Optical Response of Semiconductors in Electric Fields: Excitonic Effects*

Frank C. Weinstein,† John D. Dow, and Binneg Y. Lao‡
Joseph Henry Laboratories of Physics, Princeton University, Princeton, New Jersey 08540
(Received 12 July 1971)

The effects of a uniform electric field on the optical response functions of a semiconductor are calculated, by numerically solving the effective-mass equation for a Wannier exciton. The calculations assume an isotropic effective-mass model of direct transitions at a three-dimensional M_0 threshold. The relationship between the real and imaginary parts of the dielectric function are discussed for the purposes of understanding the physics of electroreflection and developing rules of thumb for interpreting electroreflectivity spectra. The theory is compared with the electroreflectivity data at the direct edge of Ge taken by Handler, Jasperson, and Koeppen, and yields very good agreement. The value of the momentum matrix element extracted from the fit of theory to experiment is $0.35\hbar/a_0$, in excellent agreement with experimentally measured values.

I. INTRODUCTION

The optical properties of semiconductors, and the changes in those properties induced by externally applied forces, have provided some of the most useful and experimentally accessible information about the electronic states of covalent materials.[1,2] In particular, electroabsorption has been of considerable interest since 1958, when Franz[3] and Keldysh[4] performed the first calculations of the effects of a strong electric field on the optical absorption coefficient. The subsequent introduction[5] of phase-sensitive detectors, with their high signal-to-noise ratios, greatly facilitated the measurement of externally induced changes in the optical constants and stimulated considerable experimental effort in electroabsorption, electroreflection, and other types of modulation spectroscopy.[6] However, theories of the optical response of semiconductors in an electric field[7,8] were based primarily on the one-electron Franz-Keldysh theory until 1965, when Duke[9] attempted to include electron-hole scattering, using a model potential. Subsequently, Ralph[10] obtained numerical solutions of the effective-mass equation for an exciton in a uniform electric field. Blossey[11] extended Ralph's theory to include excitonic effects at M_3 as well as M_0 critical points; in addition, he discussed the effects of thermal broadening and calculated differential electroabsorption at the M_0 critical point in lead iodide.[12] Calculations of direct[13] and indirect[14] allowed electroabsorption and "half-forbidden" and "forbidden" electroabsorption,[15] and the effects of electric fields on optical absorption edges,[16] have been performed in this laboratory.

The present work goes beyond the one-electron theory to provide a quantitative treatment of the effects of electron-hole correlations on measured electroreflection spectra,[16a] and presents a comparison with the electroreflectivity data of Handler, Jasperson, and Koeppen.[17] We have considered electroreflection for several reasons. First, electric-field modulation is interesting theoretically since the electric field modifies electron-hole correlations and changes selection rules. Second, electroreflectivity spectra, although more difficult to analyze than electroabsorption, are more readily measured. The data for Ge taken by Handler, Jasperson, and Koeppen[17] are ideal for this purpose: (i) The large number of oscillations in the electroreflectivity spectrum suggests that the electric fields were highly uniform, (ii) the range of photon energies covered, 0.7–1.2 eV, includes the M_0 edge, where the effects of excitons on the line shapes are highly important, (iii) the energy-band structures are well known in this region and the isotropic effective-mass approximation is obeyed quite well, and (iv) the broadening is small and hopefully has only a weak dependence on energy. Thus the experiment of Handler, Jasperson, and Koeppen is a prototype of electroreflectivity measurements on highly excited electronic states—states embedded in a continuous background associated with a lower-energy threshold.

We have undertaken this work to understand the principal effects of excitons on the optical properties of electric-field-perturbed semiconductors. Previous efforts[11–13] to include the quasibound and continuum exciton states in calculations of direct transitions in electric fields have been limited to calculations of electro*absorption*, and have failed to achieve better than qualitative agreement[18] with experiment. Here, we are able to obtain relatively good agreement between calculated and measured differential electro*reflectivity*, indicating that (i) the Elliott exciton theory[19] is capable of describing the small effects measured in modulation spectra and (ii) Handler, Jasperson, and Koeppen[17] have succeeded in obtaining a uniform applied electric field in the absorbing region of their Ge samples.

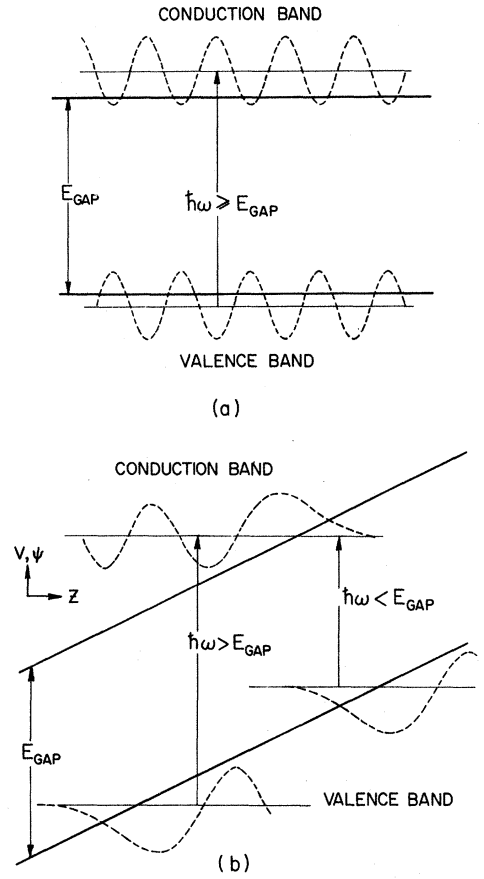


FIG. 1. Energy-band edges $V(z)$ (heavy solid lines) and wave functions $\psi(z)$ (dashed lines) of conduction and valence electrons as a function of position z. (a) With no electric field the band edges are flat. Direct optical transitions are allowed only for photon energies $\hbar\omega \geq E_{gap}$. (b) In an electric field F, the bands are tilted. Transitions now occur at all photon energies. For $\hbar\omega < E_{gap}$ the overlap in the wave functions and hence the transition probability, is exponentially small, while for $\hbar\omega > E_{gap}$ the transition probability is finite and oscillates as a function of photon energy $\hbar\omega$.

The calculation proceeds in several steps. First we calculate the imaginary part of the dielectric response function $\epsilon_2(\omega, F)$, which is proportional to the optical absorption. Next, by using the Kramers-Kronig dispersion relations, we arrive at the real part of the dielectric function $\epsilon_1(\omega, F)$. All the remaining optical response functions are then readily calculated, including the electroreflectivity. We have chosen to calculate an electroreflectivity spectrum for comparison with the measured values $\Delta R/R$. An alternative procedure would have been to obtain a "measured" spectrum for $\Delta\epsilon(w, F)$ by Kramers-Kronig-transforming the data and then comparing theory with experiment. Our principal reason for not following the latter procedure is that electroreflectivity data are often taken only over a limited range of photon energies; thus significant contributions to the dispersion integrals may arise from unmeasured parts of a spectrum. In addition to the differential electroreflectivity, we have also obtained the dependences of the various optical response functions, such as the index of refraction and the absorption coefficient, on the photon energy.

In Sec. II we summarize the principal physical results of the theory of electroabsorption. In Sec. III we discuss the Kramers-Kronig transformation and present several intuitively instructive examples of its effects on typical line shapes. Section IV is devoted to a treatment of the optical response functions of a semiconductor in a uniform field. Section V deals with the comparison of the theory with the electroreflectivity spectrum of Ref. 17; Sec. VI summarizes our results.

II. PHYSICS OF ELECTROABSORPTION

In zero applied field the wave functions of a solid are Bloch functions,

$$\psi_{n\vec{k}} = \psi_{\vec{k}} u_{n\vec{k}}(\vec{r}) , \tag{2.1}$$

where $\psi_{\vec{k}} = V^{-1/2} e^{i\vec{k}\cdot\vec{r}} = V^{-1/2}(\cos\vec{k}\cdot\vec{r} + i\sin\vec{k}\cdot\vec{r})$ is the envelope function and $u_{n\vec{k}}(\vec{r})$ is a periodic function. The (real part of the) envelope function varies sinusoidally in configuration space and is sketched for a typical conduction electron and a somewhat heavier valence electron in Fig. 1(a). The application of a uniform field $\vec{F} = F\hat{z}$ has the effect of accelerating the plane-wave envelope functions, while polarizing the periodic parts of the Bloch functions. In a static gauge ($\vec{A} = 0$, $\phi = -Fz$), the conduction and valence bands are tilted [Fig. 1(b)] and the effective-mass equation for the electron envelope function becomes[20]

$$(p_e^2/2m_e - eFz_e)\psi(\vec{r}_e) = E^e\psi(\vec{r}_e) . \tag{2.2}$$

Here we have assumed parabolic bands and an isotropic effective mass m_e for the electron; the electron's position and momentum operators are denoted by $\vec{r}_e$ and $\vec{p}_e$, respectively. The hole envelope function solves a similar effective-mass equation. Approximate[20] stationary-state wave functions for the electron are products of plane waves and Airy functions, modulated by polarized $u_{n\vec{k}}$ functions:

$$\begin{aligned}\psi_{n\vec{k}}(\vec{r}_e) = {} & V^{-5/12}(2\pi^2\lambda_e)^{1/4}\exp(ik_x x_e + ik_y y_e) \\ & \times \mathrm{Ai}(\lambda_e z_e + \lambda_e E_z^e/eF) u_{n\vec{k}}(\vec{r}_e) \\ & + F\sum_{n'\vec{k}'} M(n\vec{k}, n'\vec{k}') u_{n'\vec{k}'}(\vec{r}_e) . \end{aligned} \tag{2.3}$$

Here $\lambda_e^3 = 2m_e|e|F/\hbar^2$, $E_z^e = E_e - \hbar^2(k_x^2 + k_y^2)/2m_e$, and M is a mixing coefficient involving a matrix element of z, and is usually sufficiently small to be neglected.[21] Concentrating on optical transitions from near the top of the valence band to the

bottom of the conduction band, we find that the imaginary part of the dielectric response function is proportional to the square of the electron-hole overlap (times the density of states),

$$\epsilon_2(\omega) \propto \sum \left| \int_{-\infty}^{\infty} \mathrm{Ai}(\lambda_e z + \lambda_e E_z^e/eF)\,\mathrm{Ai}(\lambda_h z + \lambda_h E_z^h/eF)\,dz \right|^2 \delta(E_e - E_h + E_{\mathrm{gap}} - \hbar\omega) , \tag{2.4}$$

where λ_h is defined as in Eq. (2.3).

