PHYSICAL REVIEW B VOLUME 10, NUMBER 6 15 SEPTEMBER 1974

Improved calculations of the complex dielectric constant of semiconductors

Roger A. Breckenridge
NASA Langley Research Center, Hampton, Virginia 23665

Robert W. Shaw, Jr.
Booz-Allen Applied Research, Bethesda, Maryland 20014

Arden Sher*
College of William and Mary, Williamsburg, Virginia 23185

(Received 6 May 1974)

Expressions for the real, static part $\epsilon_1(0)$ and the imaginary part $\epsilon_2(\omega)$ of the dielectric constant of semiconductors in the long-wavelength limit are obtained using the isotropic nearly-free-electron band approximation (Penn model). Earlier calculations of these functions do not satisfy the Kramers-Kronig relations and yield an excessively large result for the f-sum rule. The corrected expressions eliminate these inconsistencies. Values of the energy gap between the bonding and antibonding states are obtained for diamond, silicon, and germanium, respectively. $\epsilon_1(\omega)$ is obtained from $\epsilon_2(\omega)$ through the use of the Kramers-Kronig relation. The theoretical curves for $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ are compared with experimental results.

I. INTRODUCTION

Although detailed band-structure calculations have been used to evaluate the dielectric constant for semiconductors with excellent results,[1] the appeal of approximate calculations based on the isotropic nearly-free-electron band approximation remains strong due to the simplicity of the analytical results obtained. Several authors have reported calculations of the long-wavelength limit of $\epsilon_1(0)$, the real part of the dielectric function in the static limit,[2,3], and $\epsilon_2(\omega)$, the imaginary part.[4] However, it can be shown that the expressions given previously do not satisfy the Kramers-Kronig relations and give an excessively large value for the f-sum rule. These inconsistencies motivated the present work, an effort we felt was particularly important due to the dependence of the Phillips-Van Vechten theory of ionicity on the Penn expression[2] for $\epsilon_1(0)$.

We have found that all the inconsistencies in the earlier results can be eliminated within the framework of the Penn model and that the resulting expressions for $\epsilon_1(0)$ and $\epsilon_2(\omega)$ differ from those given previously. The calculations of the dielectric-function components are presented in Sec. II. In Sec. III we verify that the Kramers-Kronig relation is satisfied for $\epsilon_1(0)$ and $\epsilon_2(\omega)$ and we calculate $\epsilon_1(\omega)$ from $\epsilon_2(\omega)$ by use of the Kramers-Kronig relation. Then we show that $\epsilon_2(\omega)$ closely satisfies the f-sum rule.[5] In Sec. IV we obtain values for the energy gap of silicon, germanium, and diamond from experimental values of $\epsilon_1(0)$, and compare our theoretical calculations to experimental results for $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$. Finally, in Sec. V we indicate the implications of our calculations relative to theoretical developments based on the Penn result.

II. CALCULATION OF $\epsilon_1(0)$ AND $\epsilon_2(\omega)$ USING THE PENN MODEL

The complex dielectric function for a solid is given in the reduced zone by Ehrenreich and Cohen[6] as

$$\epsilon(\omega, q) = 1 - \lim_{\alpha \to 0} \frac{4\pi e^2}{q^2 \Omega} \sum_{\vec{k}, l, l'} \frac{|\langle \vec{k} l | e^{-i\vec{q}\cdot\vec{r}} | \vec{k}+\vec{q}, l' \rangle|^2 [f_0(E_{\vec{k}+\vec{q}, l'}) - f_0(E_{\vec{k} l})]}{E_{\vec{k}+\vec{q}, l'} - E_{\vec{k} l} - \hbar\omega + i\hbar\alpha}, \tag{2.1}$$

where $f_0(E_{\vec{k}l})$ is the distribution function for the reduced wave vector $\vec{k}$ and band l, and Ω is the volume of the solid. $|\vec{k}l\rangle$ is a Bloch wave function which satisfies $H_0|\vec{k}l\rangle = E_{\vec{k}l}|\vec{k}l\rangle$, where H_0 is the Hamiltonian for an electron in the unperturbed periodic lattice.

$\epsilon(\omega, q)$ can be expressed in the extended zone scheme,[7]

$$\epsilon(\omega, q) = 1 - \lim_{\alpha \to 0} \frac{4\pi e^2}{q^2 \Omega} \sum_{\vec{k}, \vec{G}} \frac{|\langle \vec{k} | e^{-i\vec{q}\cdot\vec{r}} | \vec{k}+\vec{q}+\vec{G} \rangle|^2 (N_{\vec{k}+\vec{q}+\vec{G}} - N_{\vec{k}})}{E_{\vec{k}+\vec{q}+\vec{G}} - E_{\vec{k}} - \hbar\omega + i\hbar\alpha}, \tag{2.2}$$

where the distribution functions have been replaced by occupation numbers, corresponding to the condition of zero temperature, and $\vec{G}$ is a reciprocal-lattice vector. The complex dielectric function is defined here as

$$\epsilon(\omega, q) = \epsilon_1(\omega, q) - i\epsilon_2(\omega, q) , \tag{2.3}$$

so that $\epsilon_2(\omega, q)$ will be a positive quantity.

