the present work is to investigate deviations from the one-particle picture of optical excitations in a semiconductor. The dominant corrections to the one-particle absorption are due to electron-hole attraction and its exchange counterpart. The classical exciton concept points to the importance of the electron-hole attraction in the limits of the Wannier and Frenkel model. In the continuum an intermediate picture is more appropriate. The exciton exchange is due to the Coulomb repulsion between the induced charge distributions. It creates the RPA local-field effect and is responsible for the motion of the exciton in a crystal. A Green's-function treatment makes possible a very general

formulation of electron-hole interaction.

The combined electron-hole interactions introduce in the example Si significant (up to a factor of 2 enhancement on the low-energy side) modifications of the theoretical absorption strength. This quite generally seems to explain the well-known deviations of one-particle spectra in semiconductors from experiment. The electron-hole attraction is the dominant many-particle correction in the optical response to a transverse probe. When combined with the electron-hole exchange it introduces shifts of low-energy absorption structure compared to the one-electron interband transitions. The maximum shift is about $\omega_{\text{gap}}/10$ in Si compared to a value of $\omega_{\text{gap}}/4$ which we found in our earlier diamond work.[26] It thus depends on the localization properties of the electrons and seems to indicate that the empirical band structure adjustment works rather well from Ge on to smaller gap semiconductors. Here we still expect significant modifications in oscillator strength of the optical absorption.

ACKNOWLEDGMENTS

We are grateful to Professor H. Bilz, Professor M. Cardona, Professor P. Fulde, and Dr. G. Strinati, Dr. T. Arya, Dr. H. -J. Mattausch, and Dr. A. Muramatsu for many stimulating discussions. One of us (L. J. S.) would like to thank Professor H. Bilz and his colleagues for the hospitality at the Max-Planck-Institut für Festkörperforschung, Stuttgart and the Humboldt Foundation for an award.

APPENDIX: CALCULATION OF THE ELECTRON-HOLE ATTRACTION MATRIX - V^s AND THE DENSITY INTERACTION MATRIX V

It is clear from the defining formula [Eq. (2.19)] that the screened electron-hole matrix element V^s can be reduced in the Gaussian representation to a sum over four-center integrals of the type

$$\int d^3r \int d^3r' \int d^3r'' e^{-a(\vec{r}-\vec{A})^2} e^{-b(\vec{r}-\vec{B})^2} \frac{1}{|\vec{r}-\vec{r}''|} \times \epsilon^{-1}(\vec{r}''-\vec{r}') e^{-c(\vec{r}'-\vec{C})^2} e^{-d(\vec{r}'-\vec{D})^2} . \tag{A1}$$

The terms containing p-type orbitals $xR_p(r)$ from Eq. (3.4) can be derived from (A1) by the application of differential operators to the basic s-type formula. This procedure is based on the equation[60]

$$(x-A_x) e^{-a(\vec{r}-\vec{A})^2} = -\frac{1}{2a}\frac{\partial}{\partial A_x} e^{-a(\vec{r}-\vec{A})^2} . \tag{A2}$$

The local dielectric screening $\epsilon^{-1}(\vec{r}-\vec{r}')$ in Eq. (A1) is expanded in a Gaussian series as follows:

$$\int d^3r'' |\vec{r}-\vec{r}''|^{-1} \epsilon^{-1}(\vec{r}''-\vec{r}') = \left(\epsilon_0^{-1} + (1-\epsilon_0^{-1}) \sum_i A_i e^{-\alpha_i(\vec{r}-\vec{r}')^2}\right) |\vec{r}-\vec{r}'|^{-1} , \tag{A3}$$

with

$$\sum_i A_i = 1 . \tag{A4}$$

In our Si calculation we take the macroscopic dielectric constant $\epsilon_0 = 11.7$ and fit a two-Gaussian expansion to the Fourier-transformed ϵ^{-1}:

$$\int d^3r'' |r-r''|^{-1} \epsilon^{-1}(r''-r') = \frac{1}{2\pi^2} \int \frac{e^{-i\vec{q}\cdot\vec{r}}}{q^2 \epsilon(q)} d^3q , \tag{A5}$$

where $\epsilon(q)$ is approximated by the averaged (over Δ, Σ, Λ directions) screening function of Walter and Cohen.[51]