The qualitative features of electroabsorption can be deduced from this one-electron picture. As can be seen in Fig. 1(b), the absorption (which depends on the overlap of the wave functions) is finite but exponentially decreasing below the gap and oscillatory above the gap.[22] This can be understood quantitatively by carrying out the integral in Eq. (2.4) using the convolution theorem for Airy functions.[23] The result is an Airy function whose argument depends only on the reduced mass $\mu = (m_e^{-1} + m_h^{-1})^{-1}$ and on the energy difference $E_z^e - E_z^h$. The sum over the final states is readily done and gives the result[24,25]

$$\epsilon_2(\omega) = \frac{2e^2}{m^2\omega^2} |\langle c|\hat{\epsilon}\cdot\vec{p}|v\rangle|^2 \frac{2\mu}{\hbar^2} \lambda \int_{\eta}^{\infty} \mathrm{Ai}^2(t)\,dt , \tag{2.5}$$

where $\langle c|\hat{\epsilon}\cdot\vec{p}|v\rangle$ is the transition matrix element of the component of the momentum along the direction of polarization of the photon, $\hat{\epsilon}$, $\lambda^3 = 2\mu|e|F/\hbar^2$, and $\eta = -(E_{\mathrm{gap}} - \hbar\omega)\lambda/eF$. The integral in Eq. (2.5) is responsible for the frequency dependence deduced from Fig. 1(b): For $\hbar\omega < E_{\mathrm{gap}}$, $\epsilon_2(\omega)$ is exponentially small, while for $\hbar\omega > E_{\mathrm{gap}}$, $\epsilon_2(\omega)$ oscillates.

The effects of excitons are not easily represented on Fig. 1, so we transform from electron and hole coordinates to center-of-mass and relative-motion coordinates of the electron-hole pair. The center of mass has no net charge, and moves as a free particle (plane wave), unaffected by a uniform field. The potential energy for relative motion is the ramp depicted in Fig. 2 and the "exciton" envelope function[26] is

$$\psi_{\vec{K},\nu} \propto V^{-1/2} e^{i\vec{K}\cdot\vec{R}} U_\nu(\vec{r}) , \tag{2.6a}$$

where $\vec{K}$ and $\vec{R}$ are the center-of-mass wave vector and position vector, respectively, $\vec{r}$ is the electron-hole separation, and the relative-motion wave function $U_\nu(\vec{r})$ is given by

$$U_\nu(\vec{r}) = V^{-5/12} (2\pi^2\lambda)^{1/4} e^{i(k_x x + k_y y)} \mathrm{Ai}\left[\lambda z + (\lambda/eF)E_z\right] , \tag{2.6b}$$

just as in Eq. (2.3). Here ν is the set of quantum numbers k_x, k_y, E_z; $\lambda^3 = 2\mu|e|F/\hbar^2$ is defined for reduced electron-hole mass μ. In this "exciton" picture[19,27] the optical absorption is proportional to the density of states times the probability that the electron and hole are in the same unit cell,

$$\epsilon_2(\omega) \propto \sum_\nu |U_\nu(0)|^2 \delta(\hbar\omega - E_{\mathrm{gap}} - E_\nu) . \tag{2.7}$$

Substituting the wave function (2.6b) in Eq. (2.7), we again find

$$\epsilon_2(\omega) \propto \sum_{k_x k_y E_\nu} \mathrm{Ai}^2\left[(\lambda/eF)E_z\right] \times \delta(\hbar\omega - E_{\mathrm{gap}} - E_z - (\hbar^2/2\mu)(k_x^2 + k_y^2)) \tag{2.8}$$

which, when summed, is identical to the one-electron result, Eq. (2.5). In this "exciton" picture the exponential Franz-Keldysh absorption tail and the oscillatory absorption above the gap follow from the relative-motion wave function at zero electron-hole separation,

$$U(0) = V^{-5/12} (2\pi^2\lambda)^{1/4} \mathrm{Ai}\left[(\lambda/eF)E_z\right] . \tag{2.9}$$

The introduction of the electron-hole interaction $-e^2/\epsilon_0 r$ is now possible in the exciton picture (Fig. 2). The attractive Coulomb interaction draws the "particle" toward the origin, causing the relative-motion wave function $U(\vec{r})$ to be enhanced near $r = 0$. Thus the optical absorption, which is proportional to $|U(0)|^2$, is increased by the Coulomb interaction (Fig. 2) both above the gap ($E > 0$) and below (where the enhancement is an exponential function of E). In addition, the Coulomb interaction causes the principal features of the absorption spectrum to be shifted to lower energy and increases the period of the oscillations of $\epsilon_2(\omega)$ for photon energies greater than the band gap.[13] These features are evident in Fig. 3(a), where we have plotted

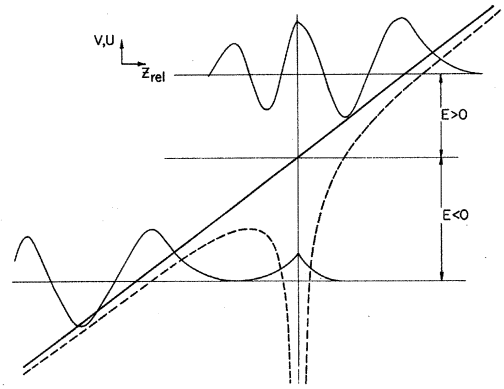
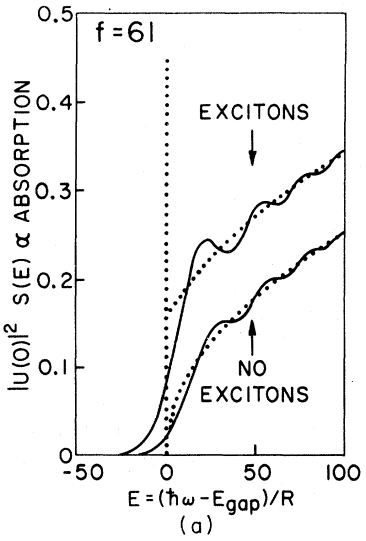


FIG. 2. Potential $V(z_{rel})$ and wave functions $U(z_{rel})$ for relative motion of electron and hole as a function of electron-hole separation z_{rel}. The ramp potential (solid line) due to the applied electric field is modified when the electron-hole Coulomb interaction is included (dashed curve). A "bound" ($E < 0$) and "continuum" ($E > 0$) wave function are plotted. Note the exponential enhancement of $U(0)$ for the "bound" wave functions.



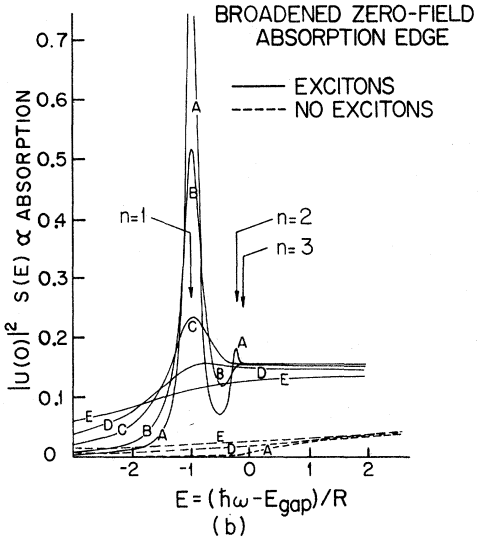


FIG. 3. Absorption strength as a function of photon energy. (a) The solid curves correspond to an applied field appropriate to the heavy hole at the Ge direct edge. The dotted curves are for zero applied field. The exciton curves lie above the one-electron curves for *all* values of photon energy. The dotted spike represents the contribution of the exciton bound states for $-1 \leq E < 0$. The continuum contribution of the zero-field excitons is finite at $E=0$, whereas the free one-electron curve shows the usual square-root threshold behavior. No broadening is included. (b) The zero-field (dotted) curves in (a) are broadened and replotted on an expanded energy scale, for various values of Γ/R. The solid curves correspond to excitons, while the dashed curves are for one-electron theory. $E=-1$ corresponds to the 1*s* exciton state. The labels are as follows: A, $\Gamma/R=0.1$; B, $\Gamma/R=0.2$; C, $\Gamma/R=0.5$; D, $\Gamma/R=1.0$; E, $\Gamma/R=2.0$. Note that the $n=2, 3, \ldots$ exciton states are not resolved for $\Gamma/R>0.1$ and that $\Gamma/R \approx 0.2$ was used in the fit to $\Delta R/R$. Observe the difference between the exciton and one-electron curves.

the absorption strength $|U(0)|^2 S(\hbar\omega - E_{gap})$ [$S(E)$ is the density of states at energy E] as a function of photon energy, for transitions from the heavy-hole band in Ge in a field F of 4.2×10^4 V/cm ($f = |e|Fa/R$ is a dimensionless measure of the strength of the field in exciton Rydbergs per electron per excition Bohr radius; the field in the figure is thus 61 times stronger than the field needed to ionize the excition.) The zero-field unbroadened absorption strengths predicted by the exciton and one-electron theories are also plotted in Fig. 3(a).

If thermal broadening is a small perturbation (i.e., $\Gamma < R$), then the broadening may be included in the theory by convoluting the zero-field absorption-strength curves with a Lorentzian of half-width Γ. This convolution effectively averages the absorption strength over an energy of order Γ and has little effect on the smooth parts of the spectrum; the amplitudes and phases of the oscillations of the $f=61$ curve are reduced and shifted, respectively—but the qualitative aspects are not significantly altered. However, the zero-field spectrum near and below $E \equiv \hbar\omega - E_{gap} = 0$ is significantly affected by the broadening, as may be seen in Fig. 3(b). This region of the spectrum is responsible for the first negative peak of differential absorption and reflectivity spectra. Thus the dominance of excitons and the effects of broadening are most apparent in the first negative peak of an electromodulation experiment, but are generally evident throughout the spectra.

III. KRAMERS-KRONIG RELATIONS

Once the imaginary part of the dielectric response function ϵ_2 is known, the real part ϵ_1 can be determined by the Kramers-Kronig relations. The explicit form of the Kramers-Kronig relations used throughout solid-state physics is

$$\epsilon_1(\omega) - 1 = \frac{2}{\pi} P\int_0^\infty \frac{\omega' \epsilon_2(\omega')}{\omega'^2 - \omega^2}\, d\omega'. \tag{3.1}$$

Here $P\int$ is the Cauchy principal-value integral, and ϵ_1 and ϵ_2 are the real and imaginary parts of the complex dielectric response function.[28] The Kramers-Kronig relations follow directly from the principle of causality, and put very few restrictions on the explicit form of $\epsilon_2(\omega)$. Analogous dispersion relations[29] connect the real and imaginary parts of the complex index of refraction $\tilde{n} = n + ik = (\tilde{\epsilon})^{1/2}$ and the magnitude and phase of the reflectivity at normal incidence

$$R^{1/2} e^{i\theta} = \frac{(n - n_0) + ik}{(n + n_0) + ik}$$

[here n_0 is the (real) refractive index of the medium of incidence]:

$$\theta(\omega) = \frac{\omega}{\pi} P\int_0^\infty \frac{\ln R(\omega')}{\omega'^2 - \omega^2}\, d\omega'. \tag{3.2}$$

This dispersion relation is of particular importance to experimenters who measure the reflectivity and wish to obtain ϵ_1 and ϵ_2.