From (2.2) and (2.3),

$$\epsilon_1(\omega, q) = 1 - \frac{4\pi e^2}{q^2\Omega} \sum_{\vec{k},\vec{G}} |\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q}+\vec{G}\rangle|^2 (N_{\vec{k}+\vec{q}+\vec{G}} - N_{\vec{k}})\, P[1/(E_{\vec{k}+\vec{q}+\vec{G}} - E_{\vec{k}} - \hbar\omega)] , \tag{2.4}$$

where P indicates the principal value, and

$$\epsilon_2(\omega, q) = -\frac{4\pi^2 e^2}{q^2\Omega} \sum_{\vec{k},\vec{G}} |\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q}+\vec{G}\rangle|^2 (N_{\vec{k}+\vec{q}+\vec{G}} - N_{\vec{k}})\, \delta(E_{\vec{k}+\vec{q}+\vec{G}} - E_{\vec{k}} - \hbar\omega) . \tag{2.5}$$

We are interested in $\epsilon_1(0, q)$ and $\epsilon_2(\omega, q)$ in the limit as q goes to zero and obtain closed form expressions for these quantities from (2.4) and (2.5) by use of the Penn model for a semiconductor. The model is equivalent to the nearly-free-electron model, isotropically extended to three dimensions, as shown in Fig. 1. The energies of the two bands are given by

$$E^{\pm}_{\vec{k}} = \tfrac{1}{2}\{E^0_{\vec{k}} + E^0_{\vec{k}'} \pm [(E^0_{\vec{k}} - E^0_{\vec{k}'})^2 + E_g^2]^{1/2}\} , \tag{2.6}$$

where $E^0_{\vec{k}} = \hbar^2 k^2/2m$ and $\vec{k}' = \vec{k} - 2k_F\hat{k}$. $k_F\hat{k}$ is the Fermi wave vector. The wave functions are

$$\psi^{\pm}_{\vec{k}}(\vec{r}) = \frac{e^{i\vec{k}\cdot\vec{r}}}{\Omega^{1/2}} \left(\frac{1 - \beta^{\pm}_{\vec{k}} e^{-i2k_F\hat{k}\cdot\vec{r}}}{[1 + (\beta^{\pm}_{\vec{k}})^2]^{1/2}}\right) , \tag{2.7}$$

where

$$\beta^{\pm}_{\vec{k}} = \frac{E_g}{E^0_{\vec{k}} - E^0_{\vec{k}'} + [(E^0_{\vec{k}} - E^0_{\vec{k}'})^2 + E_g^2]^{1/2}} .$$

In the preceding expressions, the − and + superscripts refer to $\vec{k}$ in the first Brillouin and $\vec{k}$ in the second Brillouin zone, respectively.

Since we are interested in $\epsilon_1(0, q)$ and $\epsilon_2(\omega, q)$ as q goes to zero, we use the Srinivasan condition[3] that for a given $\vec{k}_1$ then $\vec{k}_2 = \vec{k}_1 + q$ does not have an independent reciprocal-lattice vector if the angle between $\vec{k}_1$ and $\vec{k}_2$ is smaller than the smallest angle between the reciprocal-lattice vectors of interest in a real crystal. Srinivasan argues that in a real crystal there are only a discrete number of reciprocal-lattice vectors, so assigning separate reciprocal-lattice vectors to all $\vec{k}_1$ and $\vec{k}_2$ is not physical. In our calculations, this condition should definitely hold since q is always small and will finally be allowed to go to zero. There are two values of $\vec{G}$ associated with the Penn model: $\vec{G} = 0$ which corresponds to the normal process and $\vec{G} = -2k_F\hat{k}$ which corresponds to the umklapp process. Since we are using the Srinivasan condition, $\vec{k}$ and $\vec{k}+\vec{q}$ have the same reciprocal-lattice vector, $-2k_F\hat{k}$.

Substituting (2.6) and (2.7) in the general expressions, we obtain

$$\epsilon_1(0) = 1 + \lim_{q\to 0} \frac{8\pi e^2}{q^2\Omega} \sum_{\vec{k}} N_{\vec{k}}(1 - N_{\vec{k}+\vec{q}}) \frac{|\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q}\rangle|^2}{E^+_{\vec{k}+\vec{q}} - E^-_{\vec{k}}} + \lim_{q\to 0} \frac{8\pi e^2}{q^2\Omega} \sum_{\vec{k}} N_{\vec{k}}(1 - N_{\vec{k}'+\vec{q}}) \frac{|\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}'+\vec{q}\rangle|^2}{E^+_{\vec{k}'+\vec{q}} - E^-_{\vec{k}}} ; \tag{2.8}$$

and

$$\epsilon_2(\omega) = \lim_{q\to 0} \frac{4\pi^2 e^2}{q^2\Omega} \left(\sum_{\vec{k}} N_{\vec{k}}(1 - N_{\vec{k}+\vec{q}}) |\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q}\rangle|^2 \delta(E^+_{\vec{k}+\vec{q}} - E^-_{\vec{k}} - \hbar\omega) + \sum_{\vec{k}} N_{\vec{k}}(1 - N_{\vec{k}'+\vec{q}}) |\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}'+\vec{q}\rangle|^2 \delta(E^+_{\vec{k}'+\vec{q}} - E^-_{\vec{k}} - \hbar\omega)\right) , \tag{2.9}$$

where ω is restricted to positive frequencies. $N_{\vec{k}} = 1$ for $\vec{k}$ in the first Brillouin zone and $N_{\vec{k}} = 0$ for $\vec{k}$ elsewhere.

