High-frequency dielectric properties of covalent semiconductors within the nearly-free-electron approximation. I. The one-plasmon-band model

Kurt Sturm
Institut für Festkörperforschung der Kernforschungsanlage, Jülich, West Germany

Luiz E. Oliveira*
Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, United Kingdom

(Received 30 April 1980)

The macroscopic dielectric response function $\epsilon_M(\vec{k},\omega)$ is calculated analytically for covalent semiconductors in the high-frequency limit. The effects of the periodic crystal potential are included by second-order perturbation theory within the self-consistent-field approximation. A study of the plasmon line shape, i.e., of the energy-loss function $\mathrm{Im}[-1/\epsilon_M(0,\omega)]$ and of the optical absorption $\mathrm{Im}\epsilon_M(0,\omega)$ demonstrates the importance of local-field corrections in covalent semiconductors. In the long-wavelength limit, the theory predicts that the unknown effective pseudopotential form factors $U_{\vec{G}}$ for reciprocal-lattice vectors with $|\vec{G}| \geq |\vec{G}_{400}|$ may be obtained successively for increasing $|\vec{G}|$ and increasing excitation energies $\hbar\omega$ from energy-loss experiments or optical measurements. Available experimental data in two cases (Si and Ge) show absorption edges near the predicted energies associated with particular $U_{\vec{G}}$. For the first time, the anisotropic plasmon dispersion has been calculated for semiconductors and is compared with recent loss experiments on single crystals of Si, GaAs, and InSb. Owing to the similarity of the crystal potential in these compounds, the theory predicts the same trends for all three materials, namely: $\omega_p^{[100]}(\vec{k}) > \omega_p^{[110]}(\vec{k}) > \omega_p^{[111]}(\vec{k})$. For Si and GaAs this is in agreement with experiment whereas unexpectedly the experiment shows that $\omega_p^{[110]}(\vec{k}) > \omega_p^{[100]}(\vec{k}) > \omega_p^{[111]}(\vec{k})$ for InSb.

I. INTRODUCTION

In recent years, due to improved experimental techniques, electron-energy-loss spectroscopy (ELS) has been increasingly used to study high-frequency electronic excitations in solids.[1-3] The "classic" electronic excitation studied by this technique is the plasmon, which is the quantum of the longitudinal electron density fluctuations of the conduction or valence electrons in a solid. Electron-energy-loss spectroscopy has also been used to investigate interband transitions, core-electron excitations, and excitons. Here, we shall be concerned with plasmons and their interactions with interband transitions and local fields which are both a consequence of the periodic crystal potential.

In the self-consistent-field approximation (SCF) the response of the electrons to a weak scalar external potential can be calculated from the microscopic dielectric function which is obtained from the response of independent particles to a total potential. This total potential consists of the external potential and the potential induced by the charge-density fluctuations due to the external perturbation. If exchange and correlation effects are neglected, the induced potential is related to the charge-density fluctuations by Poisson's equation.

Within the homogeneous electron-gas model in which the periodic potential of the positive ions is "smeared" out to a structureless positive background that neutralizes the total charge of the electron gas, the dielectric response is given by a simple scalar function, the well-known (longitudinal) Lindhard dielectric function[4] $\epsilon_L(q,\omega)$. The assumption of quasihomogeneity, i.e., the generalization to Bloch electrons in a real crystal, leads to the Ehrenreich-Cohen dielectric functions[5] $\epsilon_{EC}(\vec{q},\omega)$, which for many years formed the basis of our understanding of the dielectric properties such as the optical properties of metals and semiconductors.

During the last ten years, however, it has become increasingly clear that $\epsilon_{EC}(0,\omega)$ is insufficient to describe accurately the optical properties of covalent semiconductors. The problem arises from the assumption of quasihomogeneity which means that the induced potential $\phi_{\mathrm{ind}}(\vec{r})$ has the same wavelength as the external potential

$$\phi_{\mathrm{ext}}(\vec{r}) = \phi^{\mathrm{ext}}_{\vec{q}} e^{i\vec{q}\cdot\vec{r}} . \tag{1.1}$$

However, in a crystal the unperturbed valence- or conduction-electron density is really a periodic function of position, although in simple metals this modulation is sufficiently weak to justify the assumption of quasihomogeneity as borne out by detailed calculations of the energy-loss spectra of simple metals.[6] In covalent semiconductors, however, the covalent-bond charges are a clear manifestation of the inhomogeneous distribution

of the valence electrons. This inhomogeneity is reflected in the dielectric response giving rise to an induced potential

$$\phi_{\mathrm{ind}}(\vec{r}) = \phi_{\vec{q}}^{\mathrm{ind}}\, e^{i\vec{q}\cdot\vec{r}} + \sum_{\vec{G}\neq\vec{0}} \phi_{\vec{q}+\vec{G}}^{\mathrm{ind}}\, e^{i(\vec{q}+\vec{G})\cdot\vec{r}}, \qquad (1.2)$$

where $\vec{G}$ is a reciprocal-lattice vector. The $\phi_{\vec{q}+\vec{G}}^{\mathrm{ind}}$ arise from local fields which are set up by the short-wavelength charge-density fluctuations $\rho(\vec{q}+\vec{G})$. Adler and Wiser[7] have shown that the inclusion of the microscopic local fields leads to the definition of a microscopic dielectric matrix $[\epsilon(\vec{k}+\vec{G}, \vec{k}+\vec{G}', \omega)]$ given by

$$\epsilon(\vec{k}+\vec{G}, \vec{k}+\vec{G}', \omega) = \delta_{\vec{G},\vec{G}'} - \frac{4\pi e^2}{|\vec{k}+\vec{G}|^2} \sum_{\sigma,\sigma'} \frac{n(E_\sigma) - n(E_{\sigma'})}{\hbar\omega + i\delta + E_\sigma - E_{\sigma'}} \langle\sigma|\, e^{-i(\vec{k}+\vec{G})\cdot\vec{r}}\,|\sigma'\rangle\langle\sigma'|\, e^{i(\vec{k}+\vec{G}')\cdot\vec{r}}\,|\sigma\rangle \qquad (1.3)$$

in which $\vec{k}$ is restricted to the first Brillouin zone and $|\sigma\rangle, E_\sigma$ denote single-particle Bloch states with σ as shorthand for all the quantum numbers. All measurable quantities such as the optical properties and the loss function may then be obtained by inversion of this dielectric matrix. Thus if we define a macroscopic dielectric function $\epsilon_M(\vec{q}, \omega)$ as the quantity that would be used in Maxwell's equations, we find that

$$\epsilon_M(\vec{q}, \omega) = 1/\epsilon^{-1}(\vec{k}+\vec{G}_0, \vec{k}+\vec{G}_0, \omega)\ , \qquad (1.4)$$

with $\vec{q} = \vec{k} + \vec{G}_0$. In particular, for the optical properties

$$\epsilon_M(0, \omega) = 1/\epsilon^{-1}(0, 0, \omega)\ . \qquad (1.5)$$

Here $\epsilon^{-1}(\vec{k}+\vec{G}_0, \vec{k}+\vec{G}_0, \omega)$ or $\epsilon^{-1}(0, 0, \omega)$ is an element of the inverted matrix so that it should be noted, for example, that both (1.4) and (1.5) are the inverse of an element of the inverse dielectric matrix. The local-field effects are manifest in the off-diagonal elements of the dielectric matrix so, if these effects are neglected, the matrix becomes diagonal and the problem reduces to

$$\epsilon_M(\vec{q}, \omega) = \epsilon_{\mathrm{BC}}(\vec{q}, \omega) = \epsilon(\vec{k}+\vec{G}_0, \vec{k}+\vec{G}_0, \omega), \qquad (1.6)$$

or to $\epsilon_L(q, \omega)$ if the periodicity of the crystal potential is neglected completely.

As we have said, local-field effects cannot be neglected in covalent semiconductors, so a calculation of their dielectric properties in general entails an extensive numerical problem. As is common in any band-structure calculation one has to limit oneself to some finite energy (or frequency) range which, together with the desired accuracy of the required $\epsilon_M(\vec{q}, \omega)$, ultimately defines the size of the electronic band-structure problem which provides the input for the calculation of the dielectric matrix. A typical calculation where these problems have been considered in detail is the investigation by Louie *et al.*[8] of the long-wavelength dielectric properties of Si for values up to the plasma frequency ω_p. Their work revealed very clearly the importance of local-field effects for high frequencies, in particular, in the evaluation of the energy-loss function $\mathrm{Im}[-1/\epsilon_M(0, \omega)]$.

It is precisely this high-frequency regime, i.e., where $\hbar\omega \gg$ the gap energy, E_{gap}, which is considered in the present investigation. However, whereas most of the previous work is restricted to the long-wavelength limit, we consider finite $\vec{q}$ in order to study the effect of the periodic crystal potential on the plasmon dispersion and compare our results with recent loss experiments on single crystals of Si,[9] GaAs, and InSb.[10]

In a previous Letter by one of the authors[11] the energy-loss function at $q = 0$ was calculated for a series of covalent semiconductors within what we would now term an effective "one-plasmon-band model." The effective 1×1 dielectric matrix obtained[6,7] by folding down the infinite dielectric matrix was calculated to second order in the crystal potential in the nearly-free-electron approximation (NFE). The results for Si were found in good agreement with those of Louie *et al.*[8]

In contrast with most simple metals, the plasmon dispersion in semiconductors reaches the Brillouin-zone boundary—where plasmon band splitting should occur—for $|\vec{k}| < q_c$, where q_c is the cutoff wave vector for plasmons in the homogeneous electron-gas model. As is frequent in ordinary electron-band-structure problems a single-band approximation breaks down in the neighborhood of the Brillouin- (or Jones-) zone boundary. Near the boundary, at least an effective "two-plasmon-band model" is needed, which would be obtained by folding down the dielectric matrix to a 2×2 matrix. Our present calculation in which we fold down the dielectric matrix to a 1×1 matrix is therefore restricted to $\vec{k}$ values well within the first Brillouin zone. A following publication[12] will be devoted to a detailed analysis of plasmon bands and band splittings at the zone boundary and to whether they can be observed in covalent semiconductors.