Because of the linearity of Kramers-Kronig transforms, they are also valid for differential quantities. In particular,

$$\Delta\epsilon_1(\omega) = \frac{2}{\pi} P\int_0^\infty \frac{\omega' \Delta \epsilon_2(\omega')}{\omega'^2 - \omega^2} d\omega' \qquad (3.3)$$

and

$$\Delta\theta(\omega) = \frac{\omega}{\pi} P\int_0^\infty \frac{\Delta R}{R} \frac{1}{\omega'^2 - \omega^2} d\omega'. \qquad (3.4)$$

Here the Δ indicates a difference between the finite-field and zero-field quantities, e.g.,

$$\Delta\epsilon_2(\omega) = \epsilon_2(\omega, F) - \epsilon_2(\omega, F = 0).$$

The Kramers-Kronig transformation is quasilocal in that isolated structures in the imaginary part of a function $g_2(\omega)$ will tend to map into corresponding isolated structures in the real part $g_1(\omega)$. Thus at frequencies "near resonance"[30] we can single out the contribution to the real part g_1 of a particular structure in the imaginary part g_2. We have depicted the real and imaginary parts of some simple functions in Figs. 4–7, in order to get an intuitive feeling for the effect of the Kramers-Kronig transformation. The real and imaginary parts of these functions are listed in Table I. The purpose of these figures is to illustrate a set of rules of thumb by which structure in $\epsilon_1(\omega)$ can be identified with structure in the more easily interpretable $\epsilon_2(\omega)$. Note that a general $\epsilon_2(\omega)$ spectrum can be written as a linear superposition of structures of the sort

TABLE I. Some simple functions and their Kramers-Kronig transforms whose graphs appear as Figs. 4–8. In the entries below, only the positive-frequency parts ($\omega > 0$) $\tilde{g}_2$ and $\tilde{g}_1$ are shown. The complete forms are $g_2(\omega) = \tilde{g}_2(\omega) - \tilde{g}_2(-\omega)$ and $g_1(\omega) = \tilde{g}_1(\omega) + \tilde{g}_1(-\omega)$. The sole exception is in lines 7(a) and 7(b), where the complete form of $\tilde{g}_1(\omega)$ is displayed.

Fig.	$\tilde{g}_2(\omega)$		$\tilde{g}_1(\omega)$	
4(a)	1, 0,	$\lvert\omega-\omega_0\rvert<1$ otherwise	$\frac{1}{\pi}\ln\left\lvert\frac{(\omega-\omega_0)-1}{(\omega-\omega_0)+1}\right\rvert$	
4(b)	-1, 1, 0,	$-1<\omega-\omega_0<0$ $0<\omega-\omega_0<1$ otherwise	$\frac{1}{\pi}\ln\left\lvert\frac{(\omega-\omega_0)^2}{(\omega-\omega_0)^2-1}\right\rvert$	
4(c)	$\frac{1}{2}(\omega-\omega_0+1)$, 0,	$\lvert\omega-\omega_0\rvert<1$ otherwise	$\frac{1}{\pi}\left(1+\frac{\omega+\omega_0+1}{2}\ln\left\lvert\frac{(\omega-\omega_0)-1}{(\omega-\omega_0)+1}\right\rvert\right)$	
4(d)	$(\omega-\omega_0+1)$, $-(\omega-\omega_0-1)$, 0,	$-1<\omega-\omega_0<0$ $0<\omega-\omega_0<1$ otherwise	$\frac{1}{\pi}\ln\left\lvert\frac{(\omega-\omega_0)-1}{(\omega-\omega_0)+1}\right\rvert+\frac{\omega-\omega_0}{\pi}\ln\left\lvert\frac{(\omega-\omega_0)^2}{(\omega-\omega_0)^2-1}\right\rvert$	
4(e)	$2(\omega-\omega_0+1)$, $-2(\omega-\omega_0)$, $2(\omega-\omega_0-1)$, 0,	$-1<\omega-\omega_0<-\frac{1}{2}$ $-\frac{1}{2}<\omega-\omega_0<\frac{1}{2}$ $\frac{1}{2}<\omega-\omega_0<1$ otherwise	$\frac{2}{\pi}\ln\left\lvert\frac{(\omega-\omega_0)^2-\frac{1}{4}}{(\omega-\omega_0)^2-1}\right\rvert-\frac{2}{\pi}(\omega-\omega_0)\ln\left\lvert\frac{(\omega-\omega_0-\frac{1}{2})^2(\omega-\omega_0+1)}{(\omega-\omega_0+\frac{1}{2})^2(\omega-\omega_0-1)}\right\rvert$	
5(a), 5(b)	$\frac{1}{1+(\omega-\omega_0)^2}$, $e^{(\omega-\omega_0)^2/2}$,	(solid) (dashed)	$\frac{\omega-\omega_0}{1+(\omega-\omega_0)^2}$ $(2/\pi)^{1/2}\,e^{-(\omega-\omega_0)^2/2}\int_0^{\omega-\omega_0} e^{t^2}\,dt$	(solid) (dashed)
5(c), 5(d)	$\tilde{g}_2(\omega+\frac{1}{2})\lvert_{5(a)}-\tilde{g}_2(\omega-\frac{1}{2})\lvert_{5(a)}$		$\tilde{g}_1(\omega+\frac{1}{2})\lvert_{5(b)}-\tilde{g}_1(\omega-\frac{1}{2})\lvert_{5(b)}$	
6(a)	$(\cos\omega)/(\omega^2+25)$		no closed form	
6(b)	$e^{-\omega/15}\cos\omega$		no closed form	
7(a)	$(\omega_g^{3/2}/\omega^2)(\omega-\omega_g)^{1/2}$, 0,	$\omega>\omega_g$ $\omega<\omega_g$	$(\omega_g^{3/2}/\omega^2)[2\omega_g^{1/2}-(\omega+\omega_g)^{1/2}]$, $(\omega_g^{3/2}/\omega^2)[2\omega_g^{1/2}-(\omega+\omega_g)^{1/2}-(\omega_g-\omega)^{1/2}]$,	$\omega>\omega_g$ $\omega<\omega_g$
7(b)	$(\omega_g^{1/2}/\omega^2)(\omega-\omega_g)^{3/2}$, 0,	$\omega>\omega_g$ $\omega<\omega_g$	$(-\omega_g^{1/2}/\omega^2)[2\omega_g^{3/2}-(\omega+\omega_g)^{3/2}]$, $(-\omega_g^{1/2}/\omega^2)[2\omega_g^{3/2}-(\omega+\omega_g)^{3/2}-(\omega_g-\omega)^{3/2}]$,	$\omega>\omega_g$ $\omega<\omega_g$
8	$\frac{\omega-\omega_0}{(\omega-\omega_0)^2+S^2}$, 0,	$\lvert\omega-\omega_0\rvert<\Delta$ otherwise	$\frac{S}{(\omega-\omega_0)^2+S^2}\frac{2}{\pi}\arctan\frac{\Delta}{S}$ $+\frac{\omega-\omega_0}{(\omega-\omega_0)^2+S^2}\frac{1}{\pi}\ln\left\lvert\frac{(\omega-\omega_0)-\Delta}{(\omega-\omega_0)+\Delta}\right\rvert$	

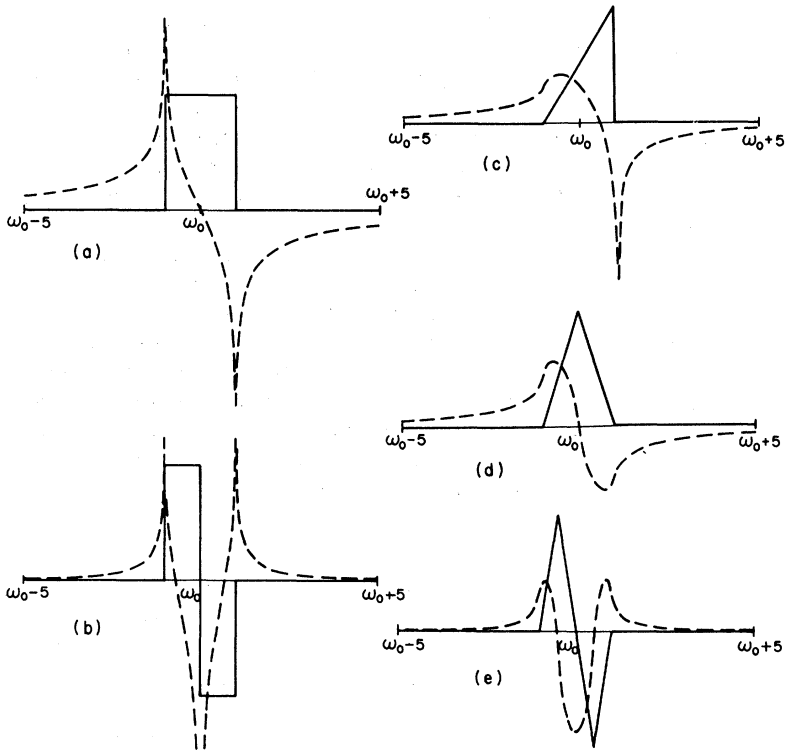


FIG. 4. Some simple functions and their Kramers-Kronig transforms. The dashed curves are the transforms of the solid curves.

considered in Table I. The approximate rules of thumb which can be deduced from these spectra are as follows: (i) Extrema in $\epsilon_2(\omega)$ correspond to zeros in $\epsilon_1(\omega)$ and vice versa; (ii) if $\epsilon_2(\omega)$ is nearly symmetric about a particular frequency ω_0, then $\epsilon_1(\omega)$ is antisymmetric about the same frequency; (iii) an abrupt rise in ϵ_2 is transformed into a high peak in ϵ_1—the height of the peak increases with the slope of ϵ_2; and (iv) a rapid variation in the slope of ϵ_2 leads to an inflection point in ϵ_1.

These rules are illustrated in Figs. 4–7. Figure 4(a) shows a rectangular pulse and its transform. Note the symmetry, the presence of the logarithmic singularity in g_1 arising from the infinite slope in g_2, and the slow falloff of g_1 with frequency. Such a rectangular threshold behavior might be associated with an M_0 or M_2 two-dimensional critical point.[31] Figure 4(b) illustrates a square wave, which is the difference between two displaced pulses like that in Fig. 4(a). This difference spectrum has the same general features of Fig. 4(a) but contains additional structure associated with the derivative (or, more precisely, the difference) and falls off more rapidly for large ω. A more slowly varying threshold is exhibited in Figs. 4(c) and 4(d); note that the inflection points in g_1 correspond to discontinuities in the slope of g_2, that the reduced slope of g_2 is manifested as a reduced height of g_1, and that the symmetry rule (ii) is obeyed. Figure 4(e) presents a symmetrical triangular wave, which has the general features of Figs. 4(c) and 4(d) as well as the more rapid falloff associated with difference spectra. Note that all the curves obey the rule that zeros in g_2 correspond to extrema in g_1. The rule is obeyed by nearly all differential spectra, both theoretical and experimental.[32]

The contrast between spectra and difference spectra is clearly shown in Fig. 5. In Fig. 5(a), we show a Lorentzian of unit width as the solid curve and a Gaussian of unit width as the dashed curve; both are centered at $\omega=\omega_0$. In Fig. 5(b), their corresponding Kramers-Kronig transforms are shown, using the same vertical scale as in Fig. 5(a). Note that while the Gaussian has much sharper tails than the Lorentzian, both their transforms have the same asymptotic "near-resonance" behavior,[30] $\sim(\omega-\omega_0)^{-1}$.[33] In Fig. 5(c), we have plotted the difference between two such Lorentzians displaced from one another by one unit (solid), and the difference between two displaced Gaussians (dashed). In Fig. 5(d), their corresponding transforms are shown. As we have seen before, the tails of these differential quantities are much sharper than their counterparts in Fig. 5(a) and 5(b). In addition, the complementarity between extrema in Fig. 5(c) and zeros in Fig. 5(d) is particularly apparent.