To perform the integrals for the calculation of the matrix elements, we assume a real crystal. Evaluation of the matrix elements yields

$$|\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q}\rangle|^2 = \frac{(1 + \beta^-_{\vec{k}}\beta^+_{\vec{k}+\vec{q}})^2}{[1 + (\beta^-_{\vec{k}})^2][1 + (\beta^+_{\vec{k}+\vec{q}})^2]}$$

and (2.10)

$$|\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}'+\vec{q}\rangle|^2 = \frac{(\beta^-_{\vec{k}} + \beta^+_{\vec{k}'+\vec{q}})^2}{[1 + (\beta^-_{\vec{k}})^2][1 + (\beta^+_{\vec{k}'+\vec{q}})^2]} .$$

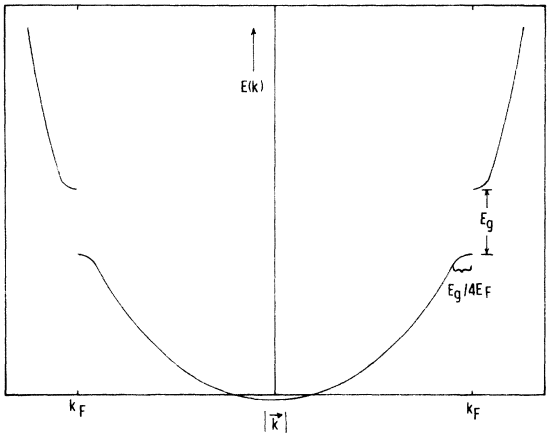

FIG. 1. Electron energy as a function of wave number for an isotropic three-dimensional nearly-free-electron model.

If the expressions in (2.10) are examined as q tends to zero, the leading term goes as

$$|\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}+\vec{q}\rangle|^2 = |\langle\vec{k}|e^{-i\vec{q}\cdot\vec{r}}|\vec{k}'+\vec{q}\rangle|^2 \cong \frac{z^2\Delta^2\eta^2}{4(x^2+\Delta^2)^2}\ , \qquad (2.11)$$

where $z=\cos\theta$, θ is the angle between $\vec{k}$ and $\vec{q}$, $x=1-k/k_F$, $\Delta=E_g/4E_F$, $\eta=q/k_F$, and $E_F=\hbar^2k_F^2/2m$.

The normal process does not contribute to $\epsilon_1(0)$[3] or $\epsilon_2(\omega)$,[4] so (2.8) and (2.9) become

$$\epsilon_1(0)=1+\frac{8\pi e^2}{k_F^2\Omega}\sum_{\vec{k}(1\text{st BZ})}\frac{z^2\Delta^2}{4(x^2+\Delta^2)^2}\frac{1}{E^+_{\vec{k}'}-E^-_{\vec{k}}} \qquad (2.12)$$

and

$$\epsilon_2(\omega)=\frac{4\pi^2e^2}{k_F^2\Omega}\sum_{\vec{k}(1\text{st BZ})}\frac{z^2\Delta^2}{4(x^2+\Delta^2)^2}\times\delta(E^+_{\vec{k}'}-E^-_{\vec{k}}-\hbar\omega)\ , \qquad (2.13)$$

respectively, where the limit as $q\to 0$ has been taken. The sum on $\vec{k}$ over the first Brillouin zone is converted into an integral, and the integrals completed by quadrature to give

$$\epsilon_1(0)=1+\frac{2}{3}\frac{E_p^2}{E_g^2}[(1+\Delta^2)^{1/2}-\Delta] \qquad (2.14)$$

and

$$\epsilon_2(E)=\frac{\pi}{2}\frac{E_p^2}{E^3}\frac{[E_g-\Delta(E^2-E_g^2)^{1/2}]^2}{(E^2-E_g^2)^{1/2}}\ , \qquad (2.15)$$

where $E_p=(4\hbar^2\pi e^2n/m)^{1/2}$ is plasma energy and $E=\hbar\omega$. Equation (2.15) holds for $E_g\le E \le 4E_F(1+\Delta^2)^{1/2}$ and $\epsilon_2=0$ for other values of E. The procedure for setting $q=0$ directly in the δ function in (2.13) is somewhat facile, but it can be shown (see Appendix A) that if q is retained to second order in the δ function, the same expression as in (2.15) is obtained which holds for the energy range $E_g<E\le 4E_F(1+\Delta^2)^{1/2}$.

