

Modeling the optical dielectric function of semiconductors: Extension of the critical-point parabolic-band approximation

Charles C. Kim,* J. W. Garland, H. Abad, and P. M. Raccah

Department of Physics, University of Illinois at Chicago, P.O. Box 4348, Chicago, Illinois 60680

(Received 29 August 1991)

A model is proposed for the line shape of the optical dielectric function of zinc-blende semiconductors. For comparison with previously proposed models, this model is used primarily with spectroscopic ellipsometry data (but also transmission data below 1.5 eV) to obtain an analytic room-temperature dielectric function for GaAs. It is found to be more generally valid than the harmonic-oscillator model, the critical-point (CP) model, or the model of Adachi. It is applicable over the entire range of photon energies, below and above the lowest band gaps, incorporates the electronic band structure of the medium, and exactly satisfies the Kramers-Kronig transformation. It goes beyond the CP parabolic-band approximation in that it correctly takes into account the full analytic form of the electronic density of states and thus does not require the use of arbitrary cutoff energies. Also, it allows one to go beyond the usual approximation of Lorentzian broadening, which is known to be incorrect for elements and compounds above very low temperatures. For these reasons, it results in excellent quantitative agreement with experimental results for the dielectric function and for its derivatives with respect to photon energy, much better than that given by earlier models. Finally, the parameters of the model are physically significant and are easily determined as functions of composition for semiconductor alloys. Application of the model to the fitting of spectroscopic data on GaAs strongly suggests that spectroscopic ellipsometry does not measure the true bulk dielectric function. It also supports the conclusion that the line-shape broadening in GaAs at room temperature is more nearly Gaussian than Lorentzian.

I. INTRODUCTION

Measurements of optical properties long have been a powerful tool in studying the electronic structure of solids. Today, a knowledge of the refractive indices and absorption coefficients of semiconductors is especially important in the design and analysis of heterostructure lasers and other wave-guiding semiconductor devices. The dielectric function, $\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega)$, fully describes the optical properties of any homogeneous medium¹⁻⁶ at all photon energies $\hbar\omega$.

Spectroscopic ellipsometry (SE) is an excellent technique with which to investigate the optical response of semiconductors and, in particular, to measure the spectral dependence of the dielectric function. However, values for $\epsilon(\omega)$ obtained from experimental data have the serious deficiency that they are not expressed as functions of electronic critical-point energies E_j , alloy compositions x , or even the photon energy $\hbar\omega$. The construction of accurate model line shapes is necessary in order to express $\epsilon(\omega)$ as a function of $\hbar\omega$ or of x and the E_j for an alloy series. Thus, the accurate modeling of $\epsilon(\omega)$ provides a decisive advantage in simulating the dielectric function of multilayer systems and in designing optoelectronic devices.

An accurate model for $\epsilon(\omega)$ also is necessary for the accurate fitting of SE data or even of modulated spectroscopy data obtained by thermoreflectance, piezoreflectance, or any form of electroreflectance. In particular, it is required for the accurate determination⁷ of

critical point energies E_j and line widths Γ_j , and for the detailed characterization⁸ of optical materials. To be useful for these purposes, a model for $\epsilon(\omega)$ must accurately describe both $\epsilon(\omega)$ and its first three derivatives with respect to photon energy. There are two reasons for this. First, the obtaining of accurate values for the E_j and Γ_j from SE data requires the fitting of second or third derivatives of that data. Second, electroreflectance⁹⁻¹¹ and photoreflectance¹² line shapes contain contributions proportional to the first and second derivatives of $\epsilon(\omega)$, for which the standard critical-point parabolic-band (CPPB) model is not accurate, as well as the Franz-Keldysh-Aspnes¹³⁻¹⁵ contribution proportional to the third derivative.