This complementarity rule is especially clear in Fig. 6, where we have plotted two different decaying oscillatory functions and their Kramers-Kronig transforms. In Fig. 6(a), the amplitude decay is

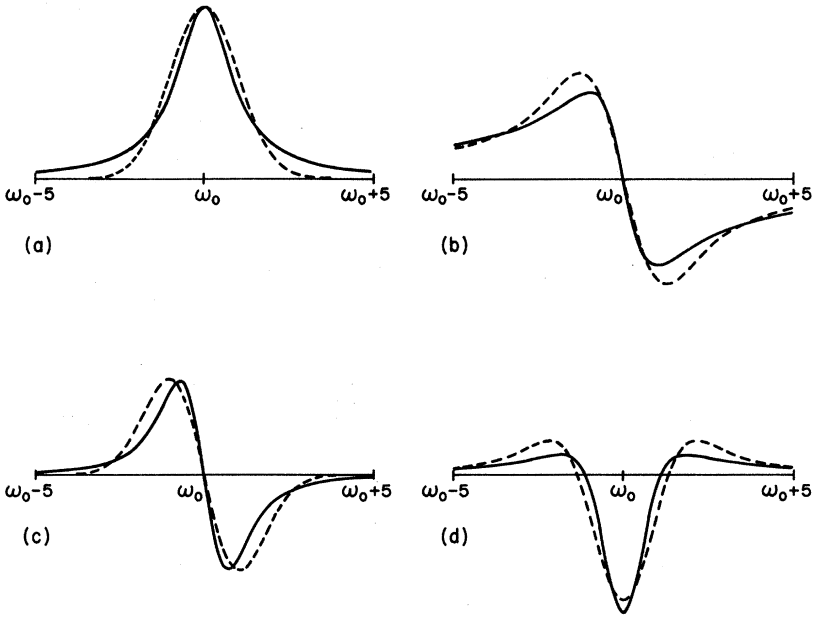


FIG. 5. Some simple absolute and differential spectra and their Kramers-Kronig transforms. The solid curves are based on Lorentzian line shapes, while the dashed curves are Gaussians. (a) The basic curves, (b) their Kramers-Kronig transform, (c) difference between two curves of type (a) displaced relative to one another by one unit width, and (d) Kramers-Kronig transform of (c).

Lorentzian, while in Fig. 6(b) it is exponential (see Table I).[34] Unlike the functions given in Figs. 4 and 5, the Kramers-Kronig transforms are not expressible in terms of simple functions. Rather, they have been obtained by numerical integration of the Kramers-Kronig relation. Note that the transforms g_1 look like phase-shifted versions of the respective g_2's, where the shift is approximately $\pi/2$ (at least asymptotically). Thus the minima and maxima in g_1 correspond to the zeros in g_2, as above. We find this behavior again, in the case of electroabsorption spectra[25] (in Ref. 25 the Coulomb interaction is neglected).

We next apply these results to some simple physical situations. It is elementary to show that for allowed band-to-band transitions at a direct edge, neglecting excitons, the imaginary part of the dielectric constant has a square-root dependence on ω. $\epsilon_2(\omega)$ is given in Table I, line 7(a), where $\hbar\omega_g = E_{gap}$ is the bandgap energy; this result assumes parabolic bands and a constant transition matrix element, and is evaluated for an M_0 critical point. We do the Kramers-Kronig integral to obtain[35] $\epsilon_1(\omega) - 1$, which also has a square-root dependence, as shown in Table I. In a similar way, $\epsilon_2(\omega)$ and $\epsilon_1(\omega) - 1$ for the "forbidden"[36] transitions in the same approximation (the matrix element is assumed to be a linear function of the electron's wave vector) vary as the $\frac{3}{2}$ power of frequency, as may be seen in Table I, line 7(b). These four functions have been plotted in Figs. 7(a) and 7(b). Note in Fig. 7(a) that the sharp rise in slope in ϵ_2 at ω_g appears in $\epsilon_1 - 1$ as a sharp peak, in agreement with our rules of thumb. In Fig. 7(b), for the forbidden case, the discontinuity in slope in ϵ_2 at $\omega = \omega_g$ shows up as an inflection point in $\epsilon_1 - 1$. Although the scale in Figs. 7(a) and 7(b) is the same, the magnitude of ϵ_2, say, for the "forbidden" case will be smaller than that of

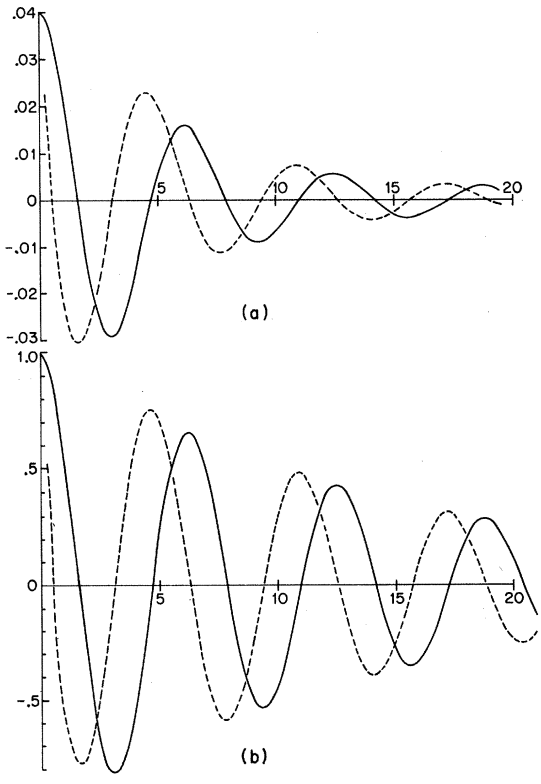


FIG. 6. Oscillatory functions and their Kramers-Kronig transforms. The dashed curves are the transforms of the solid curves. The amplitude of the solid curve is given by $(\omega^2+25)^{-1}$ in (a) and by $e^{-\omega/15}$ in (b).

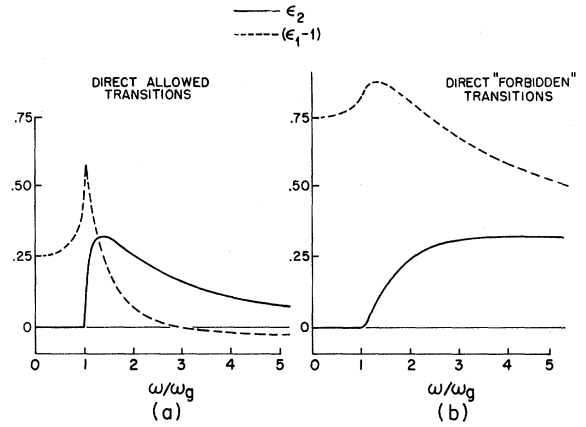


FIG. 7. Dielectric constant ϵ_2 and its Kramers-Kronig transform $\epsilon_1 - 1$ for direct transitions based on one-electron theory. Allowed transitions are shown in (a) and "forbidden" transitions in (b). $\hbar\omega_g$ is the energy gap.

the allowed case by a factor of order $(ka_L)^2$, typically about 10^{-3}, where k is the relative-motion wave vector and a_L is the lattice spacing. This picture of direct absorption concentrates on the region of the Brillouin zone near an M_0 critical point, assumes that the transition matrix element is independent of energy, and neglects deviations of the energy bands from the model of constant isotropic effective masses. At higher energies, other transitions associated with different critical points in the transition band structure [which have not been included in Figs. 7(a) and (b)] will normally mask the contribution of the direct edge.

The indirect transitions are not suited to such a straightforward analysis. It can be shown[37] that the ω dependence of ϵ_2 for indirect transitions (phonon assisted) is given by

$$\epsilon_2 \propto \frac{(\omega - \omega_g \mp \omega_{\text{phon}})^2}{\omega^2}, \quad \omega > \omega_g \pm \omega_{\text{phon}}$$
$$\propto 0, \quad \omega < \omega_g \pm \omega_{\text{phon}}. \tag{3.5}$$

In this expression ω_{phon} is the energy of a phonon, the $\pm$ signs corresponding to emission and absorption, respectively. Since for large energy ($\omega \to \infty$), ϵ_2 approaches a constant, we expect the simple Kramers-Kronig analysis used in Fig. 7 to give a logarithmic divergence at large energies. Hence, without a further assumption on the energy dependence of the matrix element we can proceed no further.

When we include electron-hole Coulomb interactions, we find the following modifications. The optical absorption (and hence ϵ_2) now contains an additional factor of $|U(0)|^2$, the probability that the electron and hole are in the same unit cell. (Although the Coulomb interaction severely affects the form of the wave function, the free-electron wave functions becoming Coulomb waves or bound-state hydrogenic wave functions, and the density of continuum states remains the same as in the one-electron case.[27]) Thus ϵ_2 has a discrete part ($\omega < \omega_g$) consisting of a series of δ functions, which corresponds to the bound states, and a continuum portion ($\omega > \omega_g$) given by

$$\epsilon_2 \propto \frac{1}{2\pi} \frac{1}{1 - e^{-2\pi/\sqrt{E}}}, \tag{3.6a}$$

where

$$E = (\hbar\omega - E_{\text{gap}})/R, \tag{3.6b}$$

$$R = \mu e^4/2\epsilon_0^2\hbar^2 \tag{3.6c}$$

is the excitonic Rydberg, μ is the reduced mass of the exciton, and ϵ_0 is the static dielectric constant. For *very* large ω, ϵ_2 tends to the one-electron value, so that we can perform the Kramers-Kronig analysis and be assured of convergence. We note that at the band gap, ϵ_2 is a constant and continuous; thus the discontinuity in slope which occurred in the one-electron case (see Fig. 7) is removed in $\epsilon_1 - 1$ as well. Furthermore, the presence of the δ functions in ϵ_2 at $\omega_n = \omega_g - R/\hbar n^2$ leads to a set of violent oscillations in $\epsilon_1 - 1$, since $\epsilon_1 - 1 \propto (\omega_n^2 - \omega^2)^{-1}$ if $\epsilon_2 \propto (\omega - \omega_n)$. Hence we must include broadening in order to describe these curves in a meaningful way. Recall that the Lorentzian-broadened zero-field ϵ_2 curves, with and without excitons, are shown in Fig. 3(b). Similar considerations apply to the cases of direct "forbidden" and indirect transitions when excitons are included. That is, the Coulomb interaction strongly modifies the absorption over a large region around the energy gap, with a similar effect on the Kramers-Kronig transform.