Equation (2.14) differs from the result given by Penn,[2]

$$\epsilon_1(0)=1+\frac{E_p^2}{E_g^2}(1-\Delta+\tfrac{1}{3}\Delta^2)\ .$$

The factor of $\frac{2}{3}$ appearing in the second term of (2.14) is of particular importance in calculations of a homopolar energy gap E_g from measured values of $\epsilon_1(0)$. The difference arises because of the two approximations that Penn employs. The first approximation is that the electron energy is the free-electron energy for $k<(1-\Delta)k_F$ and for $k>(1+\Delta)k_F$; but $E^-_{\vec{k}}\approx(1-\Delta)E_F$ for $(1-\Delta)k_F<k<k_F$ and $E^+_{\vec{k}}\approx(1+\Delta)E_F$ for $k_F<k<(1+\Delta)k_F$. Penn indicates that this approximation will introduce an error in $\epsilon_1(q)$ that is of order E_g/E_F which is sizeable. The second approximation is that Penn uses the matrix-element values at the Brillouin-zone boundary to perform his calculations. These two approximations tend to maximize the value of $\epsilon_1(0)$.

Equation (2.15) differs from the Bardasis and Hone result[4] by a factor of $\frac{1}{2}$, but has the same functional dependence on E_p, E_g, E_F, and E. In Sec. III, we present additional evidence to support the reduction by $\frac{1}{2}$ in the strength of the absorption. Furthermore, the expression for $\epsilon_2(\omega)$ given by Bardasis and Hone holds for the same region of frequency (or energy) as ours, but the derivation given by these authors does not make this constraint on the result explicit.

III. KRAMERS-KRONIG RELATION AND THE *f*-SUM RULE

A. Verification and application of the Kramers-Kronig relation

The Kramers-Kronig relation of interest can be shown to be

$$\epsilon_1(E)-1=\frac{2}{\pi}P\int_0^\infty\frac{\epsilon_2(E')E'\,dE'}{E'^2-E^2}\ , \qquad (3.1)$$

and for $E=0$,

$$\epsilon_1(0)-1=\frac{2}{\pi}P\int_0^\infty\frac{\epsilon_2(E')\,dE'}{E'}\quad . \qquad (3.2)$$

Since the δ function in $\epsilon_2(E')$ restricts its nonzero range to $E_g<E'\le 4E_F(1+\Delta^2)^{1/2}$, the principal-value sign can be dropped. From (2.13) and (3.2),

$$\epsilon_1(0)-1=\frac{2}{\pi}\frac{e^2\Delta^2k_F}{3}\int_0^\infty\int_0^1\frac{(1-x)^2\,dx}{(x^2+\Delta^2)^2}\frac{dE'}{E'}\times\delta(4E_F(x^2+\Delta^2)^{1/2}-E')\ . \qquad (3.3)$$

We obtain the following expression for $\epsilon_1(0)$:

$$\epsilon_1(0)=1+\frac{2}{3}\frac{E_p^2}{E_g^2}[(1+\Delta^2)^{1/2}-\Delta]\ , \qquad (3.4)$$

which is in complete agreement with (2.14). Thus, the Kramers-Kronig relation is satisfied by our expressions for $\epsilon_1(0)$ and $\epsilon_2(\omega)$, a condition which is not met by expressions for the dielectric constants derived previously using the Penn model. It can be easily shown, for example, that Penn's expression for $\epsilon_1(0)$ and Bardasis and Hone's expression for $\epsilon_2(E)$ do not satisfy (3.2).

For $E < E_g$ $(\omega < \omega_g)$, we examine the Kramers-Kronig relation to obtain $\epsilon_1(E)$. From (2.15) and (3.1), we obtain

$$\epsilon_1(E) - 1 = \frac{e^2\Delta^2 k_F}{6\pi E_F} \int_0^1 \frac{(1-x)^2\,dx}{(x^2+\Delta^2)^{3/2}[x^2+\Delta^2-(E/4E_F)^2]}\ , \tag{3.5}$$

where the principal-value sign has been dropped because the integral exists. We can write $\epsilon_1(E)$ as

$$\epsilon_1(E) = 1 + \tfrac{2}{3}(E_p^2/E_g^2)I(E)\ , \tag{3.6}$$

where

$$I(E) = \tfrac{3}{2}\Delta^2 \int_0^1 \frac{(1-x)^2\,dx}{(x^2+\Delta^2)^{3/2}[(x/\Delta)^2+1-(E/E_g)^2]}\ . \tag{3.7}$$

This expression has been evaluated numerically for selected values of Δ (see Sec. IV). For $E_g < E < 4E_F(1+\Delta^2)^{1/2}$ a similar procedure has been followed, except the integral has been performed exactly.