Using the representation

$$\frac{1}{r} e^{-\alpha r^2} = (\pi)^{-1/2} \int_0^\infty s^{-1/2} e^{-(\alpha+s)r^2} ds , \tag{A6}$$

the r-dependent screening contributions can be written as

$$[aAbB; cCdD]_\alpha = \int d^3r \int d^3r' e^{-a(\vec{r}-\vec{A})^2} e^{-b(\vec{r}-\vec{B})^2} \frac{e^{-\alpha|\vec{r}-\vec{r}'|^2}}{|\vec{r}-\vec{r}'|} e^{-c(\vec{r}'-\vec{C})^2} e^{-d(\vec{r}'-\vec{D})^2}$$

$$= (\pi)^{-1/2} \exp\left(-\frac{ab}{a+b}(\vec{A}-\vec{B})^2 - \frac{cd}{c+d}(\vec{C}-\vec{D})^2\right) \int_0^\infty I_x I_y I_z s^{-1/2} ds , \tag{A7}$$

where

$$I_x = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dx' e^{-p(x-P_x)^2} e^{-q(x'-Q_x)^2} e^{-(s+\alpha)(x-x')^2}$$

$$= \pi(pq)^{-1/2} \left(1 + \frac{p+q}{pq}(s+\alpha)\right)^{-1/2} \exp\left(-\frac{(s+\alpha)(P_x-Q_x)}{1+(s+\alpha)(p+q)/(pq)}\right) , \tag{A8}$$

with $p=a+b$, $q=c+d$. The points $\vec{P}$ and $\vec{Q}$ are given by

$$P_x=(aA_x+bB_x)/(a+b)\,,\quad Q_x=(cC_x+dD_x)/(c+d)\,. \tag{A9}$$

In deriving (A7) we used that the product of two Gaussians having different centers $\vec{A}$ and $\vec{B}$ is itself a Gaussian,

$$e^{-a(\vec{r}-\vec{A})^2}e^{-b(\vec{r}-\vec{B})^2}=\kappa e^{-p(\vec{r}-\vec{P})^2}\,, \tag{A10}$$

with p and $\vec{P}$ defined above, and

$$\kappa=\exp[-(\vec{A}-\vec{B})^2ab/(a+b)]\,. \tag{A11}$$

Combining (A7) and (A8) we have

$$[aAbB;cCdD]_\alpha=\pi^{5/2}(pq)^{-3/2}\exp\left(-\frac{ab}{a+b}(\vec{A}-\vec{B})^2-\frac{cd}{c+d}(\vec{C}-\vec{D})^2\right)$$
$$\times\int_0^\infty ds\left(1+\frac{p+q}{pq}(s+\alpha)\right)^{-3/2}s^{-1/2}\exp\left(-\frac{(s+\alpha)(\vec{P}-\vec{Q})^2}{1+(s+\alpha)(p+q)/(pq)}\right). \tag{A12}$$

With the substitution,

$$1+(s+\alpha)(p+q)/(pq)=(u-t^2)^{-1}\,, \tag{A13}$$

where

$$u=[1+(p+q)\alpha/(pq)]^{-1} \tag{A14}$$

follows

$$[1+(s+\alpha)(p+q)/(pq)]^{-3/2}s^{-1/2}ds=2u^{1/2}\left(\frac{pq}{p+q}\right)^{1/2}dt\,. \tag{A15}$$

Using this result in (A12), after some elementary transformations, we arrive at the final expression

$$[aAbB;cCdD]_\alpha=2\pi^{5/2}u[pq(p+q)^{1/2}]^{-1}$$
$$\times\exp\left(-\frac{ab}{a+b}(\vec{A}-\vec{B})^2-\frac{cd}{c+d}(\vec{C}-\vec{D})^2-\frac{pq}{p+q}(1-u)(\vec{P}-\vec{Q})^2\right)\Gamma_0\left(\frac{pqu(\vec{P}-\vec{Q})^2}{p+q}\right), \tag{A16}$$

where the function $\Gamma_0(t)$ is the $(m=0)$ case of the class of integrals[60]

$$\Gamma_m(t)=\int_0^1 u^{2m}e^{-tu^2}du\quad(t>0;\,m=0,1,2,\ldots)\,. \tag{A17}$$