Early models for $\epsilon(\omega)$ either have been purely phenomenological or have relied upon the parabolic-band (PB) approximation. The phenomenological models¹⁶⁻¹⁹ provide fits to $\epsilon(\omega)$ which are numerically more accurate than those given by models which incorporate the parabolic-band approximation. However, the phenomenological models completely fail to reproduce either the line shape of $\epsilon_2(\omega)$ near the E_0 critical point or the line shape of the derivatives of $\epsilon(\omega)$ near any critical point. Furthermore, because of their phenomenological, nonphysical character, they cannot be expected *a priori* to provide satisfactory fits to the dependence of $\epsilon(\omega)$ on physical parameters such as alloy composition, temperature, strain, or electric field. The models based on the PB approximation are more satisfactory in some respects, but do not yield good fits to $\epsilon(\omega)$ and still violate the full

critical-point structure of the electronic density of states, which is imposed by topological considerations. More recent composite models,^{20–22} which augment a CPPB model with phenomenological contributions at higher energies, fail to overcome fully the deficiencies of the earlier models. In particular, it has been shown that these models require²⁰ the assumption of large contributions to $\epsilon(\omega)$ from indirect transitions—an assumption which is easily shown to be false. Finally, all previous models incorporate the assumption of Lorentzian broadening, which is known^{23,24} to be wrong for elements and compounds. None of these models are capable of fitting accurately both the dielectric function and its derivatives.

In this paper we present a new model for $\epsilon(\omega)$ which (i) yields excellent simultaneous fits to $\epsilon(\omega)$ and its derivatives with respect to energy, (ii) fully incorporates the analytic form of the joint density of the states $J_{cv}(E)$ between each pair of valence and conduction bands, (iii) yields accurate values of the critical-point energies and line widths, and (iv) allows one to go beyond the approximation of Lorentzian broadening. This model allows one to express $\epsilon(\omega)$ as a function of alloy composition, as did the model of Adachi²⁰ and its generalizations,^{21,22} but does not require the imposition of arbitrary cutoff energies or the assumption of large contributions to $\epsilon(\omega)$ from indirect transitions, as did that model.

In Sec. II we present the formalism of the dielectric function for a bulk solid in the absence of line broadening. Then, we show how this formalism is generalized properly to take line broadening into account either within a Lorentzian or a Gaussian approximation. In Sec. III we briefly review the previous models for $\epsilon(\omega)$. In Sec. IV,

we develop our new model. In Sec. V we show how to apply our model to the calculation of an analytic $\epsilon(\omega)$, and compare the results of our model with experimental data for GaAs, for $\epsilon(\omega)$ and its numerical first, second, and third derivatives. A comparison of the fits obtained from our model and shown in Sec. V to those obtained from previous models and shown in Sec. III shows that our model is vastly superior numerically as well as analytically. In Sec. VI we summarize the strengths of our model, consider ways it could be improved and present a new physical result based on our fittings of GaAs spectroscopic data.

II. THE OPTICAL DIELECTRIC FUNCTION WITH LINE BROADENING

The basic formula for the transverse dielectric function in the absence of line broadening has been derived by several authors^{16,25–27} by calculating the transition probabilities. However, the proper method of introducing line broadening is not obvious within that formalism. On the other hand, it is quite obvious within the various formalisms^{28–30} used to derive the longitudinal dielectric function. We introduce line broadening using the formalism of Ehrenreich and Cohen.³⁰ This is justified because the transverse and longitudinal dielectric functions become identical in the form in the limit of infinite wavelength (or zero wave number), the appropriate limit to consider for optical response.