In Sec. II we briefly define the one-plasmon-band approximation. The dielectric properties in

the long-wavelength limit are discussed in Sec. III together with the derivation of $\epsilon_M(0,\omega)$, which to second order in the crystal potential is represented by a particularly simple formula. We also discuss there the possibility of estimating the effective crystal potential form factors from optical- or energy-loss experiments by exploiting the properties of the simple formula for $\epsilon_M(0,\omega)$. The increasing importance of local-field corrections with increasing frequency is shown by evaluating the optical absorption $\mathrm{Im}\,\epsilon_M(0,\omega)$ for Si and the loss function $\mathrm{Im}[-1/\epsilon_M(0,\omega)]$ for a number of semiconductors of the diamond or zinc-blende structures. Comparison with existing experimental data is made. In Sec. IV the more elaborated but still analytic extension of the theory to finite $\vec{k}$ is outlined. The theory is applied to study the change of the dielectric properties with increasing $\vec{k}$ and in particular the anisotropy of the plasmon dispersion as observed by recent ELS experiments on Si, GaAs, and InSb. General conclusions are drawn in Sec. V and remaining problems, in particular with respect to the plasmon dispersion, are discussed.

II. THE ONE-PLASMON-BAND APPROXIMATION

As discussed in the Introduction, the effect of the periodic potential on the dielectric properties results in a dielectric matrix in $\vec{k}$ space where rows and columns are labeled by reciprocal-lattice vectors $\vec{G}$, $\vec{G}'$. In order to make this more apparent and also to save some writing, we introduce the notation

$$\epsilon_{\vec{G},\vec{G}'}(\vec{k},\omega)\equiv\epsilon(\vec{k}+\vec{G},\ \vec{k}+\vec{G}',\omega).$$

The plasmon dispersion in the homogeneous electron-gas model is defined by

$$\epsilon_L(q,\omega)=0\ , \tag{2.1}$$

which corresponds to a pole in the loss function $\mathrm{Im}[-1/\epsilon_L(q,\omega)]$. In calculating the loss function for a real crystal, however, we have to invert the dielectric matrix. Since every element of the inverted dielectric matrix is inversely proportional to the determinant of this matrix, plasmons are now defined as the poles of

$$\det[\epsilon_{\vec{G},\vec{G}'}(\vec{k},\omega)]=0\ , \tag{2.2}$$

which defines a band-structure problem.[13] As long as $\vec{k}$ is sufficiently well inside the first zone and as we are only concerned with the lowest plasmon band, the problem can be reduced to an effective 1×1 dielectric matrix. If we keep only terms to second order in the effective crystal potential,[6,7]

$$\epsilon_M(\vec{k},\omega)=\epsilon_{\vec{0},\vec{0}}(\vec{k},\omega)-\sum_{\vec{G}\neq\vec{0}}\frac{\epsilon_{\vec{0},\vec{G}}(\vec{k},\omega)\epsilon_{\vec{G},\vec{0}}(\vec{k},\omega)}{\epsilon_L(|\vec{k}+\vec{G}|,\omega)}\ . \tag{2.3}$$

The first term of Eq. (2.3) is the Ehrenreich-Cohen approximation[5] to the dielectric properties, which is sufficiently accurate for most simple metals. The second term of (2.3) depends on off-diagonal elements of the dielectric matrix and includes the local-field effects. Since the off-diagonal elements are proportional to the Fourier components $U_{\vec{G}}$ of the effective crystal potential,

$$V(\vec{r})=\sum_{\vec{G}}U_{\vec{G}}\,e^{-i\vec{G}\cdot\vec{r}}\ , \tag{2.4}$$

in keeping only terms to second order in $U_{\vec{G}}$ we have replaced $\epsilon_{\vec{G},\vec{G}}(\vec{k},\omega)$ in the denominator by its zero-order approximation, namely by $\epsilon_L(|\vec{k}+\vec{G}|,\omega)$.

It should be noted that a calculation to second order in the effective potential is only applicable to energies $\hbar\omega$ that are large compared to the gap. The plasmon energy in covalent semiconductors satisfies this condition and explains why the plasma frequency is fairly accurately determined by the homogeneous electron gas value $\omega_p^0=(4\pi n_0e^2/m)^{1/2}$, with n_0 being the density of the valence electrons. Since $(V_{\vec{G}_{111}}/\hbar\omega_p^0)\approx0.2$ for most semiconductors (see Table I), a perturbation expansion to second order in $U_{\vec{G}}/\hbar\omega$ is expected to work well for $\omega\sim\omega_p^0$. The relation between the coefficients in Table I and $U_{\vec{G}}$ is given by[14]

$$U_{\vec{G}}=V_G^S\cos(\vec{G}\cdot\vec{\tau})+iV_G^A\sin(\vec{G}\cdot\vec{\tau})\ , \tag{2.5}$$

with $\vec{\tau}=\frac{1}{8}a(1,1,1)$ and

$$V_G^S=\tfrac{1}{2}[V_1(G)+V_2(G)]\ ,\qquad V_G^A=\tfrac{1}{2}[V_1(G)-V_2(G)]\ , \tag{2.6}$$

where the indices 1 and 2 stand for pseudopotentials due to the two types of atoms in the III-V compounds.

III. THE LONG-WAVELENGTH LIMIT

A brief account of some of the results presented in this section was previously reported in a Letter by one of the authors.[11] Here we present a detailed derivation and discussion.

For $\vec{k}$ well within the first Brillouin zone, we said in Sec. II that the individual terms of the dielectric matrix that constitute $\epsilon_M(\vec{k},\omega)$, as in

TABLE I. Empirical pseudopotential form factors derived from experimental (Refs. 14 and 15) energy-band splittings ($V_3^S = V_{G_{111}}^S$, $V_8^S = V_{G_{220}}^S$, $V_{11}^S = V_{G_{311}}^S$, etc). a is the lattice constant; $|V_3^S/\hbar\omega_p^0|$ demonstrates the smallness of the pseudopotential compared to the plasmon energy.

	V_3^S (Ry)	V_8^S	V_{11}^S	V_3^A	V_4^A	V_{11}^A	a (Å)	$\lvert V_3^S/\hbar\omega_p^0\rvert$
Diamond	−0.811	0.337	0.132	0	0	0	3.57	0.35
Si	−0.21	0.04	0.08	0	0	0	5.43	0.17
Ge	−0.23	0.01	0.06	0	0	0	5.66	0.20
Sn	−0.20	0.00	0.04	0	0	0	6.49	0.21
GaP	−0.22	0.03	0.07	0.12	0.07	0.02	5.44	0.18
GaAs	−0.23	0.01	0.06	0.07	0.05	0.01	5.64	0.20
InP	−0.23	0.01	0.06	0.07	0.05	0.01	5.86	0.21
InAs	−0.22	0.00	0.05	0.08	0.05	0.03	6.04	0.21
GaSb	−0.22	0.00	0.05	0.06	0.05	0.01	6.12	0.22
AlSb	−0.21	0.02	0.06	0.06	0.04	0.02	6.13	0.21
InSb	−0.20	0.00	0.04	0.06	0.05	0.01	6.48	0.21

(2.3), may be calculated analytically by expanding the single-particle Bloch functions and corresponding energies to second order in the effective crystal potential. A different derivation, however, may be used to calculate the long-wavelength limit which results in a particularly simple form of $\epsilon_M(0,\omega)$. This limit is of particular interest since the optical properties as well as the loss function for $\vec{k}\to\vec{0}$ derive from $\epsilon_M(0,\omega)$.

We use the identity

$$(E_\sigma - E_{\sigma'})\langle\sigma|e^{-i\vec{k}\cdot\vec{r}}|\sigma'\rangle = \left\langle\sigma\left|e^{-i\vec{k}\cdot\vec{r}}\left(\frac{\hbar^2k^2}{2m}+\frac{\hbar\vec{k}\cdot\vec{p}}{m}\right)\right|\sigma'\right\rangle \tag{3.1}$$

and a similar relation for $(E_\sigma - E_{\sigma'})\langle\sigma'|e^{i\vec{k}\cdot\vec{r}}|\sigma\rangle$, that, in the limit as $\vec{k}\to\vec{0}$, it follows from (1.3)

$$\mathrm{Im}\epsilon(0,0,\omega) = \lim_{\vec{k}\to\vec{0}} \pi\frac{4\pi e^2}{k^2}\frac{\hbar^2}{m^2}\frac{1}{(\hbar\omega)^2}\sum_{\sigma,\sigma'}[n(E_\sigma)-n(E_{\sigma'})]\,|\langle\sigma|\vec{k}\cdot\vec{p}|\sigma'\rangle|^2\delta(\hbar\omega+E_\sigma-E_{\sigma'})\,. \tag{3.2}$$

Since

$$\langle\sigma|\vec{k}\cdot\vec{p}|\sigma'\rangle = \langle\sigma|\vec{k}\cdot\vec{p}|\sigma\rangle\delta_{\sigma,\sigma'} + \frac{(1-\delta_{\sigma,\sigma'})}{E_\sigma - E_{\sigma'}}\langle\sigma|[\mathcal{H},\vec{k}\cdot\vec{p}]|\sigma'\rangle \tag{3.3}$$

and

$$[\mathcal{H},\vec{k}\cdot\vec{p}] = [V(\vec{r}),\vec{k}\cdot\vec{p}] = \sum_{\vec{G}}\hbar(\vec{k}\cdot\vec{G})U_{\vec{G}}e^{-i\vec{G}\cdot\vec{r}}\,, \tag{3.4}$$

we obtain

$$\mathrm{Im}\epsilon(0,0,\omega) = \lim_{\vec{k}\to\vec{0}}\sum_{\vec{G},\vec{G}'}(\vec{k}\cdot\vec{G})(\vec{k}\cdot\vec{G}')U_G U_{G'}^*\frac{\hbar^4}{m^2}\frac{4\pi^2e^2}{k^2}\frac{1}{(\hbar\omega)^4} \times\sum_{\sigma,\sigma'}[n(E_\sigma)-n(E_{\sigma'})]\langle\sigma|e^{-i\vec{G}\cdot\vec{r}}|\sigma'\rangle\langle\sigma'|e^{i\vec{G}'\cdot\vec{r}}|\sigma\rangle\,\delta(\hbar\omega+E_\sigma-E_{\sigma'})\,, \tag{3.5}$$

where the contribution of the first term of (3.3) has been omitted[16] since it is only nonzero for $\hbar\omega\to 0$.