They are closely related to the error function $\mathrm{erf}(t^{1/2})$ and its derivatives, specifically

$$\Gamma_0(t)=\tfrac{1}{2}(\pi/t)^{1/2}\,\mathrm{erf}(t^{1/2})\,. \tag{A18}$$

The integrals $\Gamma_m(t)$ can be expressed in terms of $\Gamma_0(t)$ by means of the recurrence formula[60]

$$\Gamma_m(t)=(2m+1)^{-1}[2t\Gamma_{m+1}(t)+e^{-t}]\,. \tag{A19}$$

In combination with the relation

$$(d/dt)\Gamma_m(t)=-\Gamma_{m+1}(t)\,, \tag{A20}$$

(A19) allows for an analytic calculation also of p-type contributions to V^s, which are expressed in Eq. (A2) as derivatives with respect to Gaussian centers. The result (A16) for the four-center integral $[aAbB;cCdD]$ can be used to calculate the complete screening integral (A1). The r-independent screening contribution proportional to ϵ_0^{-1} immediately follows by taking the $\alpha\to 0$ limit $(u\to 1)$ in Eq. (A15):

$$[aAbB;cCdD]_{\alpha=0}$$
$$=2\pi^{5/2}[pq(p+q)^{1/2}]^{-1}$$
$$\times\exp\left(-\frac{ab}{a+b}(\vec{A}-\vec{B})^2-\frac{cd}{c+d}(\vec{C}-\vec{D})^2\right)$$
$$\times\Gamma_0\left(\frac{pq(\vec{P}-\vec{Q})^2}{p+q}\right). \tag{A21}$$

This, of course, is also the expression which generates the Coulomb matrix V.