In the absence of line broadening the longitudinal dielectric function is given by the formula³⁰

$$\epsilon_l(\omega, \mathbf{q}) = 1 - \lim_{\eta \rightarrow 0} \frac{8\pi e^2}{q^2 \Omega} \sum_{c, \mathbf{k}} \sum_{v, \mathbf{k}'} |(c, \mathbf{k}|v, \mathbf{k}')|^2 \left(\frac{\delta_{\mathbf{k}, \mathbf{k}'+\mathbf{q}}}{\hbar\omega - E_c(\mathbf{k}) + E_v(\mathbf{k}') + i\eta} - \frac{\delta_{\mathbf{k}, \mathbf{k}'-\mathbf{q}}}{\hbar\omega + E_c(\mathbf{k}) - E_v(\mathbf{k}') + i\eta} \right). \quad (2.1)$$

Here,

$$(c, \mathbf{k}|v, \mathbf{k}') = \frac{1}{\Delta} \int_{\text{unit cell}} d\mathbf{r} u_{c\mathbf{k}}^*(\mathbf{r}) u_{v\mathbf{k}'}(\mathbf{r}), \quad (2.2)$$

Ω and Δ are the volumes of the solid and of the unit cell, respectively, and $u_{c\mathbf{k}}(\mathbf{r})$ and $u_{v\mathbf{k}'}(\mathbf{r})$ are the usual periodic parts of the Bloch functions in the conduction and valence bands, respectively. The sums over c and v go over all conduction (c) and valence (v) bands, the sums over \mathbf{k} and \mathbf{k}' go over the first Brillouin zone, and η is a positive infinitesimal.

For small \mathbf{q} , perturbation theory yields the result

$$|(c, \mathbf{k}|v, \mathbf{k}')|^2 = |(c, \mathbf{k}|v, \mathbf{k} \pm \mathbf{q})|^2 = \left(\frac{\hbar q P_{cv}(\mathbf{k})}{m E_{cv}(\mathbf{k})} \right)^2, \quad (2.3)$$

where m is the free-electron mass, $E_{cv}(\mathbf{k}) = E_c(\mathbf{k}) - E_v(\mathbf{k})$,

$$P_{cv}(\mathbf{k}) = \frac{1}{\Delta} \left| \int u_{c\mathbf{k}}^*(\mathbf{r}) p_\mu u_{v\mathbf{k}}(\mathbf{r}) d^3r \right|,$$

and p_μ is the component of the momentum operator in the direction of \mathbf{q} . Substitution of Eq. (2.3) into Eq. (2.1) leads to the final result

$$\epsilon_l(\omega) = 1 - \lim_{\eta \rightarrow 0} \frac{8\pi \hbar^2 e^2}{m^2} \sum_{k, c, v} \left(\frac{P_{cv}(\mathbf{k})}{E_{cv}(\mathbf{k})} \right)^2 \left(\frac{1}{\hbar\omega - E_{cv}(\mathbf{k}) + i\eta} - \frac{1}{\hbar\omega + E_{cv}(\mathbf{k}) + i\eta} \right). \quad (2.4)$$

This result can also be expressed in the form

$$\epsilon_1(\omega) = 1 + i \lim_{\eta \rightarrow 0} \frac{8\pi\hbar^2 e^2}{m^2} \sum_{\mathbf{k}, c, v} \left(\frac{P_{cv}(\mathbf{k})}{E_{cv}(\mathbf{k})} \right)^2 \left(\int_0^\infty ds e^{i(\hbar\omega - E_{cv}(\mathbf{k}) + i\eta)s} - \int_0^\infty ds e^{i(\hbar\omega + E_{cv}(\mathbf{k}) + i\eta)s} \right), \quad (2.5)$$

which can be derived directly within a Green's-function formalism.³¹ The optical dielectric function is identical with this dielectric function, except that the direction of the momentum is in the direction of the electric field of the light for the optical dielectric function. The summation over \mathbf{k} in Eq. (2.5) can be replaced by an integration over the energy $E \equiv E_{cv}(\mathbf{k})$ by introducing the joint density of states, $J_{cv}(E)$. One finds the result

$$\epsilon(\omega) = 1 + i \frac{8\pi\hbar^2 e^2}{m^2} \sum_{c, v} \int J_{cv}(E) dE \left(\frac{P_{cv}(E)}{E} \right)^2 \left(\int_0^\infty ds e^{i(\hbar\omega - E + i\eta)s} - \int_0^\infty ds e^{i(\hbar\omega + E + i\eta)s} \right) \quad (2.6)$$