Finally, before leaving our discussion of the Kramers-Kronig transformation, it will be instructive to illustrate the effects of truncating the limits on the integral at $\omega = \omega_0 \pm \Delta$. Our choice, a Lorentzian line shape of width S, is an extreme case since it falls off very slowly for large ω; it is amenable to simple calculation, however (see Table I, line 8). The Kramers-Kronig integral is readily performed, yielding the result quoted in Table I. In the limit of no truncation ($\Delta \to \infty$), we recover the expected result, a Lorentzian of width S. In Fig. 8, we have plotted $g_1(\omega)$ for several values of Δ. We note the striking difference between $g_1(\omega)$ for finite values of Δ and the $\Delta = \infty$ case. The logarithmic singularities and the fact that $g_1(\omega)$ is negative for most values of ω may be surprising. Also note that the maximum value (at $\omega = \omega_0$) is strongly affected by the cutoff. Thus, for example, if we represent a typical peak in a $\Delta R/R$ spectrum by a Lorentzian of half-width $S = 10$ meV, then by truncating the Kramers-Kronig integration at $20S = 0.2$ eV above the peak, we expect the error in the peak value of the Kramers-Kronig transform

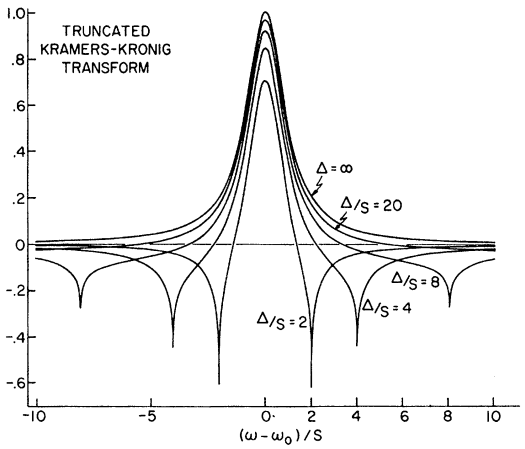


FIG. 8. Result of truncating Kramers-Kronig integral of Lorentzian line shape of width S, for various values of cutoff Δ.

to be about 4% and, more importantly, the detailed behavior of the tails will be inaccurate even at energies as low as 50 meV above the peak. This shows that considerable care must be taken to adequately approximate the tails of a spectrum whose Kramers-Kronig transform is sought. Failure to do so may result in significant errors and may introduce spurious features.

IV. OPTICAL RESPONSE FUNCTIONS

Before calculating the differential electroreflectivity, $\Delta R/R$, it is necessary to know the values of $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ in order to evaluate the coefficients of $\Delta\epsilon_1$ and $\Delta\epsilon_2$. Since the reflectivity at normal incidence is

$$R=\frac{(n-n_0)^2+k^2}{(n+n_0)^2+k^2}\,, \tag{4.1}$$

where n and k are the real and imaginary parts of the complex refractive index of the medium and n_0 is the (real) refractive index of the medium of incidence, one can show by differentiation[38] that

$$\frac{\Delta R}{R}=\frac{\partial \ln R}{\partial \epsilon_1}\,\Delta\epsilon_1+\frac{\partial \ln R}{\partial \epsilon_2}\,\Delta\epsilon_2\,, \tag{4.2}$$

where

$$c_1=c[(\epsilon_1-n_0^2)A_+ +\epsilon_2 A_-]\,,$$
$$c_2=c'[(\epsilon_1-n_0^2)/A_+ -\epsilon_2/A_-]\,,$$
$$c=n_0/[(\epsilon_1-n_0^2)^2+\epsilon_2^2]\,,$$
$$c'=2\epsilon_2 c/|\epsilon|^2\,,$$
$$A_\pm=\pm 2^{1/2}[|\epsilon|\pm\epsilon_1]^{1/2}/|\epsilon|\,.$$

In deriving this symmetrical form, use was made of the usual relations

$$|\epsilon|=(\epsilon_1^2+\epsilon_2^2)^{1/2}\,,\quad n=\left(\frac{|\epsilon|+\epsilon_1}{2}\right)^{1/2}\,,$$

$$k=\left(\frac{|\epsilon|-\epsilon_1}{2}\right)^{1/2}\,, \tag{4.3}$$

which are the solutions of $\epsilon_1=n^2-k^2$ and $\epsilon_2=2nk$. This relation (4.2), together with an analogous one for $\Delta\theta$, the variable conjugate to $\Delta R/R$, can be inverted to yield $\Delta\epsilon_1$ and $\Delta\epsilon_2$. {Experimentally $\Delta R/R$ is measured and then $\Delta\theta$ is obtained by a Kramers-Kronig integration [see Eq. (3.4) and Ref. 1]}.

The optical functions were calculated as follows. The value of [16]

$$\epsilon_2(\omega,F)=\frac{4\pi^2e^2}{m^2\omega^2}\,|\langle c|\hat{\epsilon}\cdot\vec{p}|v\rangle|^2 \times|U(0)|^2 S(\hbar\omega-E_{\rm gap}) \tag{4.4}$$

was transformed using the Kramers-Kronig integral to yield $\epsilon_1(\omega,F)$. Since the above expression for ϵ_2 only includes the contribution due to direct transitions at the fundamental edge, the remaining oscillator strength of all the higher-energy transitions was simulated by adding a term $G\delta(\omega-\omega_A)$ to ϵ_2. The value $\hbar\omega_A=4$ eV was arbitrarily chosen since the maximum value of the optical absorption of Ge occurs near this phonon energy. The strength G was then adjusted so that the static dielectric constant acquired its correct value $\epsilon_1(0)=\epsilon_0=15.8$.[39] Typically the direct-edge contributions to ϵ_1 amounted to less than 2% of the total.

The remaining optical constants (n, k, α, and R) were calculated from ϵ_1 and ϵ_2 using Eqs. (4.1), and (4.3) and the relation between the absorption coefficients α and k, i.e.,

$$\alpha=(2\omega/c)\,k\,. \tag{4.5}$$

These functions are plotted in Figs. 9(a)—9(f). Note the different scales on the ordinates of these curves. The curves corresponding to the dispersive parts of optical constants, viz., ϵ_1, n, and R, are primarily given by the simple curves

$$\epsilon_1=1+\frac{2}{\pi}\,G\,\frac{\omega_A}{\omega_A^2-\omega^2}\,,$$
$$n=\epsilon_1^{1/2}\,,\quad R=\left(\frac{n-n_0}{n+n_0}\right)^2\,. \tag{4.6}$$

We have taken $n_0=1.33$, the index of refraction of water, since $\Delta R/R$ for the data we shall examine was measured in an electrolytic bath. Upon these curves are superimposed the (small) direct-edge contributions. In contrast to the dispersive parts, the absorptive parts of the optical constants ϵ_2, k, and α are completely determined by the direct transition, since the only other source of oscillator strength is the δ function at $\hbar\omega_A=4$ eV, which lies far to the right on this energy scale. Note too that these curves are quite smooth in the vicinity

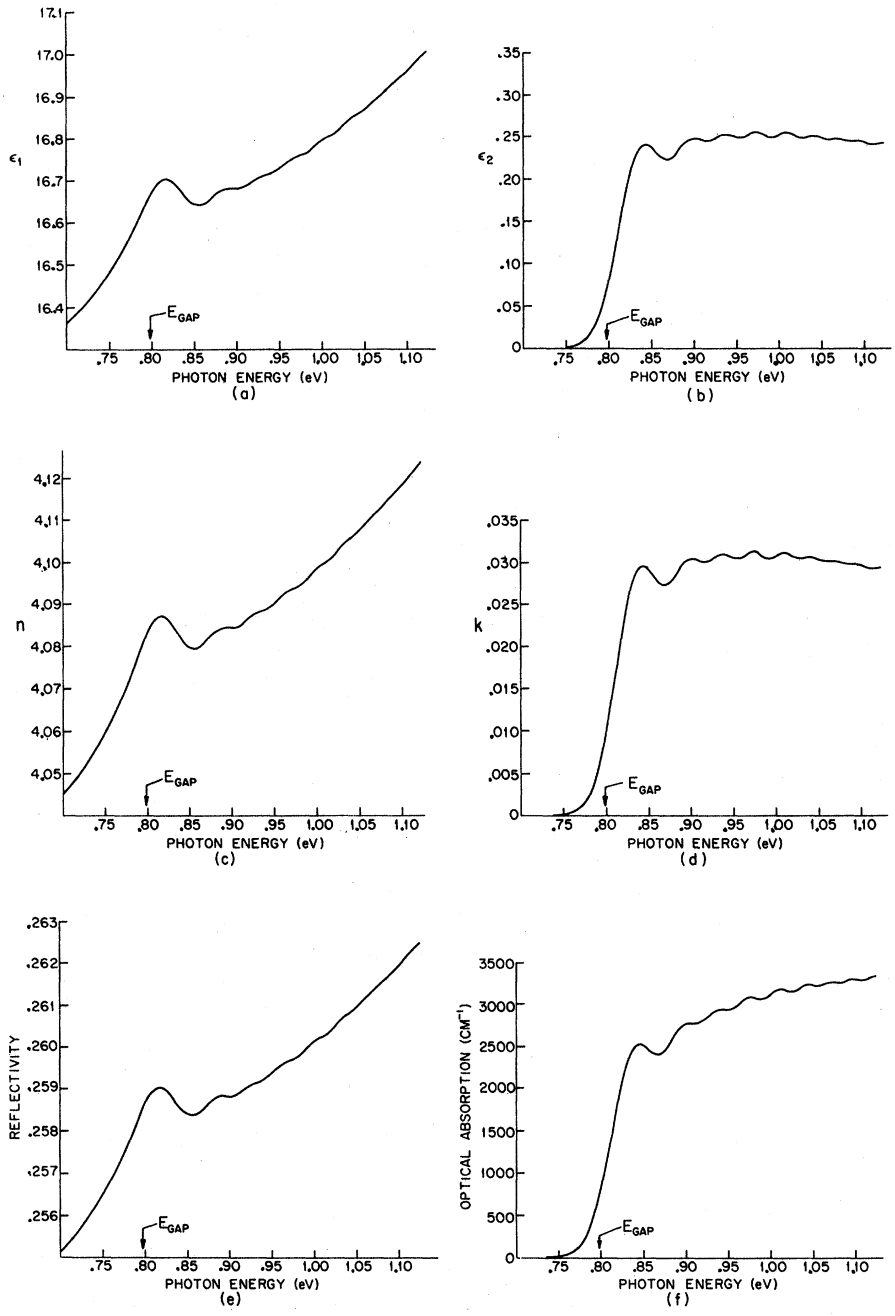


FIG. 9. Complete optical response functions as a function of photon energy. (a) Real part of the dielectric function ϵ_1. (b) Imaginary part of the dielectric function ϵ_2. (c) Real part of the index of refraction n. (d) Imaginary part of the index of refraction k. (e) Reflectivity R. (f) Optical absorption α (in cm^{-1}).

of the gap since the electric fields and the thermal broadening have smoothed (ionized) the discrete exciton lines which would have been present in the zero-field low-temperature case. The absolute magnitude of α agrees well with accepted experimental values.[27]

In Figs. 10(a) and 10(b), the coefficients $c_1(\omega)$ and $c_2(\omega)$ are plotted [see Eq. (4.2)]. Note particularly the difference in the two vertical scales. As has been noted by others,[17] the *principal contribution to* $\Delta R/R$ *near the direct edge of* Ge *comes from* $\Delta\epsilon_1$.