*Permanent address: University of California, San Diego, La Jolla, California 92093.
[1]J. C. Phillips, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1966), Vol. 18, p. 55.
[2]F. Bassani and G. P. Parravacini, *Electronic States and Optical Transitions in Solids*, edited by R. A. Bullinger (Pergamon, New York, 1975).
[3]M. L. Cohen and V. Heine, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1970), Vol. 24, p. 1.
[4]M. Cardona and F. Pollak, Phys. Rev. 142, 530 (1966).
[5]D. L. Greenaway and G. Harbeke, *Optical Properties and Band Structure of Semiconductors*, edited by B. R. Pamplin (Pergamon, New York, 1968).
[6]R. J. Elliott, Phys. Rev. 108, 1384 (1957).
[7]J. C. Phillips, Phys. Rev. 136, A1705 (1964).
[8]B. Velicky and J. Sak, Phys. Status Solidi 16, 147 (1966).
[9]E. O. Kane, Phys. Rev. 180, 852 (1969).
[10]G. F. Koster and J. C. Slater, Phys. Rev. 95, 1667 (1954).
[11]Y. Toyozawa, M. Inoue, T. Jnui, M. Ozaki, and E. Hanamura, J. Phys. Soc. Jpn. 22, 1327 and 1349 (1967).
[12]J. Hermanson, Phys. Rev. 166, 893 (1966).
[13]R. M. Martin, J. A. Van Vechten, J. E. Rowe, and D. E. Aspnes, Phys. Rev. B 6, 2500 (1972), and refereneces therein.
[14]K. Cho, Phys. Rev. B 14, 4463 (1976); Inst. Phys. Conf. Ser. 43, 841 (1979).
[15]M. H. Denisov and V. P. Makarov, Phys. Status Solidi B 56, 9R (1973).
[16]U. Rössler, in *Rare Gas Solids*, edited by M. L. Klein and J. A. Venables (Academic, London, 1976), Vol. 1, and references in this review.
[17]M. Altarelli and F. Bassani, J. Phys. C 4, L328 (1971).
[18]L. Resca and S. Rodriquez, Phys. Rev. B 17, 3334 (1978).
[19]A short version of this work has been published by W. Hanke and L. J. Sham in Phys. Rev. Lett. 43, 387 (1979).
[20]For recent reviews of the concepts of dielectric response see W.Hanke, Adv. Phys. 27, 287 (1978) and *Festkörperprobleme XIX, Advances in Solid State Physics*, edited by J. Treusch (Vieweg, Braunschweig, 1979), p. 43.
[21]L. J. Sham and T. M. Rice, Phys. Rev. 144, 807 (1966).
[22]S. L. Adler, Phys. Rev. 126, 413 (1962); N. Wiser, *ibid.* 129, 62 (1963).
[23]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; Phys. Rev. B 8, 4585 (1973); 8, 3591 (1973).
[24]L. J. Sham, Phys. Rev. B 6, 3584 (1972).
[25]J. A. Van Vechten and R. M. Martin, Phys. Rev. Lett. 28, 446 (1972).
[26]W. Hanke and L. J. Sham, Phys. Rev. B 12, 4501 (1975).
[27]K. Arya and S. Jha, Phys. Rev. B 10, 4485 (1974).
[28]S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. Lett. 34, 155 (1975).
[29]N. E. Brener, Phys. Rev. B 12, 1487 (1975).
[30]P. E. Van Camp, V. E. Van Doren, and J. T. Devreese, J. Phys. C 9, L79 (1976).
[31]R. D. Turner and J. C. Inkson, J. Phys. C 9, 3583 (1976).
[32]W. Schäfer, M. Schreiber, and J. Treusch, Inst. Phys. Conf. Ser. 43, 1211 (1979).
[33]F. Nizzoli, J. Phys. C 11, 673 (1978).
[34]See also S. K. Sinha, in *Dynamical Properties of Solids*, edited by G. K. Horton and A. A. Maradudin (North-Holland, Amsterdam, 1979), Vol. 3.
[35]W. Kohn and L. J. Sham, Phys. Rev. 140, A1133 (1965).
[36]M. Cardona and F. H. Pollak, in *Physics of Optoelectronic Materials*, edited by W. A. Albers (Plenum, New York, 1971), p. 111.
[37]J. L. Freeouf, Phys. Rev. B 7, 3810 (1972).
[38]M. Welkowsky and R. Braunstein, Phys. Rev. B 5, 497 (1972).
[39]E. Matatagni, A. G. Thompson, and M. Cardona, Phys. Rev. 176, 950 (1968).
[40]M. Chandrasekhar and F. H. Pollak, Phys. Rev. B 15, 2127 (1977).
[41]O. K. Andersen, Phys. Rev. B 12, 3060 (1975).
[42]W. R. Heller and A. Marcus, Phys. Rev. 84, 852 (1951).
[43]J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1954).
[44]E. O. Kane and A. B. Kane, Phys. Rev. B 17, 2691 (1978).
[45]D. J. Chadi, Phys. Rev. B 16, 3572 (1977).
[46]W. Kohn, Phys. Rev. B 7, 4388 (1973).
[47]B. Ortenburger and W. E. Rudge, IBM Research report (unpublished).
[48]J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B 10, 5095 (1974).
[49]H. Hattori, H. Kuriyama, T. Katayawa, and M. Koto, J. Phys. Soc. Jpn. 20, 988 (1965).
[50]P. W. Anderson, Phys. Rev. Lett. 21, 13 (1968).
[51]J. P. Walter and M. L. Cohen, Phys. Rev. B 2, 1821 (1970).
[52]H. R. Phillip and H. Ehrenreich, Phys. Rev. 129, 1550 (1963).
[53]R. R. L. Zucca, J. P. Walter, Y. R. Shen, and M. L. Cohen, Solid State Commun. 8, 627 (1970).
[54]A. Daunois and D. E. Aspnes, Phys. Rev. B 18, 1824 (1978).
[55]K. L. Shaklee, J. E. Rowe, and M. Cardona, Phys. Rev. 174, 828 (1968).
[56]S.K. Sinha, R. P. Gupta, and D. L. Price, Phys. Rev. B 9, 2564 (1974).
[57]Y. Onodera, Prog. Theor. Phys. 49, 37 (1973).
[58]D. L. Johnson, Phys. Rev. B 9, 4475 (1974).
[59]E. G. Maksimov and I. I. Mazin, Solid State Commun. 27, 527 (1978).
[60]I. Shavitt, in *Methods in Computational Physics*, edited by B. Alder, S. Feshbach, and M. Rotenberg (Academic, New York, 1963), p. 1.
[61]M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, New York, 1958), p. 398.