B. f-sum rule

The f-sum rule can be written in the form

$$\frac{2}{\pi}\int_0^\infty E\,\epsilon_2(E)\,dE = E_p^2\ . \tag{3.8}$$

We can use the expression for $\epsilon_2(\omega)$ obtained earlier [Eq. (2.15)] to sum the oscillator strength and obtain

$$\frac{2}{\pi}\int_0^\infty E\,\epsilon_2(E)\,dE = E_p^2\left\{(1+\Delta^2)^{1/2} - 2\Delta + \Delta^2 \ln\left[\frac{1+(1+\Delta^2)^{1/2}}{\Delta}\right]\right\}\ . \tag{3.9}$$

For typical semiconductors, Δ is usually near 0.1, we therefore use this value to obtain

$$\frac{2}{\pi}\int_0^\infty E\,\epsilon_2(E)\,dE = 0.835E_p^2\ .$$

Evidently the simple two-band model proposed by Penn artifically reduces the total oscillator strength either by reducing the density of states at the gap or by eliminating the possibility of transitions to other bands. To estimate the importance of additional bands, we have used a nearly-free-electron three-band model which is isotropically extended to three dimensions to determine the effect of including more energy states on the f-sum rule. For $\Delta = 0.1$, essentially the same result is obtained which implies that the inclusion of more energy states does not improve the situation. A plausible explanation for why the f-sum rule is not satisfied exactly is that the nearly-free-electron model yields approximate wave functions and energies because a perturbation expansion has been employed to obtain them.

The Bardasis and Hone result for $\epsilon_2(\omega)$ is a factor of 2 larger than ours and for $\Delta = 0.1$, gives

$$\frac{2}{\pi}\int_0^\infty E\,\epsilon_2(E)\,dE = 1.67E_p^2\ ;$$

a result which is in significant disagreement with the f-sum rule. If one does not restrict the range of energy over which the expression for $\epsilon_2(E)$ is valid in conducting the integral, the discrepancy is even larger. The present calculation brings $\epsilon_2(E)$ is closer agreement with the f-sum rule but leaves open the question of why the total oscillator strength is too small in the Penn model.

IV. COMPARISON OF THEORETICAL CALCULATIONS WITH EXPERIMENTAL RESULTS

To obtain E_g, the energy gap between the bonding and the antibonding states, (2.14) can be approximated as

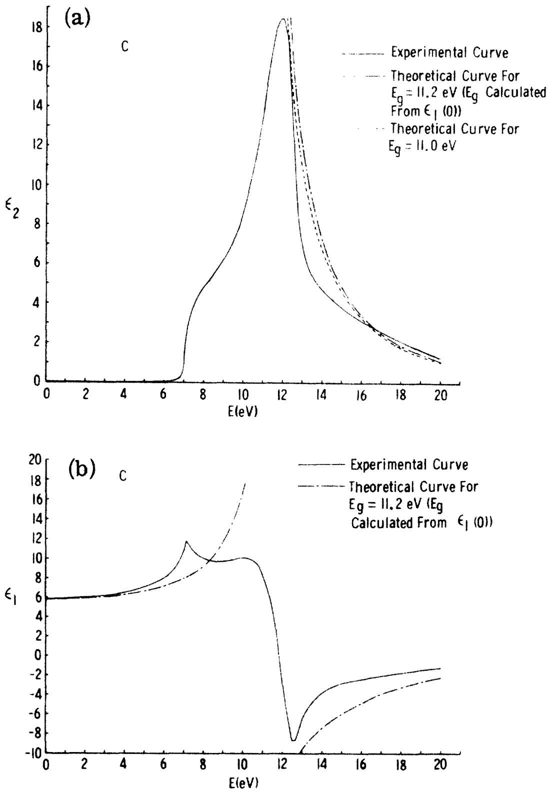

FIG. 2. Imaginary and real parts of the dielectric constant in the long-wavelength limit $(q \to 0)$ for diamond, as measured optically and as predicted by theory.

$$\epsilon_1(0) \cong 1 + \frac{2}{3}\frac{E_p^2}{E_g^2}\left[1 - \left(\frac{E_g}{4E_F}\right)\right]. \qquad (4.1)$$

Experimental values of $\epsilon_1(0)$ are 5.7, 12, and 16 for diamond, silicon, and germanium, respectively. From (4.1), we calculate that E_g is 11.2 eV for diamond and 3.92 eV for silicon. If we include Van Vechten's empirical value[8] D to take into account d-state effects in germanium, then (4.1) becomes

$$\epsilon_1(0) \cong 1 + \frac{2}{3}\frac{DE_p^2}{E_g^2}\left[1 - \left(\frac{E_g}{4E_F}\right)\right]. \qquad (4.1')$$

For germanium we use the Van Vechten value of $D = 1.25$ to obtain a value of 3.53 eV for E_g. We note that our values for E_g are lower than the values of 13.6, 4.8, and 4.3 eV for diamond, silicon, and germanium, respectively, obtained from the Penn expression for $\epsilon_1(0)$ because of the presence of the $\frac{2}{3}$ in our expression.