As we are interested in $\mathrm{Im}\epsilon(0,0,\omega)$ to second order in the pseudopotential, the last double summation in (3.5) is evaluated for free electrons. It then follows that only terms with $\vec{G}'=\vec{G}$ contribute, and using the definition of the Lindhard dielectric function $\epsilon_L(G,\omega)$ the final result may be written as

$$\mathrm{Im}\epsilon(0,0,\omega) = \lim_{\vec{k}\to\vec{0}}\left(4\sum_{\vec{G}}\frac{(\vec{k}\cdot\vec{G})^2}{k^2G^2}\frac{|U_{\vec{G}}|^2E_G^2}{(\hbar\omega)^4}\mathrm{Im}\epsilon_L(G,\omega)\right)\,, \tag{3.6}$$

with $E_G = \hbar^2G^2/2m$. Since this formula is only valid for semiconductors for $\hbar\omega \gg E_{\mathrm{gap}}$, we do not want to rely on the Kramers-Kronig relation for the derivation of $\mathrm{Re}\epsilon(0,0,\omega)$ but prefer to calculate it directly. From (1.3) one has

$$\mathrm{Re}\epsilon(\vec{k},\vec{k},\omega)=1-\frac{4\pi e^2}{k^2}\sum_{\sigma,\sigma'}\frac{n(E_\sigma)-n(E_{\sigma'})}{\hbar\omega+E_\sigma-E_{\sigma'}}\langle\sigma|e^{-i\vec{k}\cdot\vec{r}}|\sigma'\rangle\langle\sigma'|e^{i\vec{k}\cdot\vec{r}}|\sigma\rangle\,. \tag{3.7}$$

The above expression can quite generally be rewritten by using the identity

$$\frac{1}{x+a}=\frac{1}{x}-\frac{a}{x^2}+\frac{a^2}{x^2(x+a)} \tag{3.8}$$

and the f-sum rule[4]

$$\sum_{\sigma,\sigma'}n(E_\sigma)(E_{\sigma'}-E_\sigma)(\langle\sigma|e^{-i\vec{q}\cdot\vec{r}}|\sigma'\rangle\langle\sigma'|e^{i\vec{q}\cdot\vec{r}}|\sigma\rangle+\langle\sigma|e^{i\vec{q}\cdot\vec{r}}|\sigma'\rangle\langle\sigma'|e^{-i\vec{q}\cdot\vec{r}}|\sigma\rangle)=\frac{\hbar^2q^2}{m}n_0 \tag{3.9}$$

which is obtained from the double commutator

$$[e^{-i\vec{q}\cdot\vec{r}},[\mathcal{H},e^{i\vec{q}\cdot\vec{r}}]]=\hbar^2q^2/m\,.$$

We find that

$$\mathrm{Re}\epsilon(\vec{k},\vec{k},\omega)=1-\frac{\omega_p^2}{\omega^2}-\frac{4\pi e^2}{k^2}\sum_{\sigma,\sigma'}\frac{n(E_\sigma)-n(E_{\sigma'})}{\hbar\omega+E_\sigma-E_{\sigma'}}\frac{(E_\sigma-E_{\sigma'})^2}{(\hbar\omega)^2}\,|\langle\sigma|e^{-i\vec{k}\cdot\vec{r}}|\sigma'\rangle|^2\,. \tag{3.10}$$

It is obvious that in the limit $\vec{k}\to\vec{0}$ the last term in Eq. (3.10) represents band-structure contributions. If these are sufficiently weak, then the plasmon energy in the long-wavelength limit is approximately given by the homogeneous electron-gas result.

The evaluation of the limit $\vec{k}\to\vec{0}$ proceeds as for $\mathrm{Im}\epsilon(0,0,\omega)$ and yields

$$\mathrm{Re}\epsilon(0,0,\omega)=\lim_{\vec{k}\to\vec{0}}\Bigg(1-\frac{\omega_p^2}{\omega^2}-\sum_{\vec{G},\vec{G}'}(\vec{k}\cdot\vec{G})(\vec{k}\cdot\vec{G}')U_{\vec{G}}U_{\vec{G}'}^*\,\frac{\hbar^4}{m^2}\frac{1}{(\hbar\omega)^2}$$
$$\times\frac{4\pi e^2}{k^2}\sum_{\sigma,\sigma'}\frac{n(E_\sigma)-n(E_{\sigma'})}{\hbar\omega+E_\sigma-E_{\sigma'}}\frac{1}{(E_\sigma-E_{\sigma'})^2}\langle\sigma|e^{-i\vec{G}\cdot\vec{r}}|\sigma'\rangle\langle\sigma'|e^{i\vec{G}'\cdot\vec{r}}|\sigma\rangle)\,. \tag{3.11}$$

If we now use the fact that

$$\frac{1}{(E_\sigma-E_{\sigma'})^2}\,\frac{1}{\hbar\omega+E_\sigma-E_{\sigma'}}=\frac{1}{\hbar\omega(E_\sigma-E_{\sigma'})^2}-\frac{1}{(\hbar\omega)^2(E_\sigma-E_{\sigma'})}+\frac{1}{(\hbar\omega)^2(\hbar\omega+E_\sigma-E_{\sigma'})} \tag{3.12}$$

and proceed as for $\mathrm{Im}\epsilon(0,0,\omega)$, the final result can be cast in the form

$$\mathrm{Re}\epsilon(0,0,\omega)=\lim_{\vec{k}\to\vec{0}}\left(1-\frac{\omega_p^2}{\omega^2}+4\sum_{\vec{G}}\frac{(\vec{k}\cdot\vec{G})^2}{k^2G^2}\,\frac{|U_{\vec{G}}|^2E_G^2}{(\hbar\omega)^4}\,\mathrm{Re}[\epsilon_L(G,\omega)-\epsilon_L(G,0)]\right)\,. \tag{3.13}$$

Note that the first term of (3.12) does not contribute to $\mathrm{Re}\epsilon(0,0,\omega)$ since it is odd in ω. Equations (3.6) and (3.13) may thus be represented together by

$$\epsilon(0,0,\omega)=\lim_{\vec{k}\to\vec{0}}\left(1-\frac{\omega_p^2}{\omega^2}+4\sum_{\vec{G}}\frac{(\vec{k}\cdot\vec{G})^2}{k^2G^2}\,\frac{|U_{\vec{G}}|^2E_G^2}{(\hbar\omega)^4}\,[\epsilon_L(G,\omega)-\epsilon_L(G,0)]\right)\,. \tag{3.14}$$

If local-field effects are neglected, then, as in the first term in Eq. (2.13),

$$\epsilon_M(0,\omega)=\epsilon(0,0,\omega)\,. \tag{3.15}$$

The calculation of the off-diagonal elements is easily carried out in a similar manner, leading to

$$\epsilon(0,\vec{G},\omega)=\lim_{\vec{k}\to\vec{0}}\left(-2\,\frac{(\vec{k}\cdot\vec{G})}{k^2}\,\frac{U_{\vec{G}}^*E_G}{(\hbar\omega)^2}\,[\epsilon_L(G,\omega)-\epsilon_L(G,0)]\right) \tag{3.16}$$

and

$$\epsilon(\vec{G},0,\omega)=\lim_{\vec{k}\to\vec{0}}\left(-2\,\frac{(\vec{k}\cdot\vec{G})}{G^2}\,\frac{U_{\vec{G}}^*E_G}{(\hbar\omega)^2}\,[\epsilon_L(G,\omega)-\epsilon_L(G,0)]\right). \tag{3.17}$$

Local-field effects are now included by inserting (3.14), (3.16), and (3.17) into (2.3), which leads to

$$\epsilon_M(0,\omega)=\lim_{\vec{k}\to 0}\left[1-\frac{\omega_p^2}{\omega^2}+4\sum_{\vec{G}}{}'\frac{(\vec{k}\cdot\vec{G})^2}{k^2G^2}\,\frac{|U_{\vec{G}}|^2E_G^2}{(\hbar\omega)^4}\,\epsilon_L^2(G,0)\left(\frac{1}{\epsilon_L(G,0)}-\frac{1}{\epsilon_L(G,\omega)}\right)\right]. \tag{3.18}$$

We can see the difference between (3.14) and (3.18), for instance, by considering the optical absorption. Without local-field effects we have

$$\mathrm{Im}\epsilon\,(0,0,\omega)=4\sum_{\vec{G}}\frac{(\vec{k}\cdot\vec{G})^2}{k^2G^2}\,\frac{|U_{\vec{G}}|^2E_G^2}{(\hbar\omega)^4}\,\mathrm{Im}\epsilon_L(G,\omega)\,, \tag{3.6}$$

while if local-field effects are included,

$$\mathrm{Im}\epsilon_M(0,\omega)=4\sum_{\vec{G}}\frac{(\vec{k}\cdot\vec{G})^2}{k^2G^2}\,\frac{|U_{\vec{G}}|^2E_G^2}{(\hbar\omega)^4}\times\left|\frac{\epsilon_L(G,0)}{\epsilon_L(G,\omega)}\right|^2\mathrm{Im}\epsilon_L(G,\omega)\,. \tag{3.19}$$

The physical basis of the factor $|\epsilon_L(G,0)/\epsilon_L(G,\omega)|^2$ can be viewed as follows[17]: $U_{\vec{G}}$, which represents the effective crystal potential in the single-particle Hamiltonian, is a statically screened potential, i.e.,

$$U_{\vec{G}}=\tilde{U}_{\vec{G}}/\epsilon_L(G,0)\,, \tag{3.20}$$

where $\tilde{U}_{\vec{G}}$ is the bare potential. Then

$$|U_{\vec{G}}|^2\left|\frac{\epsilon_L(G,0)}{\epsilon_L(G,\omega)}\right|^2=\left|\frac{\tilde{U}_{\vec{G}}}{\epsilon_L(G,\omega)}\right|^2\,,$$

which means that the inclusion of local-field effects results in a "dynamical screening"[18] of the crystal potential.