V. COMPARISON WITH EXPERIMENT

The differential optical response functions were calculated as follows. The imaginary part of the dielectric function $\epsilon_2(\omega, F, \Gamma = 0)$ [given in Eq. (4.4)] was evaluated using programs described elsewhere.[13,16] The zero-field[19] value of $\epsilon_2(\omega, F = 0, \Gamma = 0)$ was subtracted from Eq. (4.4) to give $\Delta\epsilon_2(\omega, F, \Gamma = 0)$ and the difference convoluted with a Lorentzian to give $\Delta\epsilon_2(\omega, F, \Gamma)$. The Kramers-Kronig integration was then carried out to give

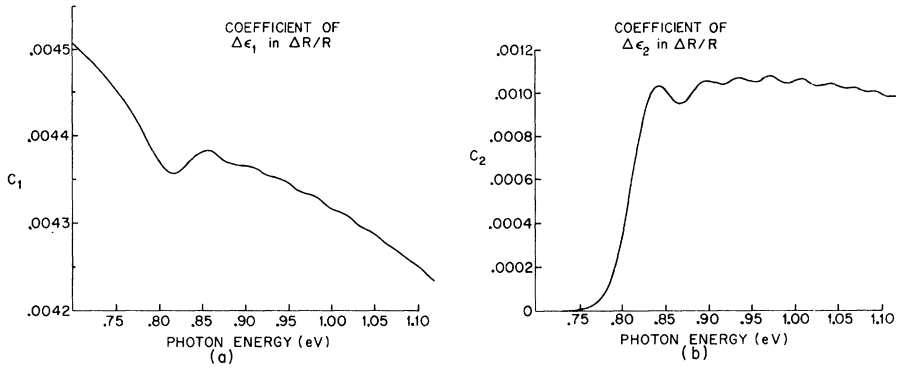


FIG. 10. Coefficients of the differential dielectric constants in the differential electroreflectivity $\Delta R/R = c_1\Delta\epsilon_1 + c_2\Delta\epsilon_2$ as a function of photon energy. c_1 is shown in (a), c_2 in (b). Note the difference in the two vertical scales.

$\Delta\epsilon_1(\omega, F, \Gamma)$. For a detailed discussion of these routines, see the Appendix. These calculations were carried out independently for the light- and heavy-hole valence bands, and the contributions added.[40] In Figs. 11(a) and 11(b) we have plotted the contributions of the light and heavy holes to $\Delta\epsilon_2$ and $\Delta\epsilon_1$, respectively. The dominant features of these two figures are the following: (i) The light hole contributes roughly 30% of the total, (ii) the period of the oscillations in energy of the light hole is longer than that due to the heavy hole by the ratio $(m_h/m_l)^{1/3}$ in the asymptotic limit $\hbar\omega - E_{gap} > 50R$, and (iii) the two contributions are of opposite sign at energies of 0.1 – 0.15 eV above the gap, leading to "beats." In Fig. 11(c), $\Delta\epsilon_1$ and $\Delta\epsilon_2$ are replotted on one graph (only the sum of the light- and heavy-hole contributions is plotted). Many of the general features of Kramers-Kronig transforms discussed in Sec. III are evident here. We note that the two functions are of the same general magnitude, that zeros in one curve tend to correspond to extrema in the other, and vice versa, and that while $\Delta\epsilon_2$ decreases to zero very quickly for photon energies below the gap, $\Delta\epsilon_1$ falls off much more slowly, as we have seen in Figs. 4 and 5.

In Fig. 12(a), $\Delta\epsilon_2$ is plotted for both the exciton theory (dashed line) and the Franz-Keldysh theory (dotted line). Several points are immediately apparent. (i) The magnitude of $\Delta\epsilon_2$ including excitons is approximately twice as large as that of the Franz-Keldysh theory. (This difference in magnitude is particularly pronounced in the region near and below the absorption edge.) Thus, in extracting a quantitative value of the square of the constant transition matrix element by comparing with experiment, one would obtain in the Franz-

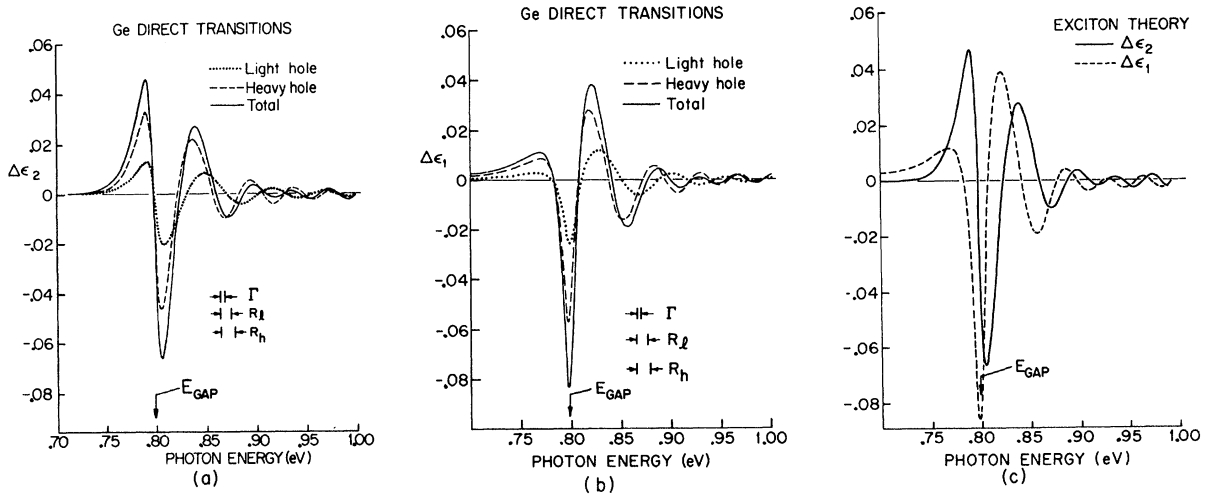


FIG. 11. Differential dielectric function as a function of photon energy, calculated from exciton theory. Broadening of $\Gamma = 3$ meV has been included. The light-hole (dotted) and heavy-hole (dashed) contributions are summed to give the solid curve. The imaginary part $\Delta\epsilon_2$ is shown in (a), while its Kramers-Kronig transform $\Delta\epsilon_1$ is shown in (b). The exciton Rydbergs for the light and heavy holes (R_l, R_h) and the broadening width Γ are also shown. (c) The sum of light- and heavy-hole contributions from (a) and (b) replotted as one graph. The solid curve is $\Delta\epsilon_2$, while the dashed curve is $\Delta\epsilon_1$.

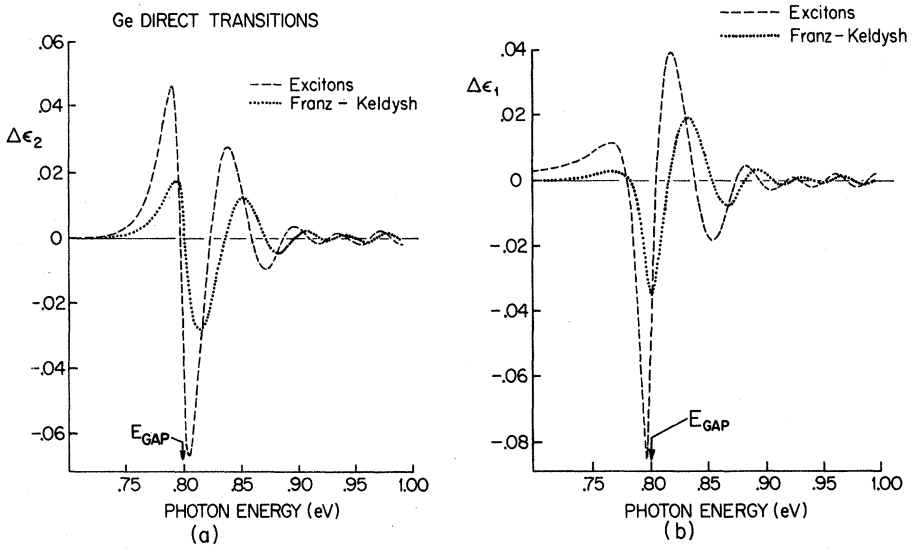


FIG. 12. Differential dielectric function calculated from exciton theory (dashed) and from one-electron Franz-Keldysh theory (dotted) as a function of photon energy. The imaginary part $\Delta\epsilon_2$ is shown in (a), the real part $\Delta\epsilon_1$ in (b). The momentum matrix element is taken to be $0.352\hbar/a_0$ in each case. Note the difference in the amplitudes.

Keldysh theory a value nearly 1.4 that predicted by the exciton theory. (ii) The curve including excitons lies at lower energies than the one-electron result; this is especially evident near the band edge, but persists to high energy. It is important to account for this shift of the spectrum toward lower energy when determining accurate values of the direct energy gap. (iii) The first negative peak, which corresponds to the broadened zero-field excitons in the difference $\Delta\epsilon_2$, is somewhat narrower and sharper than that of the Franz-Keldysh curve. Inasmuch as the dependence of ϵ_2, with excitons included, is rather smooth in this region, this peak is a sensitive indicator of the importance of the electron-hole interaction. In a similar fashion we have plotted $\Delta\epsilon_1$ in Fig. 12(b); the dashed line represents the exciton theory and the dotted line the one-electron theory. The three excitonic features of the $\Delta\epsilon_2$ curve are also manifested here: (i) a larger magnitude, (ii) a shift to lower energy, and (iii) a large negative peak. The large negative peak is due to the abrupt absorption threshold for excitons which, for large Γ, gives an abrupt step-function-like threshold compared with one-electron theory's soft square-root threshold [see Fig. 3(b)]. According to the rules of thumb for Kramers-Kronig transforms, this more abrupt absorption threshold in $\epsilon_2(\omega, F=0, \Gamma)$ for excitons causes $\epsilon_1(\omega, F=0, \Gamma)$ to be larger, contributing a larger negative peak to $\Delta\epsilon_1$. In addition, note that the exciton effects are especially pronounced below the band edge, where $\Delta\epsilon_1$ is considerably larger than that of the one-electron theory.

With these values of $\Delta\epsilon_1$ and $\Delta\epsilon_2$ it is possible to evaluate the differential electroreflectivity $\Delta R/R$ (see Sec. IV). In Fig. 13, the calculated values of $\Delta R/R(\omega)$ have been plotted (dashed line) and compared with the experimental results of Handler, Jasperson, and Koeppen[17] (solid line) and their best fit based on the one-electron Franz-Keldysh theory (dotted line). In Fig. 14, these results have been plotted semilogarithmically.