Our value of E_g for diamond agrees fairly well with Harrison's value of 10.8 eV for the covalent bonding energy of diamond,[9] although he has taken a bond-orbital-model approach to the calculation of the static dielectric constant. Harrison's covalent bonding energy values are 3.0 and 2.7 eV for silicon and germanium, respectively, substantially smaller than the values we obtain. However, Harrison indicates that for group-IV semiconductors with a sufficiently small gap a Penn-type model probably yields better results, but that his model is better for large gaps.[10]

Since $\epsilon_{\parallel} = \epsilon_{\perp}$ in the limit of small q,[11] the energy gaps obtained from (4.1) and (4.1′) can be used to generate theoretical curves for $\epsilon_2(E)$ by the use of (2.15), where E is the optical energy. For germanium, we must multiply E_p^2 by Van Vechten's D to take into account the effects of d core states on $\epsilon_2(E)$. Figures (2a), (3a), and (4a) serve to compare theoretical curves with experimental results. To illustrate the dependence of the theoretical result on the value of E_g, these figures include an additional curve for an arbitrarily selected E_g. The experimental results for silicon and germanium are those of Philipp and Ehrenreich,[12] while the experimental results for diamond are those of Philipp and Taft.[13] The theoretical curves for E_g calculated from $\epsilon_1(0)$ agree fairly well with the experimental results but do not exhibit the structure of the experimental curves, because all interactions have been ignored with the exception of the interactions of the electron with the electromagnetic field and the electron gas through the self-consistent-field approximation.[6] Heine and Jones[14] have proposed that the higher-energy peak of $\epsilon_2(E)$ is associated with E_g which seems to be consistent with our results.

The theoretical curves for $\epsilon_1(E)$ are obtained from the Kramers-Kronig relation involving $\epsilon_2(E)$. The theoretical values of E_g for diamond, silicon, and germanium are substituted into (3.5). For $E < E_g$, the integral is evaluated numerically; for $E_g < E < 4E_F(1+\Delta^2)^{1/2}$, the integral is performed exactly. In Figs. (2b), (3b), and (4b), the theoretical curves are shown so that a comparison can be drawn between them and the experimental curves. In all cases, the theoretical values are far too low except at low energies, but the theoretical curves do exhibit some qualitative agreement with the experimental curves.

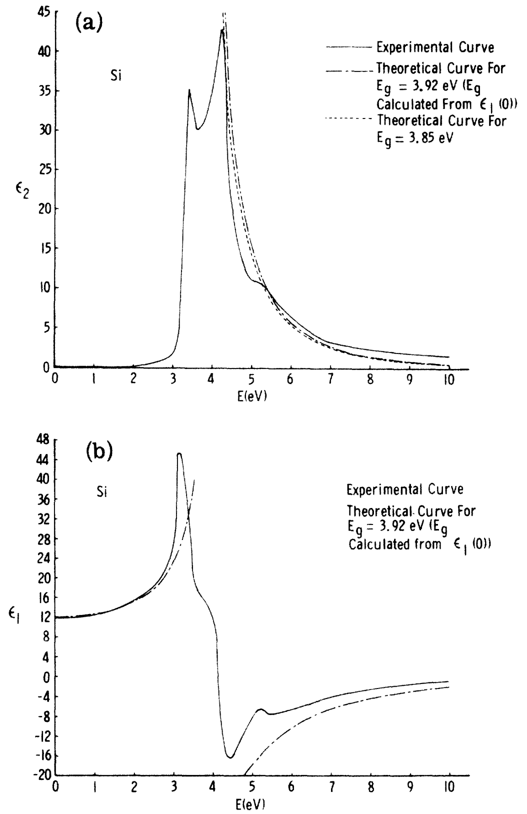

FIG. 3. Imaginary and real parts of the dielectric constant in the long-wavelength limit ($q \to 0$) for silicon, as measured optically and as predicted by theory.

V. CONCLUDING REMARKS

We have obtained corrected expressions for the real and imaginary parts of the dielectric constant of semiconductors in the long-wavelength ($q \to 0$) limit through the use of the Penn model. Since Phillips's theory of ionicity[15] is based on Penn's expression for $\epsilon_1(0)$ rather than the correct result given by (2.14), the values of E_h, C, and E_g in his theory are subject to some error. E_h and C are the average homopolar energy gap and the average heteropolar energy gap, respectively.[8] (For example, Phillips's values for E_h, C, and E_g for gal-

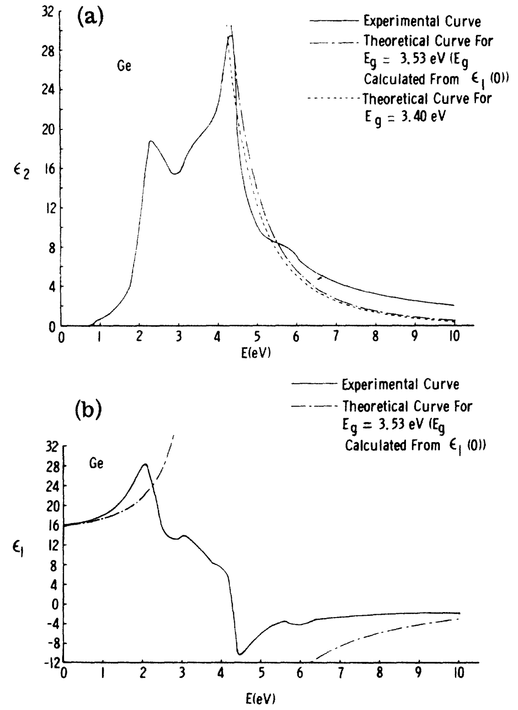

FIG. 4. Imaginary and real parts of the dielectric constant in the long-wavelength limit ($q \to 0$) for germanium, as measured optically and as predicted by theory (with the Van Vechten correction for core d-state effects included).