The natural question which now arises is how many terms have to be included in the $\vec{G}$ summation. This question is easy to answer in the case of the optical absorption, i.e., $\mathrm{Im}\epsilon_M(0,\omega)$, because for a given $\hbar\omega$ only those terms for which the point (G,ω) falls in the particle-hole excitation spectrum of the free-electron gas contribute, due to the fact that only for these points is $\mathrm{Im}\epsilon_L(G,\omega)$ nonzero. Obviously the strength of the pseudopotential is reflected in the strength of the optical absorption. From Fig. 1 it is seen that with increasing ω an increasing number of sets of reciprocal-lattice vectors, i.e., Fourier coefficients $U_{\vec{G}}$, have to be taken into account. Actually one can predict the onset of the absorption associated with a particular $U_{\vec{G}}$ from

$$\hbar\omega(G)=E_F\left[(G/k_F)^2-2(G/k_F)\right]\,. \tag{3.21}$$

Here E_F and k_F are free-electron parameters defined in the usual way as functions of the electron density n_0. For the covalent semiconductors only the lowest three symmetric and antisymmetric form factors have been determined from optical data at low frequencies ($\omega\ll\omega_p^0$) within the empirical pseudopotential method.[19] So far, nothing is known of the higher coefficients and they are usually set equal to zero. According to (3.21), one finds for $\vec{G}_4=(2\pi/a)(4,0,0)$, $\hbar\omega(G_4)=17.8$ eV.

Recent experimental data[20] on $\mathrm{Im}\epsilon_M(0,\omega)$ obtained from ELS data by a Kramers-Kronig analysis are shown in Fig. 2 together with theoretical curves as calculated[21] from Eqs. (3.6) and (3.19) with U_{G_4} set equal to zero, as is usually assumed. We then observe very good agreement between theory and experiment only below about 19 eV if local-field effects are included. We believe this is a clear indication that $V_{G_{400}}\neq 0$. Figure 3 shows the difference between the theoretical curve and the experimental curve in Fig. 2 and the form of the additional contribution expected from $|V_{400}|=1$ Ry. Scaling the theoretical curve down to the experimental curve would in principle allow a determination of V_{400}, but since the loss experiment was only carried out up to $\hbar\omega=25$ eV we do not know how reliable the $\mathrm{Im}\epsilon_{\mathrm{expt}}(0,\omega)$ is. A search for more experimental data revealed that

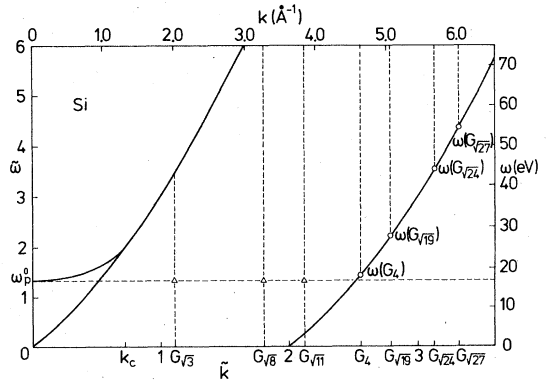

FIG. 1. Electron gas excitation spectrum with electron density appropriate to Si. Only the points (G,ω) within the particle-hole spectrum contribute to $\mathrm{Im}\epsilon_M(0,\omega)$ (for example, at $\omega=\omega_p^0$ only the first three sets of reciprocal-lattice vectors are important). $\omega(G)$ correspond to the onset of interband transitions associated with a certain $U_{\vec{G}}$. $\tilde{\omega}$ and $\tilde{k}$ are given in units of E_F and k_F, respectively.

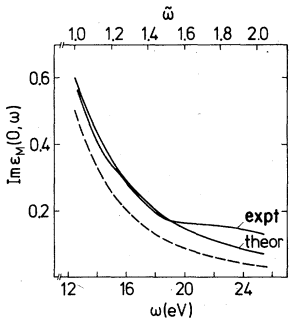

FIG. 2. Calculated $\mathrm{Im}\epsilon_M(0,\omega)$ for Si with pseudopotential from Ref. 21 compared with experiment from Ref. 20. Dashed curve is without local-field corrections [Eq. (3.6)]. $\tilde{\omega}$ is given in units of the Fermi energy E_F.

a fairly strong onset of additional interband absorption[22] occurs in Ge at about 25.5 eV in close agreement with $\hbar\omega(G_{19}) = 25.3$ eV. Figure 2 also shows that the importance of local-field effects increases with increasing frequency and seems to become the dominant absorption mechanism at high ω. To demonstrate this and the qualitatively different behavior of simple metals we have plotted in Fig. 4 the ratio

$$R(\omega) = \mathrm{Im}\epsilon_M(0,\omega)/\mathrm{Im}\epsilon(0,0,\omega) \qquad (3.22)$$

for Si and Al. If only the dominant pseudopotential coefficient is considered, which for Si is $U_{G_{111}}$ and for Al is $U_{G_{200}}$, one has

$$R(\omega) \approx |\epsilon_L(G_0,0)/\epsilon_L(G_0,\omega)|^2, \qquad (3.23)$$

which is also shown as dashed curves in Fig. 4, G_0 being G_{111} or G_{200} for Si and Al, respectively.

The energy-loss function $\mathrm{Im}[-1/\epsilon_M(0,\omega)]$ depends on both $\mathrm{Re}\epsilon_M(0,\omega)$ and $\mathrm{Im}\epsilon_M(0,\omega)$. In this case the extent of the $\vec{G}$ summation in (3.18) is less clear, but in practice we have to limit ourselves to those pseudopotential coefficients which are known. The dominant structure in the energy-loss spectrum originates from the plasmon. As already noted, the peak position is fairly well given by ω_p^0, the homogeneous electron-gas value of the plasmon frequency, whereas the height and linewidth of the loss function are mainly determined by $\mathrm{Im}\epsilon_M(0,\omega)$. Up to ω_p^0, $\mathrm{Im}\epsilon_M(0,\omega)$ is determined by the known pseudopotential coefficients. This explains why the overall agreement between the experimentally[2,20,23-25] determined loss function and the theoretical curve—as shown in Figs. 5 and 6—is very good in most cases,

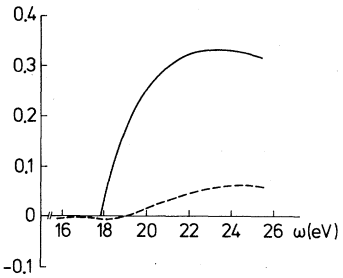

FIG. 3. Solid curve: contribution to $\mathrm{Im}\epsilon_M(0,\omega)$ starting at $\omega(G_{400})$ when $|V_{400}| = 1$ Ry is assumed; dashed curve: difference between experimental and theoretical $\mathrm{Im}\epsilon_M(0,\omega)$ when $\mathrm{Im}\epsilon_M(0,\omega)$ was calculated with $V_{400} = 0$, as in Fig. 2.

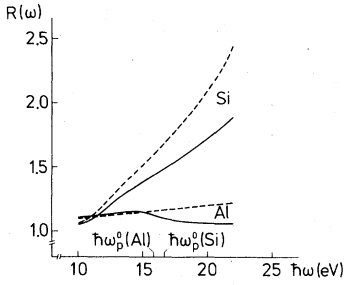

FIG. 4. $R(\omega)$ calculated from Eqs. (3.22) (solid curves) and (3.23) (dashed curves) with pseudopotential from Ref. 21 for Si and Ref. 19 for Al.

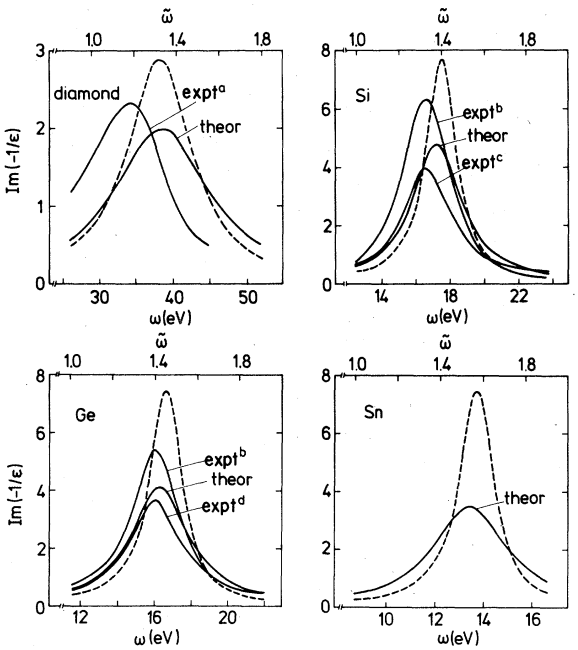

FIG. 5. Energy-loss spectra for semiconductors of the diamond structure for $\vec{k} = \vec{0}$. Theoretical curves with pseudopotential from Refs. 14 and 15. Experiments: (a) Ref. 2, (b) Ref. 23, (c) Ref. 20, (d) Ref. 25. Dashed curves: theory without local-field effects. $\tilde{\omega}$ is given in units of the Fermi energy E_F.