for the optical dielectric function, where $P_{cv}(E)$ is an average of $P_{cv}(\mathbf{k})$ over the surface in \mathbf{k} space given by $E_{cv}(\mathbf{k}) = E$, with the direction of p_μ in the equation for $P_{cv}(\mathbf{k})$ taken to be along that of the electric field of the light. Equation (2.6) is valid in the absence of line broadening and of indirect transitions. In actuality, due to defect scattering, electron-phonon scattering, electron-electron scattering, etc., the line width is finite. Formally, this results from the replacement of the infinitesimal η by a function $\gamma(s)$, which yields the result

$$\epsilon(\omega) = 1 + i \frac{8\pi\hbar^2 e^2}{m^2} \sum_{c, v} \int J_{cv}(E) dE \left(\frac{P_{cv}(E)}{E} \right)^2 \left(\int_0^\infty ds e^{i\{(\hbar\omega - E + i\gamma(s))s\}} - \int_0^\infty ds e^{i\{(\hbar\omega + E + i\gamma(s))s\}} \right). \quad (2.7)$$

In order to obtain a usable formula for $\epsilon(\omega)$, we expand $\gamma(s)$ as a power series in $s = t/\hbar$, where t is time, and truncate the series after the first two terms:

$$\gamma(s) = \Gamma + 2\sigma^2 s. \quad (2.8)$$

In previous models for $\epsilon(\omega)$, for simplicity $\gamma(s)$ has been replaced by the constant Γ . This approximation, which is referred to as Lorentzian broadening, yields the result

$$\epsilon(\omega) = 1 - \frac{8\pi\hbar^2 e^2}{m^2} \sum_{c, v} \int J_{cv}(E) dE \left(\frac{P_{cv}(E)}{E} \right)^2 \left(\frac{1}{\hbar\omega - E + i\Gamma} - \frac{1}{\hbar\omega + E + i\Gamma} \right). \quad (2.9)$$

However, it has been found theoretically that for the case of line broadening induced either by electron-phonon scattering³² or by dilute impurity scattering,^{33,34} $\gamma(s)$ is best approximated by $2\sigma^2 s$, where σ is the root-mean-square scattering t matrix. This result has been supported by studies of the line shape of $\epsilon(\omega)$ near the E_1 critical points of GaAs and CdTe.²³ This approximation yields the result,

$$\epsilon(\omega) = 1 + i \frac{8\pi\hbar^2 e^2}{m^2} \sum_{c, v} \int J_{cv}(E) dE \left(\frac{P_{cv}(E)}{E} \right)^2 \left(\int_0^\infty ds e^{i(\hbar\omega - E + 2i\sigma^2 s)s} - \int_0^\infty ds e^{i(\hbar\omega + E + 2i\sigma^2 s)s} \right). \quad (2.10)$$

It is known as Gaussian broadening because of the Gaussian time decay in Eq. (2.10). Lorentzian and Gaussian broadening each are a special case of Eq. (2.8).

Equations (2.9) and (2.10) form the basis for our modeling of $\epsilon(\omega)$. Notice that the factor $1/E^2$ in Eqs. (2.7), (2.9), and (2.10) is kept inside the integration as the effect of line broadening is introduced. If the dielectric function is derived by calculating the transition probability without the effect of line broadening, it leads immediately to the formula

$$\begin{aligned} \epsilon_2(\omega) &= \frac{8\pi^2 \hbar^2 e^2}{m^2} \\ &\times \sum_{c, v} \int J_{cv}(E) dE \left(\frac{P_{cv}(E)}{E} \right)^2 \delta(\hbar\omega - E) \\ &= 8 \left(\frac{\pi e}{m\omega} \right)^2 \sum_{c, v} P_{cv}(\hbar\omega)^2 J_{cv}(\hbar\omega) \end{aligned} \quad (2.11)$$