The following fitting scheme has been used. The contributions to $\Delta R/R$ from both the light and heavy holes have been added separately and by a least-squares calculation the best coefficients of these contributions have been determined. Thus, in effect, the momentum matrix elements are deter-

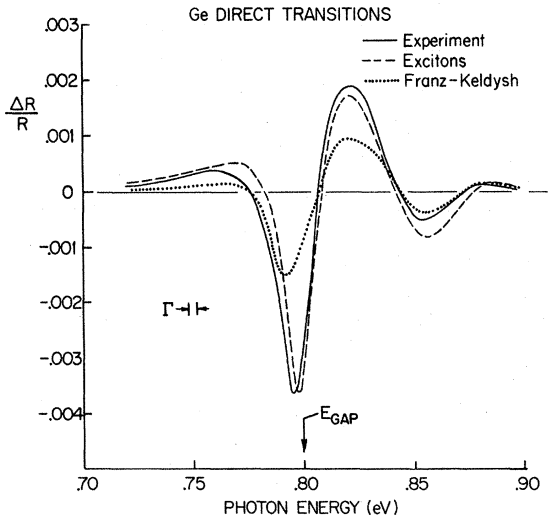


FIG. 13. The differential electroreflectivity as a function of photon energy. The solid line is the experimental result of Handler, Jasperson, and Koeppen (Ref. 17). The dashed curve is the best fit from exciton theory, while the dotted curve is the best fit from one-electron theory found by the authors of Ref. 17.

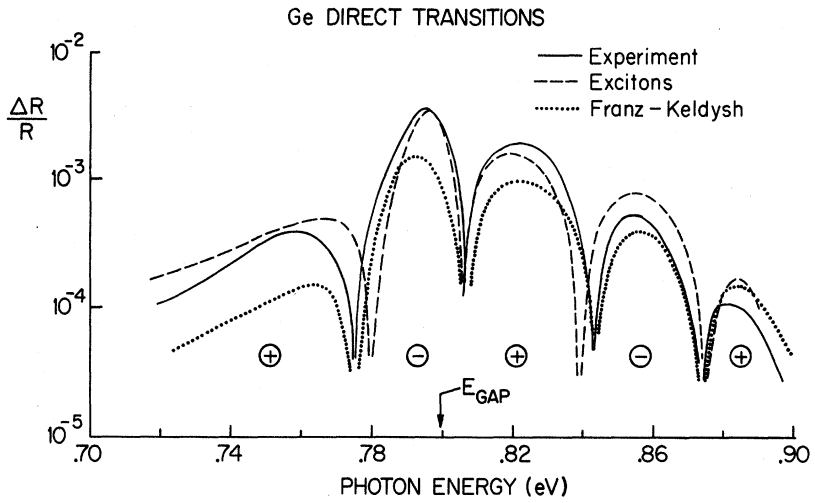


FIG. 14. Same data as in Fig. 13 replotted on semilogarithmic axes. The sign of each peak is indicated in a circle underneath it.

mined independently for the light and heavy holes. The matrix elements are the only free parameters available since the effective masses and field strengths are given experimentally. For a complete listing of the various parameters used see Table II. In this way it was found that the magnitude of the heavy- and light-hole momentum elements are $0.354\hbar/a_0$ and $0.350\hbar/a_0$, respectively, where a_0 is the Bohr radius of hydrogen. These values are to be compared with numbers given by Aspnes and Frova[41] of (i) $0.37\hbar/a_0$ (obtained from optical absorption data[42]), (ii) $0.36\hbar/a_0$ (extracted from Vrehen's data[43] on the strain-split heavy-hole exciton), and (iii) $0.39\hbar/a_0$ (predicted by $\vec{k}\cdot\vec{p}$ perturbation theory[44]). This remarkable agreement between the present theory and experiment is convincing proof of the importance of exciton correlations even in strong-field electroreflectivity experiments, a fact which has been suspected for some time.[41]

At this point let us contrast our fitting procedure with that of Handler, Jasperson, and Koeppen.[17] Those authors used one-electron theory and chose their fitting constants for a best fit in the region of beats between the light and heavy holes, which occurs at energies 150 meV above the gap. This fit accurately describes the zeros of the $\Delta R/R$ spectrum, and gives a large difference (more than 100%)[44a] in the relative contributions of the light and heavy holes, but they have not published an absolute value for the transition matrix element. In contrast, the present fit emphasizes the band edge and first few peaks, gives a good fit in this region of the spectrum, and predicts equal matrix elements (within 2%) whose magnitude accords well with accepted experimental values[41] and with the predictions of $\vec{k}\cdot\vec{p}$ theory.[44] Moreover, the effective-mass theory of direct optical transitions is most applicable near the edge region where the bands are most parabolic, the effective masses are most isotropic, and the broadening and its variation with energy are relatively small.

The present fit of exciton theory to experiment satisfies the criteria for a good fit of Ref. 13. However, a fit of the Franz-Keldysh theory to experiment using our fitting procedure is deficient in three ways: (i) The matrix element is twice as large as the accepted value; (ii) the contributions of the electron and hole have opposite signs; and (iii) the deviations from experiment are five times greater than in the exciton theory. This is further evidence of the inadequacy of the one-electron theory.

The sole deficiency of the present work is that

TABLE II. Parameters used in calculating fits to data, Figs. 13 and 14: μ_e and μ_h, exciton reduced masses for light- and heavy-hole valence bands; F, electric field; f_e and f_h, electric field in exciton Rydbergs per electron exciton Bohr radius; Γ, broadening; E_{gap}, direct band gap; $\langle c|\hat{\epsilon}\cdot\vec{p}|vl\rangle$ and $\langle c|\hat{\epsilon}\cdot\vec{p}|vh\rangle$, momentum matrix elements; a_0, Bohr radius of hydrogen.

	Ref. 17 (without excitons)	This work (with excitons)
μ_e	$0.022m_0$	$0.0195m_0$
μ_h	$0.038m_0$	$0.033m_0$
F	4.2×10^4 V/cm	4.2×10^4 V/cm
f_e	174	174
f_h	61	61
Γ	3 meV	3 meV
E_{gap}	0.79	0.798
$\langle c\|\hat{\epsilon}\cdot\vec{p}\|vh\rangle$	...	$0.354\hbar/a_0$
$\langle c\|\hat{\epsilon}\cdot\vec{p}\|vl\rangle$	...	$0.350\hbar/a_0$
$\|\langle c\|\hat{\epsilon}\cdot\vec{p}\|vl\rangle/\langle c\|\hat{\epsilon}\cdot\vec{p}\|vh\rangle\|^2$ (001 polarization)	2.4[a]	0.98

[a]Reference 44a.

it fails to fit the nodes in the "beat" region of the Ge electroreflectivity spectrum well. However, based on our experience with these calculations, we believe that these beats could be described faithfully by the theory if broadening were introduced as a free parameter (possibly energy dependent), and if the energy dependence of the exciton reduced masses were taken into account. Let us emphasize that, in contrast with the work of Ref. 17, our fit is not based on a least-squares variation of the parameters (except for the transition matrix elements), because of the expense of the calculations. The fact that we have obtained a relatively good fit using the parameters of Handler, Jasperson, and Koeppen[17]—numbers obtained by optimizing the agreement between one-electron theory and experiment—demonstrates that one-electron theory, when fitted to the higher-energy oscillations of electroreflectivity spectra, yields good estimates of effective masses μ and broadening Γ, but poor values of the transition matrix element and the band gap. Fortunately the latter two quantities are scale factors which do not enter into the evaluation of the broadened absorption strength $|U(0)|^2 S(E)$ and are quite easily determined by merely sliding the theoretical $\Delta R/R$ spectrum (plotted semilogarithmically) over the experimental spectrum. Finally, with the rapid interpolator recently developed,[45] it will soon be unnecessary to recalculate $\Delta R/R$ for each value of the fitting parameters (e.g., applied field, effective mass, and broadening); thus it is now becoming practical to do a least-squares fit based on exciton theory.

Therefore we feel that both the qualitative and the quantitative features of electroflection spectra can be understood with the excition theory. An even more quantitative, extremely accurate theory of electroreflection spectra will probably be attained when it is possible to make inexpensive least-squares fits of exciton theory to experiment, varying the broadening Γ and its dependence on photon energy in accordance with accurate *a priori* calculations of exciton-phonon scattering. After comparison with such a theory, electroreflectivity data for various semiconductors should yield precise values of energy gaps, effective masses, and band-warping parameters.

VI. SUMMARY

The theory of optical transitions in uniform electric fields including electron-hole final-state correlations has been used to calculate the differential electroreflection spectrum of Ge near the direct edge. The results are in good agreement with experiment, yielding the correct line shapes and accurate values for the matrix elements which determine the strength of the transitions. In addition, from a discussion of the qualitative aspects of electroabsorption, and a set of rules of thumb for the effects of the Kramers-Kronig transformation, we are able to elucidate the physical reasons behind the qualitative behavior of electroreflectivity spectra. The principal uncertainty at the present time lies in our simplified treatment of broadening. Together with a more refined theory of broadening, the exciton theory should provide a basis from which to extract detailed information on the band structure of semiconductors, using highly developed experimental techniques for measuring differential reflectivity in uniform electric fields.

ACKNOWLEDGMENTS

We gratefully acknowledge helpful and enlightening ing discussions with D. E. Aspnes, P. Sheng, and J. J. Hopfield.

APPENDIX

All the calculations described in this paper were done numerically since no explicit closed form expressions for ϵ_2 or ϵ_1 exist. Using the programs described in Refs. 13 and 16, $\epsilon_2(\omega, F)$ was calculated over the range of energy 0.70–1.20 eV, using double-precision arithmetic. The corresponding value due only to the Coulomb interaction is

$$\epsilon_2(\omega, F=0) = \frac{1}{Ra^3}\ \frac{1}{2\pi}\left\{1-\exp\left[-\left(\frac{R}{\hbar\omega - E_{\text{gap}}}\right)^{1/2}\right]\right\}^{-1} \theta(\hbar\omega - E_{\text{gap}}) + \sum_{n=1}^{\infty} (\pi a^3 n^3)^{-1}\ \ \delta\left(\hbar\omega - E_{\text{gap}} + \frac{R}{n^2}\right), \quad \text{(A1)}$$

where R and a are the exciton Rydberg and Bohr radius, respectively, and $\theta(x)$ is the unit step function. The difference between $\epsilon_2(\omega, F)$ and $\epsilon_2(\omega, F=0)$ was convoluted with a Lorentzian of half-width Γ,

$$\frac{\Gamma}{\pi[(\hbar\omega)^2 + \Gamma^2]}\ . \quad \text{(A2)}$$

The convolution of the continuum contribution was performed by using a 32-point Gaussian integration routine,[46] an inherently more accurate calculation than one based on Simpson's rule. This method requires evaluation of the integrand at relatively few points, but has the disadvantage that these points are not evenly spaced. This difficulty was met by using a five-point Lagrangian interpolation routine[47] to evaluate the integrand at points intermediate between the uniform mesh on which it was originally computed. At energies below the gap the discrete

exciton contribution to $\epsilon_2(\omega, F=0)$ must be subtracted from $\epsilon_2(\omega, F)$, which, of course, is continuous across the gap. These δ-function contributions convert the convolution integral into an infinite sum over discrete contributions. This sum was done explicitly, again using double-precision arithmetic, over the first 50 terms and the remainder summed by an Euler-Maclaurin formula.[48] The discrete exciton contribution to the convolution integral was found to be roughly 5% or less of the continuum contribution, except in the immediate vicinity of the gap, where the two contributions become comparable.