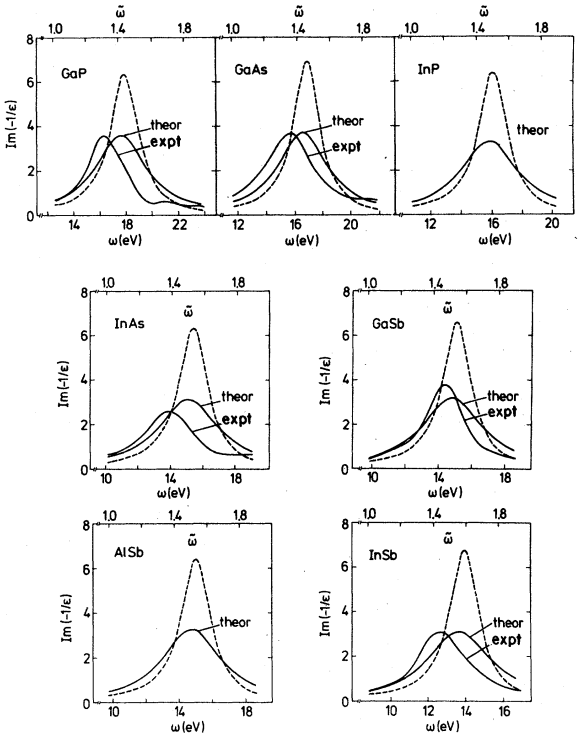

FIG. 6. Energy-loss spectra for semiconductors of the zinc-blende structure for $\vec{k}=\vec{0}$. Theoretical curves with pseudopotential from Ref. 14. Experiment from Ref. 24. Dashed curves: theory without local-field effects. $\tilde{\omega}$ is given in units of the Fermi energy E_F.

provided local-field effects have been included, of course. However, in almost all cases the position of the theoretical plasmon peak is slightly higher than the position of the experimental peak. For the semiconductors of the diamond structure (Fig. 5) we believe that inclusion of the higher $V_{\vec{G}}$, in particular $V_{G_{400}}$ and $V_{G_{331}}$, would considerably improve the agreement with experiment, as one can easily show that the effect of the pseudopotential form factors is to shift the theoretical plasmon peak to lower energies. The reason is that any excitation above the plasmon frequency shifts the plasmon frequency down and vice versa. However, for the III-V semiconductors of the zinc-blende structure, interband transitions from d-core states can be observed in the loss spectrum near 20 eV. Since dynamical effects of core electrons are excluded in the pseudopotential model their effect cannot be accounted for by the present theory, but as discussed by Philipp and Ehrenreich,[26] they also shift the plasma frequency downward in energy.

In Table II we compare the experimental results for the plasmon linewidth $\Delta_{1/2}$ and for the height of the loss function $\mathrm{Im}(-\epsilon^{-1})$ in the long-wavelength limit with our theoretical results. It is apparent from Table II that inclusion of local-field effects to second order in the effective crystal potential leads to a substantial increase in the long-wavelength plasmon linewidth, in good agreement with experiment in most cases. Also, the results we obtained for the height of the loss peak are quite satisfactory, given the present stage of experimental knowledge. Table II also displays the results for the long-wavelength plasmon linewidth

$$\Delta^{P}_{1/2} \approx \hbar\omega_p \epsilon_2^{P}(\omega_p) \tag{3.24}$$

and for the height of the loss peak

$$\mathrm{Im}(-\epsilon^{-1})^{P}_{\max} \approx [\epsilon_2^{P}(\omega_p)]^{-1} \tag{3.25}$$

TABLE II. Experimental and theoretical results for the plasmon linewidth $\Delta_{1/2}$ (in eV) and for the height of the loss function $\mathrm{Im}(-\epsilon^{-1})$ in the long-wavelength limit. LF—this work, including local-field effects, NLF—this work, neglecting local-field effects, P—phenomenological Penn model (Ref. 27).

Crystal	$\Delta^{\mathrm{expt}}_{1/2}$	$\Delta^{\mathrm{LF}}_{1/2}$	$\Delta^{\mathrm{P}}_{1/2}$	$\Delta^{\mathrm{NLF}}_{1/2}$	$\mathrm{Im}(-\epsilon^{-1})^{\mathrm{expt}}_{\max}$	$\mathrm{Im}(-\epsilon^{-1})^{\mathrm{LF}}_{\max}$	$\mathrm{Im}(-\epsilon^{-1})^{\mathrm{P}}_{\max}$	$\mathrm{Im}(-\epsilon^{-1})^{\mathrm{NLF}}_{\max}$
Diamond	13.6[a]	15.7	20.7	10.5	2.3[a]	2.0	1.5	2.9
Si	3.6,[b] 3.8[c]	3.5	4.5	2.2	6.3,[b] 3.9[c]	4.8	3.6	7.6
Ge	3.4,[b] 3.6[d]	3.9	4.9	2.1	5.4,[b] 3.6[d]	4.1	3.2	7.4
GaP	3.5[e]	4.7	7.4	2.7	3.5[e]	3.6	2.2	6.3
GaAs	4.1[e]	4.3	7.1	2.3	3.6[e]	3.7	2.2	6.8
InAs	3.7[e]	4.6	7.1	2.2	2.7[e]	3.1	2.0	6.3
GaSb	2.8[e]	4.4	5.2	2.1	3.7[e]	3.2	2.7	6.6
InSb	3.1[e]	4.2	5.1	1.9	3.1[e]	3.1	2.5	6.7

[a] Reference 2.
[b] Reference 23.
[c] Reference 20.
[d] Reference 25.
[e] Reference 24.

obtained by using a phenomenological two-band Penn model according to which,[27] for $\vec{k}=\vec{0}$,

$$\epsilon_2^P(\omega)=\pi D(\hbar\omega_p)^2\frac{E_g^2}{(\hbar\omega)^3[(\hbar\omega)^2-E_g^2]^{1/2}}\ . \quad (3.26)$$

Here, E_g is an average energy gap between the valence and conduction states, D is a parameter which describes core d-electron effects, and both E_g and D are fitted to the experimentally determined dielectric constant $\epsilon_1(\vec{k}=\vec{0},\ w=0)=\epsilon_0$. Formulas (3.24) and (3.25) are only valid for sufficiently sharp lines; for broader lines, the linewidth, for example, is overestimated. This arises from the fact that $\mathrm{Im}\epsilon_M(\vec{0},\omega)$ is not constant over the width of the peak, which is the underlying assumption in the derivation of (3.24) and (3.25). Although the phenomenological Penn model has been highly successful in describing chemical trends, it is clear from Table II that we should go beyond the spherical model to obtain agreement with experiment. Moreover, this spherical model cannot explain the anisotropy of the plasmon dispersion which is studied in the next section using an analytic extension of the present pseudopotential scheme to finite $\vec{k}$.

IV. THE ANISOTROPY OF THE PLASMON DISPERSION

It is well known from any band-structure calculation that the band energies not only depend on the modulus of $\vec{k}$ but also on the direction of $\vec{k}$. ELS experiments on single crystals are therefore expected to show that $\omega_p(\vec{k})$ depends on the direction of $\vec{k}$ with respect to the symmetry axes of the crystal. Recent ELS measurements in Si, GaAs, and InSb confirm this. In order to explain this behavior we extend the theory presented in Sec. III to finite $\vec{k}$ values.[6] Again, the matrix elements in (1.3) may be readily evaluated by using second-order perturbation theory in the pseudopotential. In what follows it is convenient to work with energies and wave vectors in units of the free-electron parameters: the Fermi energy E_F and the Fermi wave vector k_F, respectively, which are defined in the usual way as functions of the (valence) electron density n_0.

If we decompose the elements of the dielectric matrix into an absorptive part $A_{\vec{G},\vec{G}'}(\vec{k},\omega)$ and a dispersive part $H_{\vec{G},\vec{G}'}(\vec{k},\omega)$

$$\epsilon_{\vec{G},\vec{G}'}(\vec{k},\omega)=H_{\vec{G},\vec{G}'}(\vec{k},\omega)+iA_{\vec{G},\vec{G}'}(\vec{k},\omega)\ , \quad (4.1)$$

we obtain for points $(\vec{k},\omega)$ outside the free-electron particle-hole excitation spectrum when $\omega>2k+k^2$,

$$A_{\vec{0},\vec{0}}(\vec{k},\omega)=\frac{3\omega_p^2}{8k^2}\sum_{\vec{G}\neq\vec{0}}|U_{\vec{G}}|^2\int d^3p\ n(p)\left(\frac{1}{p^2-(\vec{p}+\vec{G})^2}-\frac{1}{(\vec{p}+\vec{k})^2-(\vec{p}+\vec{k}+\vec{G})^2}\right)^2$$
$$\times[\delta(\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2)-\delta(\omega+(\vec{p}+\vec{k}+\vec{G})^2-p^2)] \quad (4.2)$$

and

$$H_{\vec{0},\vec{0}}(\vec{k},\omega)=\epsilon_L(k,\omega)+[I(\vec{k},\omega)+I(\vec{k},-\omega)]. \quad (4.3)$$

The free-electron contribution $\epsilon_L(k,\omega)$ has been split off in (4.3), so the $I(\vec{k},\omega)$ and $I(\vec{k},-\omega)$ contain only the contribution that depends on the crystal potential and we have

$$I(\vec{k},\omega)=-\frac{3\omega_p^2}{8\pi k^2}\sum_{\vec{G}\neq\vec{0}}|U_{\vec{G}}|^2F(\vec{G};\vec{k},\omega)\ , \quad (4.4)$$

with

$$F(\vec{G};\vec{k},\omega)=\int d^3p\Bigg\{n(p)\Bigg[\frac{1}{[\omega+p^2-(\vec{p}+\vec{k})^2]^2}\left(\frac{1}{\omega+(\vec{p}+\vec{G})^2-(\vec{p}+\vec{k})^2}+\frac{1}{\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2}\right)$$
$$+\frac{2}{[\omega+p^2-(\vec{p}+\vec{k})^2][p^2-(\vec{p}+\vec{G})^2][\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2]}$$
$$-\frac{1}{[p^2-(\vec{p}+\vec{G})^2]^2}\left(\frac{1}{\omega+(\vec{p}+\vec{G})^2-(\vec{p}+\vec{k})^2}-\frac{1}{\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2}\right)\Bigg]$$
$$-\tfrac{1}{2}\delta(1-p)\Bigg[\frac{1}{[\omega+p^2-(\vec{p}+\vec{k})^2][p^2-(\vec{p}+\vec{G})^2]}+\left(\frac{1}{4G}\right)\frac{1}{\omega+p^2-(\vec{p}+\vec{k})^2}\ln\left|\frac{G+2}{G-2}\right|\Bigg]\Bigg\}. \quad (4.5)$$