lium arsenide are 4.32, 2.90, and 5.20 eV, respectively.[16] The correct values of Phillips's parameters are 3.55, 2.39, and 4.28 eV.)

However, we have calculated the ionicities that would result from our expression for 34 crystals, and we find that they differ only slightly from Phillips's values although the gaps E_h and C are significantly different. The reason for this is that the ionicity as defined by Phillips is not sensitive to the constant which multiplies $(E_p^2/E_g^2) \times [1-(E_g/4E_F)]$ in the expression for $\epsilon_1(0)$. In fact, if the term $E_g/4E_F$ is neglected, the ionicity can be shown to be independent of this constant. An investigation of the impact of the correction in $\epsilon_1(0)$ on the Phillips–Van Vechten theory is underway and will be reported elsewhere.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the helpful comments by C. L. Fales. They also thank W. F. White for his suggestions relative to the computer programs used and W. J. Debnam for his efforts. Finally, they wish to express their appreciation to the National Aeronautics and Space Administration who supported the research from which this paper originates.

APPENDIX A

In this appendix we derive the expression (2.15) presented above for $\epsilon_2(\omega)$ when terms to second order in q in the δ function, $\delta(E^{+}_{\vec{k}'+\vec{q}} - E^{-}_{\vec{k}} - \hbar\omega)$ are retained. We begin with equation (2.13) for $\epsilon_2(E)$,

$$\epsilon_2(E) = \lim_{q\to 0} \frac{4\pi^2 e^2}{k_F^2 \Omega} \sum_{\vec{k}(\text{1st BZ})} \frac{z^2\Delta^2}{4(x^2+\Delta^2)^2} \times \delta(E^{+}_{\vec{k}'+\vec{q}} - E^{-}_{\vec{k}} - E)\ , \tag{A1}$$

where $E=\hbar\omega$. To second order in q, the energy difference is

$$E^{+}_{\vec{k}'+\vec{q}} - E^{-}_{\vec{k}} \cong E_F\{4(x^2+\Delta^2)^{1/2} - 2x[1+(x^2+\Delta^2)^{-1/2}]\eta z + \eta^2 z^2 (x^2+\Delta^2)^{-1/2} + \eta^2\}\ , \tag{A2}$$

where $\eta = q/k_F$ and $x = 1 - k/k_F$.

We convert the sum on $\vec{k}$ in the first Brillouin zone to an integral, and obtain

$$\epsilon_2(E) \cong \lim_{q\to 0} \frac{e^2\Delta^2 k_F}{2} \int_{-1}^{1}\int_{0}^{1} \frac{(1-x)^2 z^2\, dz\, dx}{(x^2+\Delta^2)^2} \times \delta(E_F[4(x^2+\Delta^2)^{1/2} - 2x[1+(x^2+\Delta^2)^{-1/2}]\eta z + \eta^2 z^2 (x^2+\Delta^2)^{-1/2} + \eta^2] - E)\ , \tag{A3}$$

where the lower limit on the x integration has been set equal to 0. Although this is an approximation, since q is small but not equal to zero, it can be shown to introduce only terms of higher order in q than we will retain.

The δ function can be written

$$\delta(f(x)) = \left(1 \Big/ \left|\frac{df}{dx}\right|\right) \delta(x - x_0)\ , \tag{A4}$$

where

$$f(x) = E_F\{4(x^2+\Delta^2)^{1/2} - 2x[1+(x^2+\Delta^2)^{-1/2}]\eta z + \eta^2 z^2 (x^2+\Delta^2)^{-1/2} + \eta^2\} - E$$

and $f(x_0) = 0$.

We first examine $f(x)$ and investigate the case where $E = E_g$. At $x=0$, $f(0) > 0$ since the limit as $q \to 0$ has not been taken. For $x \neq 0$, $4E_F(x^2+\Delta^2)^{1/2} - E_g$ is positive and always dominates for sufficiently small η (or q). Therefore, $f(x) > 0$ for $E = E_g$ and the δ function cannot be satisfied. This means that the point $E = E_g$ contributes nothing to

$\epsilon_2(E)$.