The broadened $\Delta\epsilon_2(\omega, \Gamma)$ obtained in this way was then transformed by the Kramers-Kronig integral [Eq. (3.3)] to give $\Delta\epsilon_1(\omega, \Gamma)$. Since a principal-value integration is involved, steps must be taken to ensure that the divergent behavior of the integrand is treated properly. The most effective method is to perform a subtraction,[49] although an analog of Simpson's rule exists.[50] Thus, to calculate $P\int_A^B dx\, f(x)/(x-a)$, we replace $f(x)$ by $g(x)(x-a) + f(a)$, where $g(x) \equiv [f(x)-f(a)]/(x-a)$. The integral becomes $\int_B^A g(x)dx + f(a)\ln|(B-a)/(A-a)|$, where $g(x)$ is well behaved at $x=a$, and hence can be integrated by any of the standard methods. In our calculations Gaussian integration was used. The above procedure of subtractions is particularly accurate if the interval (A, B) is not too large. Thus it is best to break the original integral into several intervals, to use the subtraction technique in the interval containing $x=a$, and to do the other parts in the usual fashion.

One further precaution must be noted. Since in the "near-resonance" region[30] the Kramers-Kronig integrating factor falls off rather slowly at high frequencies [as $(\omega-\omega_0)^{-1}$], some scheme of extrapolating the integrand $\epsilon_2(\omega)$ is needed. As pointed out earlier, a premature cutoff in the integration may lead to spurious results (see Fig. 8). In our calculation it was impractical to broaden each value of $\Delta\epsilon_2$ at the high-frequency end of the spectrum, needed for the Kramers-Kronig integration. Instead, $\Delta\epsilon_2(\omega, \Gamma)$ in the tail region was fitted to the unbroadened result after adjusting the amplitude by an appropriate factor. This factor was determined by a least-squares fit of the form Ax^b to the ratio of broadened to unbroadened values at the highest values at which broadening had been calculated. However, the contribution of this region is small (< 1%) and the total results are not very sensitive to the scheme used. The high-frequency dependence of the unbroadened $\Delta\epsilon_2$ was obtained from the Franz-Keldysh value, with a suitable amplitude and phase shift included[16] (typically, the amplitude was about 15% greater than the Franz-Keldysh value, and shifted in energy by about 1.5 exciton Rydbergs).

---

*Work supported by the U.S. Air Force Office of Scientific Research under Contract No. AF49(638)1545.

†National Science Foundation Graduate Fellow. Present address: Cavendish Laboratory, Free School Lane, Cambridge, England.

‡Present address: Department of Physics and Astronomy, University of Maryland, College Park, Md. 20742.

[1]M. Cardona, *Modulation Spectroscopy* (Academic, New York, 1969).

[2]J. C. Phillips, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1966), Vol. 18, p. 55.

[3]W. Franz, Z. Naturforsch. 13a, 484 (1958).

[4]L. V. Keldysh, Zh. Eksperim. i Teor. Fiz. 34, 1138 (1958) [Sov. Phys. JETP 7, 788 (1958)].

[5]B.O. Seraphin, in *Proceedings of the Seventh International Conference on the Physics of Semiconductors*, edited by M. Hulin (Dunod, Paris, 1964), p. 165.

[6]B. O. Seraphin, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. Beer (Academic, New York, 1970), Vol. VI.

[7]For a comprehensive review see D. E. Aspnes and N. Bottka, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. Beer (Academic, New York, 1970), Vol. VI.

[8]J. Callaway, Phys. Rev. 130, 549 (1963); K. S. Viswanathan and J. Callaway, *ibid.* 143, 564 (1966).

[9]C. B. Duke, Phys. Rev. Letters 15, 625 (1965); C. B. Duke and M. E. Alferieff, Phys. Rev. 145, 583 (1966).

[10]H. I. Ralph, J. Phys. C 1, 378 (1968).

[11]D. F. Blossey, Bull. Am. Phys. Soc. 14, 429 (1969); Ph. D. thesis (University of Illinois, 1969) (unpublished).

[12]D. F. Blossey, Phys. Rev. B 3, 1382 (1971).

[13]J. D. Dow, B. Y. Lao, and S. Newman, Phys. Rev. B 3, 2571 (1971).

[14]B. Y. Lao, J. D. Dow, and F. C. Weinstein, Phys. Rev. Letters 26, 499 (1971); Phys. Rev. B (to be published).

[15]J. D. Dow, B. Y. Lao, and F. C. Weinstein, in *Proceedings of the Tenth International Conference on the Physics of Semiconductors, Cambridge, Mass.*, 1970 (U. S. Atomic Energy Commission, Washington, D. C., 1970), p. 183.

[16]J. D. Dow and D. Redfield, Phys. Rev. B 1, 3358 (1970).

[16a]F. C. Weinstein, J. D. Dow, and B. Y. Lao, Phys. Status Solidi 43b, K105 (1971); 43, K177 (1971).

[17]P. Handler, S. N. Jasperson, and S. Koeppen, Phys. Rev. Letters 23, 1387 (1969).

[18]Good agreement has been obtained from indirect transitions in Ge and Si (see Ref. 14).

[19]R. J. Elliott, Phys. Rev. 108, 1384 (1957); R. S. Knox, *Theory of Excitons* (Academic, New York, 1963).

[20]Of the two factors in the Bloch function, the field generally affects the plane-wave (envelope) part considerably more, since the periodic part is atomlike in character and therefore only weakly polarized by typical laboratory-strength fields.

[21]We are thus neglecting terms of order $R/\Delta E$, where R is the exciton Rydberg, $R \simeq 14$ meV in Ge, while ΔE

is a typical interband energy.

[22]K. W. Boer, H. J. Hansche, and V. Kummel, Z. Physik 155, 170 (1959).

[23]D. Aspnes, Phys. Rev. 147, 544 (1966), see p. 559.

[24]K. Tharmalingam, Phys. Rev. 130, 2204 (1963).

[25]D. Aspnes, Phys. Rev. 153, 972 (1967); see also Ref. 23.

[26]At this stage we have not yet included the Coulomb interaction $-e^2/\epsilon_0 r$ between the electron and hole.

[27]T. P. McLean, in *Progress in Semiconductors*, edited by A. F. Gibson *et al.* (Wiley, New York, 1961), Vol. 5, p. 54.

[28]For a derivation of this formula, see, e.g., L. Landau and E. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon, New York, 1960), p. 256ff.

[29]See, e.g., F. Stern, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), Vol. 15, p. 331ff.

[30]We define "near resonance" as follows. Suppose the imaginary part $g_2(\omega)$ has some structure centered at $\omega=\omega_0$ and characterized by a width γ. Then the near-resonant region of the real part $g_1(\omega)$ consists of frequencies ω such that $\omega-\omega_0 \lesssim \gamma$.

[31]Just as excitons modify the Van Hove behavior near critical points in three dimensions [see Fig. 3(a)], they will also affect the steplike behavior at an M_0 or M_2 two-dimensional critical point predicted by one-electron theory. The logarithmic singularity in $\epsilon_1(\omega)$ will accordingly be smoothed out as well.

[32]See Ref. 1, pp. 245–255, for experimental curves illustrating this point.

[33]The asymptotic behavior of the Kramers-Kronig transform $g_1(\omega)$ of an arbitrary structure $g_2(\omega)$ centered at some $\omega=\omega_0$ and characterized by a width γ may be simply described. Recall that

$$g_1(\omega)=\frac{2}{\pi}P\int_0^\infty \frac{\omega' g_2(\omega')\,d\omega'}{(\omega'-\omega)(\omega'+\omega)} .$$

Since $g_2(\omega')$ is only appreciable in the range $|\omega'-\omega_0|<\gamma$, and the linewidth γ is normally small compared with ω_0, we can make two different approximations to $g_1(\omega)$, depending on whether or not the frequency ω is nearly resonant (see Ref. 30) with the characteristic absorption frequency ω_0:

(i) Near-resonance approximation: If we have $\omega\approx\omega_0$, then we obtain $\omega'/(\omega'+\omega)\approx\frac{1}{2}$ and

$$g_1(\omega)\approx -\frac{1}{\omega-\omega_0}\left(\frac{1}{\pi}\int_0^\infty g_2(\omega')\,d\omega'\right).$$

(ii) High-frequency approximation: If we have $\omega\gg\omega_0$, then we obtain $(\omega'^2-\omega^2)^{-1}\approx-\omega^{-2}$ and

$$g_1(\omega)\approx -\frac{1}{\omega^2}\left(\frac{2}{\pi}\int_0^\infty \omega' g_2(\omega')\,d\omega'\right) .$$

The different frequency dependences $[(\omega-\omega_0)^{-1}$ vs $\omega^{-2}]$ arise from the symmetry property $g_2(-\omega)=-g_2(\omega)$ [e.g., which holds for $\epsilon_2(\omega)$ by time-reversal invariance]. In case (i), only the structure in $g_2(\omega)$ occurring at positive ω contributes appreciably to the integral, while in case (ii), both structures (at positive and negative ω) contribute with opposite signs, leading to a cancellation effect and a more rapid falloff at high frequency.

[34]We require $\tilde{g}_2(-\omega)=-\tilde{g}_2(\omega)$. This introduces a discontinuity in g_2 at $\omega=0$, leading to a logarithmic singularity in $g_1(\omega)$ [as we have seen in Fig. 4(b)]. However, this does not appreciably affect the behavior of the zeros of these functions at higher values of ω.

[35]D. L. Dexter, Nuovo Cimento 48B, 409 (1967).

[36]See R. Knox, Ref. 19.

[37]J. Bardeen, L. H. Hall, and F. J. Blatt, in *Photoconductivity Conference*, edited by E. Breckenridge (Wiley, New York, 1954), p. 146.

[38]B. O. Seraphin and N. Bottka, Phys. Rev. 145, 628 (1966).

[39]M. Cardona [J. Appl. Phys. 36, 2181 (1965)] has used a similar technique to fit ϵ_1 in CdTe.

[40]These curves assume equal light- and heavy-hole matrix elements and have been evaluated for a broadening parameter $\Gamma=3$ meV.

[41]D. E. Aspnes and A. Frova, Phys. Rev. B 2, 1037 (1970).

[42]G. G. Macfarlane, T. P. McLean, J. E. Quarrington, and V. Roberts, Proc. Phys. Soc. (London) 71, 863 (1958).

[43]Q. H. F. Vrehen, Phys. Rev. 145, 675 (1966).

[44]M. Cardona and F. H. Pollak, Phys. Rev. 142, 530 (1966).

[44a]We are grateful to J. Grover and P. Handler [Phys. Rev. B (to be published)] for pointing out that the value 1.16 quoted in Refs. 16a and 17 differs from the ratio of matrix elements by a mass ratio to the $\frac{4}{3}$ power.

[45]J. D. Dow and H. Weston (unpublished).

[46]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (U.S. GPO, Washington, D.C., 1964), p. 887.

[47]Reference 46, p. 879.

[48]Reference 46, p. 16.

[49]See, for example, M. L. Goldberger and K. Watson, *Collision Theory* (Wiley, New York, 1964), p. 566 for a discussion of subtracted dispersion relations.

[50]L. M. Delves, Computer J. 10, 389 (1968).