In calculating $\epsilon_M(\vec{k},\omega)$ to second order in the crystal potential it is sufficient to evaluate the off-diagonal elements to first order and we obtain

$$A_{\vec{0},\vec{G}}(\vec{k},\omega)=\frac{3\omega_p^{0\,2}}{8k^2}U_{\vec{G}}^{*}\int d^3p\; n(p)\left(\frac{1}{p^2-(\vec{p}+\vec{G})^2}-\frac{1}{(\vec{p}+\vec{k})^2-(\vec{p}+\vec{k}+\vec{G})^2}\right)$$
$$\times[\delta(\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2)-\delta(\omega+(\vec{p}+\vec{k}+\vec{G})^2-p^2)] \tag{4.6}$$

and

$$H_{\vec{0},\vec{G}}(\vec{k},\omega)=-\frac{3\omega_p^{0\,2}}{8\pi k^2}U_{\vec{G}}^{*}[J(\vec{k},\omega)+J(\vec{k},-\omega)], \tag{4.7}$$

with

$$J(\vec{k},\omega)=\int d^3p\; n(p)\left(\frac{1}{[\omega+p^2-(\vec{p}+\vec{k})^2][\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2]}\right.$$
$$\left.+\frac{1}{[p^2-(\vec{p}+\vec{G})^2][\omega+p^2-(\vec{p}-\vec{k})^2]}+\frac{1}{[p^2-(\vec{p}+\vec{G})^2][\omega+p^2-(\vec{p}+\vec{k}+\vec{G})^2]}\right) \tag{4.8}$$

and similar expressions for $A_{\vec{G},\vec{0}}(\vec{k},\omega)$ and $H_{\vec{G},\vec{0}}(\vec{k},\omega)$. The above integrals may be evaluated *analytically* as shown in detail in the Appendix. The integrations are conveniently carried out by introducing cylindrical coordinates $(p_\parallel, p_\perp, \phi)$. With an appropriate choice of the polar axis, the ϕ integration is performed first, followed by the $p_\perp$ and finally $p_\parallel$ integration. All the integrals required can be found in standard integral tables.

Although the final result is very lengthly, its analytic form allows a very quick calculation of the loss function $\mathrm{Im}[-1/\epsilon_M(\vec{k},\omega)]$ as a function of $\vec{k}$. It is a simple exercise to work out the limit $\vec{k}\to\vec{0}$ and check that we recover the results of Sec. III.

In Fig. 7 the calculated loss function of Si is shown for three values of $\vec{k}$ along the [111] direction together with $\mathrm{Re}\epsilon_M(\vec{k},\omega)$ and $\mathrm{Im}\epsilon_M(\vec{k},\omega)$. The dashed line gives the position of the maximum of the loss function, which even for the broad peaks is close to the zeros of $\mathrm{Re}\epsilon_M(\vec{k},\omega)$. $\mathrm{Im}\epsilon_M(\vec{k},\omega)$ remains almost constant along the dashed curve, explaining the small change in width of the loss peak for the different $\vec{k}$ values. Thus the dispersion of the loss peak reflects fairly accurately the dispersion of the plasmon in these materials. In Fig. 8 recent experimental dispersion curves of the loss peak $\Delta E=\max[\mathrm{Im}(-1/\epsilon_M)]$ for single crystals of Si,[9] GaAs, and InSb (Ref. 10) along the three principal symmetry axes are compared with their theoretical counterparts. Owing to the similarities of the crystal potential and hence the band structure in these materials the theory predicts in all three cases that the anisotropy should be of the form

$$\omega_p^{[100]}(\vec{k})\geq\omega_p^{[110]}(\vec{k})\geq\omega_p^{[111]}(\vec{k})\ . \tag{4.9}$$

This agrees with the experimental findings in Si and GaAs. In InSb, however, the experiment shows

$$\omega_p^{[110]}(\vec{k})\geq\omega_p^{[100]}(\vec{k})\geq\omega_p^{[111]}(\vec{k})\ . \tag{4.10}$$

At present we cannot offer any explanation for this discrepancy, apart from the fact that it reflects the limited knowledge of the pseudopotential, i.e., of the band structure at these high energies (see discussion in Sec. III).

V. CONCLUSIONS

We have shown that the application of the nearly-free-electron pseudopotential theory to calculating the dielectric properties of covalent semiconductors within the self-consistent-field approximation yields a quantitative description of the optical properties and the loss function in the long wavelength limit. It is demonstrated that local-field effects due to the inhomogeneous distribution of valence electrons in semiconductors (as manifest in the covalent-bond charges) must be included and that their influence increases at high frequencies.

The present simple theory suggests that the so-far unknown pseudopotential form factors for $|\vec{G}|$ values greater than $|\vec{G}_{311}|$ may be successively determined from either optical or loss experiments for $\vec{k}\to\vec{0}$. Of course such a procedure only makes sense for excitation energies well below the onset of core excitations. Good candidates for such an investigation are diamond, Si, and Ge. The new form factors could then be used (and tested) to evaluate the plasmon dispersion more precisely.

With the present limited knowledge of the pseudopotential form factors the anisotropy of the plas-

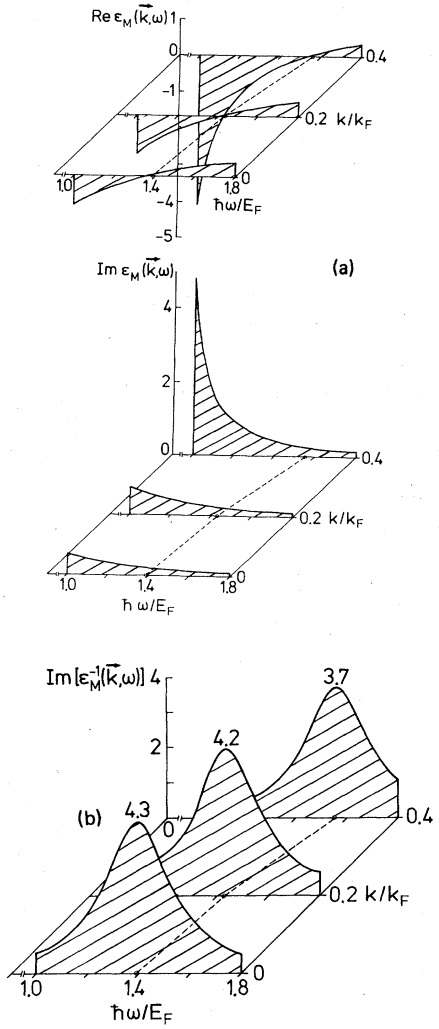

FIG. 7. Perspective plot of $\mathrm{Re}\epsilon_M(\vec{k},\omega)$, $\mathrm{Im}\epsilon_M(\vec{k},\omega)$, and $\mathrm{Im}[-1/\epsilon_M(\vec{k},\omega)]$ of Si (pseudopotential from Ref. 21) for three values of $\vec{k}$ along the [111] direction. The dashed line in the $(\vec{k},\omega)$ plane represents the dispersion of the maximum of the loss function.

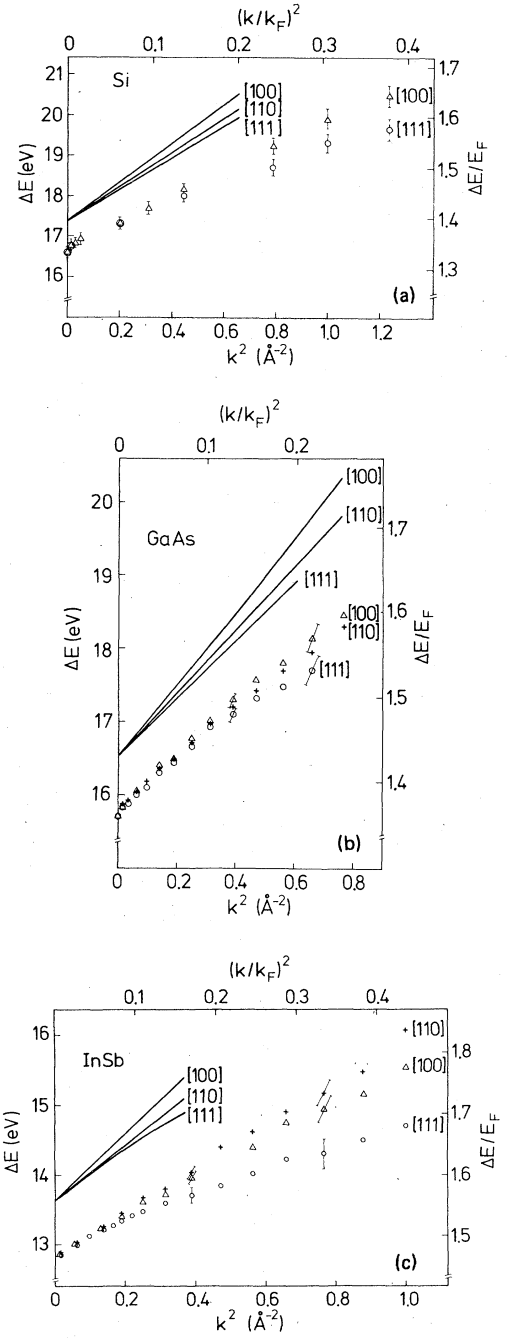

FIG. 8. Energy-loss dispersion along the [111], [110], and [100] directions. Si: experiment—Ref. 9, pseudopotential—Ref. 21; GaAs and InSb: experiment—Ref. 10, pseudopotential—Ref. 14.

mon dispersion can be satisfactorily understood for single crystals of Si and GaAs. InSb, which is believed to have a similar band structure, unexpectedly shows a different anisotropy for the [100] and [110] direction from Si and GaAs, whereas it is similar in the [111] direction which shows the least dispersion in all three cases.