Next we set $f(x)$ equal to zero to determine x_0:

$$4(x^2+\Delta^2)^{1/2}-2x[1+(x^2+\Delta^2)^{-1/2}]\eta z + \eta^2 z^2 (x^2+\Delta^2)^{-1/2}+\eta^2 - E/E_F = 0 \ . \tag{A5}$$

If terms involving η and η^2 are neglected, the solution is

$$x_1 = [(E/4E_F)^2-\Delta^2]^{1/2} \ .$$

Now let $x_0 = x_1 + \alpha$, where α is small relative to x_1, and x_0 is the solution to (A5). We make use of binomial expansions and find that to second order in η,

$$x_0 \cong x_1 + A(E)\eta z + B_1(E)\eta^2 z^2 + B_2(E)\eta^2 \ , \tag{A6}$$

where

$$A(E)=\frac{1}{2}\left(1+\frac{E}{4E_F}\right) ,$$

$$B_1(E)=\frac{1}{4}\left[\left(\frac{E}{4E_F}\right)^2-\Delta^2\right]^{-1/2} \times\left[-1+\left(\frac{E}{4E_F}\right)^2+\left(\frac{E}{4E_F}\right)-\frac{1}{2}\Delta^2+\frac{1}{2}\Delta^2\left(\frac{4E_F}{E}\right)^2\right],$$

and

$$B_2(E)=-\frac{1}{4}\left(\frac{E}{4E_F}\right)\left[\left(\frac{E}{4E_F}\right)^2-\Delta^2\right]^{-1/2} .$$

It can be shown that the expression for x_0 is valid for all $E > E_g + O(\eta^n)$ where $0 < n < 2$ and the region of invalidity vanishes in the limit as $\eta \to 0$ except for the point E_g.

Finally, we have

$$\delta(E_F\{4(x^2+\Delta^2)^{1/2}-2x[1+(x^2+\Delta^2)^{-1/2}]\eta z+\eta^2 z^2(x^2+\Delta^2)^{-1/2}+\eta^2\}-E) \cong \frac{(x^2+\Delta^2)^{3/2}}{E_F\{4x(x^2+\Delta^2)-2\eta z[\Delta^2+(x^2+\Delta^2)^{3/2}]-\eta^2 z^2 x\}}\,\delta(x-x_0) \ , \tag{A7}$$

where x_0 is given by (A6). From (A3) and (A7), we have

$$\epsilon_2(E)=\lim_{q\to 0}\frac{e^2\Delta^2 k_F}{2}\int_{-1}^{1}\int_{0}^{1}\frac{(1-x)^2 z^2\delta(x-x_0)\,dz\,dx}{E_F(x^2+\Delta^2)^{1/2}\{4x(x^2+\Delta^2)-2\eta z[\Delta^2+(x^2+\Delta^2)^{3/2}]-\eta^2 z^2 x\}} \tag{A8}$$

The expression we obtain is

$$\epsilon_2(E)=\frac{\pi}{2}\,\frac{E_p^2}{E^3}\,\frac{[E_g-\Delta(E^2-E_g^2)^{1/2}]^2}{(E^2-E_g^2)^{1/2}} \ , \tag{A9}$$

for $E_g < E \leq 4E_F(1+\Delta^2)^{1/2}$ and $\epsilon_2(E)=0$ for other values of E. Hence, a careful evaluation of the limit yields our earlier result except directly at the energy gap, where the limit is not defined.

It can be shown that the Kramers-Kronig relation for $\epsilon_1(0)$ and $\epsilon_2(E)$ is satisfied through the use of (A8). In addition, the agreement with the sum rule is the same as that obtained in (3.9).

---

*Supported in part by N.S.F. Grant No. GH-41082 and by NASA Grant No. NGL-47-006-055.

[1] John P. Walter and Marvin L. Cohen, Phys. Rev. B 5, 3101 (1972); S. J. Sramek and Marvin L. Cohen, Phys. Rev. B 6, 3800 (1972).

[2] D. R. Penn, Phys. Rev. 128, 2093 (1962).

[3] G. Srinivasan, Phys. Rev. 178, 1244 (1969).

[4] A. Bardasis and Daniel Hone, Phys. Rev. 153, 849 (1967).

[5] P. Nozières and D. Pines, Phys. Rev. 113, 1254 (1959).

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

[7] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge U.P., London, 1965), Sec. 5.6. The expression for $\epsilon(\omega, q)$ given by Ziman involves $e^{i\vec{q}\cdot\vec{r}}$ in the matrix element as a result of the way in which the perturbing potential and wave functions were selected.

[8] J. A. Van Vechten, Phys. Rev. 182, 891 (1969).

[9] Walter A. Harrison, Phys. Rev. B 8, 4487 (1973).

[10] Walter A. Harrison (private communication).

[11] David Pines, *Elementary Excitation in Solids* (Benjamin, New York, 1963), Chap. 4.

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

[13] H. R. Philipp and E. A. Taft, Phys. Rev. 136, A1445 (1964).

[14] V. Heine and R. O. Jones, J. Phys. C 2, 719 (1969).

[15] J. C. Phillips, Rev. Mod. Phys. 42, 317 (1970).

[16] J. C. Phillips, *Bonds and Bands in Semiconductors* (Academic, New York, 1973), Chap. 2.