The particular advantage of the present theory lies in the fact that all calculations can be performed analytically. The results are particularly simple and apparent in the long-wavelength limit when they can be expressed in terms of the Lindhard function. They offer an easy check for high frequencies on numerical *ab initio* computations based on a previous numerical band-structure calculation. Such a check with a previous numerical calculation of the loss function of Si (Refs. 8 and 11) shows the rather close agreement of the present results with the extensive numerical calculation, which in principle is more precise, pro-

vided a sufficiently large number of bands is taken into account. However fine, details of the band structure will be lost in the calculation of dielectric properties since the optical absorption, for example, is an average value over all electronic excitations of a given energy $\hbar\omega$. Only excitations corresponding to a large accessible volume in phase space will survive as a particular structure in the optical absorption. This might explain why a local pseudopotential,[19] as used here, and the use of pseudo-wave-functions instead of proper orthogonalized plane waves (OPW's), is sufficient,[28] in particular with respect to the present state of the experimental accuracy.

It should be noted that for semiconductors this theory is only applicable for energies large compared to the gap. For energies of the order of the gap, a tight-binding approach is more suitable as recent calculations of the optical properties of diamond and Si have shown,[29,30] in which, in particular, exchange and correlation effects were also considered, giving rise to excitonic effects. At high energies, i.e., $\omega \sim \omega_p^0$, we conclude from the very good agreement between our theory (in which exchange and correlations are neglected) and experiment for $\vec{k} \to \vec{0}$ that exchange and correlation effects should not influence the results appreciably.

For increasing $\vec{k}$ the discrepancies between theory and experiments become noticeable; for example, the calculated absolute value of the dispersion is too high compared with experiments. This quite generally seems to be a weakness of the SCF approximation, which is already known from the plasmon dispersion in simple metals.[31] Exchange and correlation effects tend to reduce the dispersion.[32] As demonstrated by Fig. 9, the strongest dispersion is obtained from the random-phase approximation (RPA) exemplified by the solution for the homogeneous electron gas with a density n_0 appropriate to Al which is not very different from the valence electron density n_0 of Si. The SCF gives a quadratic dispersion for small $\vec{k}$ which becomes steeper for large $\vec{k}$. For Al a quadratic dispersion is experimentally observed for small $\vec{k}$, but with a smaller dispersion constant, which becomes steeper, but less than in the SCF, for large $\vec{k}$. The semiconductors behave qualitatively different from metals; there the plasmon dispersion is less steep for large $\vec{k}$. An almost dispersionless plasmon is found in ionic crystals such as the alkali halides.[33,34] Nevertheless in all these cases the long-wavelength plasmon energy is fairly accurately determined by the homogeneous electron-gas value, and an NFE approximation lends itself as the next step of improvement, giving very good results for $\vec{k} \to \vec{0}$ and also explaining band-structure effects such as the anisotropy in several cases.[35]

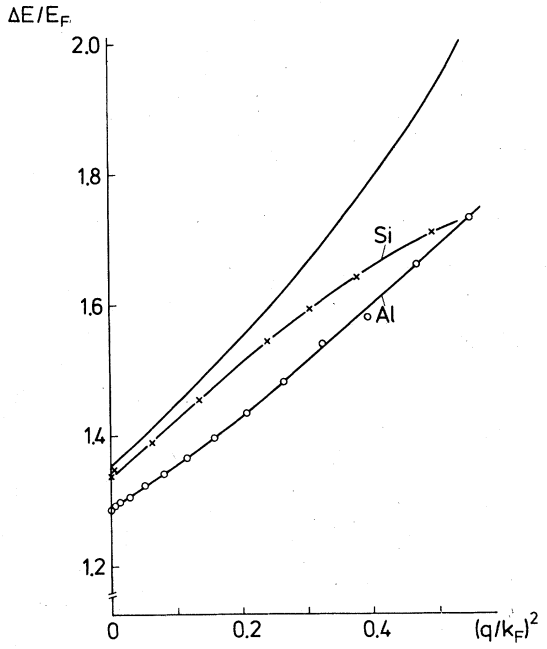

FIG. 9. Energy-loss dispersion for homogeneous electron gas (solid curve) with density appropriate for Al. Experiment for Si is from Ref. 9 and for Al from Ref. 37.

We have seen that the coupling to local fields provides an important decay mechanism for plasmons in semiconductors in addition to interband transitions. Another effect is that local fields should be responsible for the existence of plasmon bands in the sense that at a given $\vec{k}$, two (or more) plasmons may be experimentally observable in the loss spectrum, and in particular at the zone boundary, a gap between plasmon bands should open. First theoretical estimates of plasmon band gaps yield sizable values[36] (2 eV for Ge), indicating that this should be observable, but unfortunately lifetime effects were not considered. In the following paper[12] the present theory will be extended to a two-plasmon-band model to investigate the question as to whether a plasmon band structure can be observed in semiconductors.

ACKNOWLEDGMENTS

We are very much indebted to Dr. B. Buxton for many helpful comments and for a critical reading of the manuscript. One of the authors (L.E.O.) wishes to thank the Brazilian Research Council CNPq for financial support and is grateful for the hospitality extended to him by the Institut für Festkörperforschung der Kernforschungsanlage where part of this work was done.

APPENDIX

Here we indicate briefly how the integrals appearing in Eqs. (4.2)–(4.8) can be evaluated analytically.

1. Evaluation of $A_{\vec{0},\vec{0}}(\vec{k},\omega)$ and $A_{\vec{0},\vec{G}}(\vec{k},\omega)$

We first perform the momentum integration of Eq. (4.2) by introducing cylindrical coordinates $(p_\parallel, p_\perp, \phi)$ and choosing $\vec{q}=\vec{k}+\vec{G}$ as the polar axis. By using the properties of the δ function, Eq. (4.2) may be written as

$$A_{\vec{0},\vec{0}}(\vec{k},\omega)=\frac{3\omega_p^{0\,2}}{8k^2}\sum_{\vec{G}\neq\vec{0}}|U_{\vec{G}}|^2\int_{-1}^{1}dp_\parallel[L_1(p_\parallel)\delta(\omega-2p_\parallel q-q^2)-L_2(p_\parallel)\delta(\omega+2p_\parallel q+q^2)]$$

$$=\frac{3\omega_p^{0\,2}}{8k^2}\sum_{\vec{G}\neq\vec{0}}\frac{|U_{\vec{G}}|^2}{2q}\left[L_1\left(\frac{\omega-q^2}{2q}\right)\Theta\left(1-\left(\frac{\omega-q^2}{2q}\right)^2\right)-L_2\left(-\frac{\omega+q^2}{2q}\right)\Theta\left(1-\left(\frac{\omega+q^2}{2q}\right)^2\right)\right] \tag{A1}$$

with

$$\Theta(x)=\begin{cases}1, & x>0\\ 0, & x<0\end{cases}.$$

Hence,

$$L_1(x)=\int_0^{(1-x^2)^{1/2}}dp_\perp p_\perp\int_0^{2\pi}d\phi\left(\frac{1}{\alpha_1(x)-2k_\perp p_\perp\cos\phi}-\frac{1}{\alpha_2(x)-2k_\perp p_\perp\cos\phi}\right)^2, \tag{A2}$$

where $p_\parallel$, $p_\perp$ and $k_\parallel$, $k_\perp$ mean vector components parallel and perpendicular to $\vec{q}$, respectively, and

$$\alpha_1(x)=\omega-k^2-2k_\parallel x, \tag{A3}$$

$$\alpha_2(x)=\omega+k^2-2k_\parallel q-2k_\parallel x. \tag{A4}$$

A similar expression for $L_2(x)$ is obtained by changing ω to $-\omega$ in the definitions of α_1 and α_2. The integrals in Eq. (A2) are easily performed by carrying out first the ϕ integration[38] and then the $p_\perp$ integration; it follows then that

$$L_1(x)=\frac{\pi}{2k_\perp^2}\left(\frac{\psi_1^2}{(1-\psi_1^2)^{1/2}}+\frac{\psi_2^2}{(1-\psi_2^2)^{1/2}}-\frac{(\psi_1+\psi_2)^2}{(1-\psi_1^2)^{1/2}+(1-\psi_2^2)^{1/2}}\right), \tag{A5}$$

with

$$\psi_{1/2}(x)=2k_\perp(1-x^2)^{1/2}/\alpha_{1/2}(x). \tag{A6}$$

The integration of Eq. (4.6) is entirely analogous to the one just performed and yields

$$A_{\vec{0},\vec{G}}(\vec{k},\omega)=\frac{3\omega_p^{0\,2}}{8k^2}\frac{U_{\vec{G}}^*}{2q}\left[K_1\left(\frac{\omega-q^2}{2q}\right)\Theta\left(1-\left(\frac{\omega-q^2}{2q}\right)^2\right)-K_2\left(-\frac{\omega+q^2}{2q}\right)\Theta\left(1-\left(\frac{\omega+q^2}{2q}\right)^2\right)\right], \tag{A7}$$

where

$$K_1(x)=\frac{\pi}{2k_\perp^2}\{\alpha_1[1-(1-\psi_1^2)^{1/2}]-\alpha_2[1-(1-\psi_2^2)^{1/2}]\}, \tag{A8}$$

with the same definitions as before for $\alpha_{1/2}$, $\psi_{1/2}$, and a similar expression for $K_2(x)$ again obtained by simply changing ω to $-\omega$ in $\alpha_{1/2}$.

2. Evaluation of $H_{\vec{0},\vec{0}}(\vec{k},\omega)$ and $H_{\vec{G},\vec{0}}(k,\omega)$

We begin with the integration of (4.8). As the integration of the first two terms in $J(\vec{k},\omega)$ follows the same path we restrict ourselves to the third term in $J(\vec{k},\omega)$ which can be written as

$$J_3(\vec{k},\omega)=-\int_{-1}^{1}dp_\parallel\frac{M(p_\parallel)}{G^2+2Gp_\parallel}, \tag{A9}$$

where

$$M(x)=\int_0^{(1-x^2)^{1/2}}dp_\perp p_\perp\int_0^{2\pi}d\phi\,\frac{1}{\alpha(x)-2q_\perp p_\perp\cos\phi}, \tag{A10}$$

with

$$\alpha(x)=\omega-q^2-2q_\parallel x \tag{A11}$$

and $\vec{q}=\vec{k}+\vec{G}$, with $q_\parallel$, $q_\perp$ and $p_\parallel$, $p_\perp$ denoting vector components parallel and perpendicular to the polar axis defined by $\vec{G}$. The integration in (A10) is carried out as before by first performing the ϕ integration,[38] which gives

$$M(x)=\frac{\pi}{2q_\perp^2}\{1-[1-\psi^2(x)]^{1/2}\Theta(1-\psi^2(x))\}\alpha(x), \tag{A12}$$

where

$$\psi(x)=2q_\perp(1-x^2)^{1/2}/\alpha(x). \tag{A13}$$

If we substitute (A12) into (A9), we may write

$$J_3(\vec{k},\omega)=J_{31}(\vec{k},\omega)+J_{32}(\vec{k},\omega), \tag{A14}$$

where

$$J_{31}(\vec{k},\omega)=\frac{\pi q_{\parallel}}{2q_{\perp}^2 G}\int_{-1}^{1}dx\left(1-\frac{(\omega-q^2)/2q_{\parallel}+G/2}{x+G/2}\right), \tag{A15}$$

which is trivial to integrate, and

$$J_{32}(\vec{k},\omega)=\frac{\pi}{4q_{\perp}^2 G}\int_{-1}^{1}\frac{dx}{x+G/2}\,\alpha(x)[1-\psi^2(x)]^{1/2}\times\Theta(1-\psi^2(x)). \tag{A16}$$

If we define

$$P(x)=\alpha^2(x)-4q_{\perp}^2(1-x^2), \tag{A17}$$

we then have

$$J_{32}(\vec{k},\omega)=\frac{\pi}{4q_{\perp}^2 G}\int_{-1}^{1}\frac{dx}{x+G/2}\,\frac{P(x)}{\sqrt{P(x)}}\times\mathrm{sgn}[\alpha(x)]\Theta(P(x)). \tag{A18}$$

When $P(x)>0$ in the interval $[-1,1]$, $\alpha(x)$ cannot change sign in this interval [otherwise in at least one point x_i, $\alpha(x_i)$ would be zero and $P(x_i)<0$], and using the identity

$$\frac{P(x)}{x-\gamma}=\frac{P'(x)}{2}+\frac{P'(\gamma)}{2}+\frac{P(\gamma)}{x-\gamma} \tag{A19}$$

for $P(x)=ax^2+2bx+c$ (as it is the case), we see that (A18) is of the form

$$\int\frac{dx}{x-\gamma}\,\frac{P(x)}{\sqrt{P(x)}}=\tfrac{1}{2}\int dx\,\frac{P'(x)}{\sqrt{P(x)}}+\frac{P'(\gamma)}{2}\int\frac{dx}{\sqrt{P(x)}}+P(\gamma)\int\frac{dx}{(x-\gamma)\sqrt{P(x)}}, \tag{A20}$$

where all integrals are standard.[38]

When $P(x)<0$ in some interval (x_1,x_2) contained in the interval $[-1,1]$, it is easy to see that $x_{1/2}$ are the roots of $P(x)=0$. We then split (A18) in two integrals from -1 to x_1 and x_2 to $+1$ (assuming $x_2>x_1$) and proceed as before.

All the other integration corresponding to $H_{\delta,\delta}(\vec{k},\omega)$ are evaluated in a similar manner. It is important to note that all the important limiting cases were used to check the accuracy of the very lengthy analytical formula obtained, which we do not reproduce here as it offers no further direct physical insight.

---

*On leave from Universidade Federal de Alagoas, Departamento de Fisica, Maceió 57000, Alagoas, Brazil.

[1]H. Raether, in *Springer Tracts in Modern Physics* edited by G. Höhler (Springer, New York, 1965), Vol. 38, p. 85.

[2]J. Daniels, C. Festenberg, H. Raether and K. Zeppenfeld, in *Springer Tracts in Modern Physics* edited by G. Höhler (Springer, New York, 1970), Vol. 54, p. 77.

[3]S. E. Schnatterly, Solid State Phys. 34, 275 (1979).

[4]See, for example, D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1964).

[5]H. Ehrenreich and M. H. Cohen, Phys. Rev. 115, 786 (1959).

[6]K. Sturm, Z. Phys. B 25, 247 (1976); 28, 1 (1977); 29, 27 (1978); Solid State Commun. 25, 797 (1978).

[7]S. L. Adler, Phys. Rev. 126, 413 (1962); N. Wiser, *ibid.* 129, 62 (1963).

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

[9]J. Stiebling and H. Raether, Phys. Rev. Lett. 40, 1293 (1978).

[10]R. Manske, J. Phys. C 13, 911 (1980).

[11]K. Sturm, Phys. Rev. Lett. 40, 1599 (1978).

[12]L. E. Oliveira and K. Sturm, following paper, Phys. Rev. B 22, 6283 (1980).

[13]W. M. Saslow and G. F. Reiter, Phys. Rev. B 7, 2995 (1973).

[14]M. L. Cohen and T. K. Bergstresser, Phys. Rev. 141, 789 (1966).

[15]W. Saslow, T. K. Bergstresser, and M. L. Cohen, Phys. Rev. Lett. 16, 354 (1966).

[16]For simple metals this term represents the Drude absorption

$$\mathrm{Im}\,\epsilon_D(\omega)=\{(\omega_p^2\,1/\tau)/[\omega(\omega^2+1/\tau^2)]\}m/m_{\mathrm{opt}},$$

where $1/\tau$ is a global description of phonon-assisted transitions. If phonon (and impurity) scattering is neglected, then $1/\tau\to 0$ and

$$\mathrm{Im}\,\epsilon_D(\omega)=\pi(\omega_p^2/\omega)\delta(\omega)\,m/m_{\mathrm{opt}},$$

where m_{opt} is the optical mass.

[17]K. Sturm, J. Phys. C 12, 53 (1979).

[18]J. J. Hopfield, Phys. Rev. 139, A419 (1965).

[19]For a survey of the pseudopotential concept see V. Heine, M. L. Cohen, and D. Weaire, Solid State Phys. 24 (1970).

[20]J. Stiebling, Z. Phys. B31, 355 (1979); data for 18.6 eV $\leq\hbar\omega\leq$ 25.6 eV are from J. Stiebling, Diploma thesis, University of Hamburg, 1978 (unpublished).

[21]We used here a refined version of the pseudopotential of Ref. 14: J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B 10, 5095 (1974).

[22]B. Feuerbacher, R. P. Godwin, T. Sasaki, and M. Skibowski, J. Opt. Soc. Am. 58, 1434 (1968).

[23]K. Zeppenfeld and H. Raether, Z. Phys. 193, 471 (1966).

[24]C. v. Festenberg, Z. Phys. 227, 453 (1969).

[25]H. J. Hinz, Doctoral thesis, University of Hamburg, 1978 (unpublished).

[26]H. R. Philipp and H. Ehrenreich, Phys. Rev. 129, 1550 (1963).

[27]K. C. Pandey and J. C. Phillips, Solid State Commun. 19, 69 (1976). Values for E_g and D were taken from J. A. Van Vechten, Phys. Rev. 182, 891 (1969); 187, 1007 (1969). Notice that Eq. (3.26) in the present work differs by a factor of 2 from Eq. (3) in the paper by Pandey and Phillips but agrees with J. C. Phillips, Rev. Mod. Phys. 42, 317 (1970), Eq. (16.1) and A. Bardasis and D. Hone, Phys. Rev. 153, 849 (1967), Eq. (2.15). Also, in Eq. (3.25) we ignored the constant that Pandey and Phillips arbitrarily subtracted from $[\epsilon_2^P(\omega_p)]^{-1}$ to obtain $(\mathrm{Im}-\epsilon^{-1})_{\max}$.

[28]It has been pointed out by J. Appelbaum [Phys. Rev. 144, 435 (1966)] and A. O. E. Animalu [*ibid.* 163, 557 (1967)] that, in principle, matrix elements must be evaluated with the true wave functions which differ from the pseudo-wave-functions by core orthogonalization terms. For solids with tightly bound cores these corrections are generally small. We neglect them here in any case.

[29]W. Hanke, Adv. Phys. 27, 287 (1978) and references therein.

[30]W. Hanke and L. J. Sham, Phys. Rev. Lett. 43, 387 (1979).

[31]K. Sturm, Solid State Commun. 27, 645 (1978).

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

[33]M. Crenzburg, Z. Phys. 196, 433 (1966).

[34]J. R. Fields, P. C. Gibbons, and S. E. Schnatterly, Phys. Rev. Lett. 38, 430 (1977).

[35]The recently measured anisotropy of the plasmon dispersion in Al could also be explained with the present theory (see Ref. 6).

[36]R. Girlanda, M. Parrinello, and E. Tosatti, Phys. Rev. Lett. 36, 1386 (1976); see also E. Tosatti, in *Interaction of Radiation with Condensed Matter* (IAEA, Vienna, 1977), Vol. 1, p. 281.

[37]M. Urner-Wille and H. Raether, Phys. Lett. 58A, 265 (1976).

[38]W. Gröbner and N. Hofreiter, *Integral Tables*, 2nd ed. (Springer, Vienna, 1957), Vol. 1, pp. 37 and 38; *ibid.* (Springer, Vienna, 1958), Vol. 2, p. 99.