for the imaginary part of $\epsilon(\omega)$, which also can be obtained from Eq. (2.7) in the limit $\gamma(s) \rightarrow 0$. In some papers²⁰⁻²² in which line broadening is introduced phenomenologically starting from Eq. (2.11), the formula obtained for $\epsilon_2(\omega)$ contains a factor $1/\omega^2$ rather than the factor \hbar^2/E^2 . This mistake leads to a value of $\epsilon(\omega)$ which goes to infinity as ω approaches zero, rather than to the correct value of zero. Other authors^{22,35,36} have proposed approximations to circumvent the divergence arising from the spurious factor $1/\omega^2$. However, this divergence never appears if line broadening is introduced properly, as is done here.

III. PREVIOUS MODELS FOR THE DIELECTRIC FUNCTION

In the past, models for the dielectric function either have been purely phenomenological or have relied upon the CPPB approximation. Among the phenomenologi-

cal models both the Lorentzian oscillator¹⁶ and the harmonic oscillator^{17,18} (HO) are well known. These models approximate the continuum of transitions possible between band states in a solid, with the resultant critical-point structure, by a small discrete set of transitions and thus do not fully incorporate the band structure. In this respect, the CPPB models, largely developed by Cardona and Aspnes,^{37,38} more accurately represent the band structure. Recently, Adachi²⁰ developed a new model for $\epsilon(\omega)$ in which a CPPB model is augmented by discrete HO transitions at higher energies. However, none of these models are capable of describing the dielectric function and its derivatives simultaneously. Here we briefly review the HO model of Erman *et al.*,¹⁸ the CPPB model, the model of Adachi²⁰ and more recent improvements in those models, pointing out those features in each model which limit its validity.

In the harmonic oscillator model of Erman *et al.*,¹⁸ $\epsilon(\omega)$ is given by

$$\epsilon(\omega) = 1 - \sum_{k=1}^n A_k \left(\frac{1}{\hbar\omega - E_k + i\Gamma_k} - \frac{1}{\hbar\omega + E_k + i\Gamma_k} \right), \quad (3.1)$$

where E_k is the energy of a harmonic oscillator and Γ_k is its linewidth. Equation (3.1) is much simpler than Eq. (2.9) and corresponds to replacing $J_{cv}(E)$ in Eq. (2.9) by a sum of Lorentzians. In principle, each single transition from a lower band to a higher band could be represented by a harmonic oscillator, but in practice the minimum possible number of oscillators is used to represent the dielectric function. This model fails to describe $\epsilon(\omega)$ below or around the band edge at E_0 because $\epsilon_2(\omega)$ has a broadened square-root singularity at the band edge with little background, whereas the HO model yields only a sum of Lorentzian peaks for $\epsilon_2(\omega)$. Figure 1 shows the result of fitting $L(\omega_j)$, the dielectric function of GaAs as determined from the spectral data in the manner detailed in Sec. V. Following Erman *et al.*, seven harmonic oscillators were used. The numerical derivatives of $L(\omega_j)$ also are compared with those of the fit. The fit to $L(\omega_j)$ is excellent between 2.6 eV and 5 eV, but is not good near E_0 or above 5 eV. The resultant numerical derivatives of that fit fail to fit the numerical derivatives of $L(\omega_j)$, $[L(\omega_j)]_{\text{num}}^{(n)}$. Furthermore, in order to obtain this fit to $L(\omega_j)$, one is forced to include transitions at 3.48 eV and 3.836 eV which do not correspond to true critical-point (CP) transitions. This model contains 21 free parameters, with $\epsilon(\omega)$ depending nonlinearly on 14 of the 21.

The root-mean-square (rms) fractional error σ_0 of this fit to $L(\omega_j)$, defined by

$$\sigma_0^2 = \frac{\sum_j |\epsilon(\omega_j) - L(\omega_j)|^2}{\sum_j |L(\omega_j)|^2} \quad (3.2)$$

is 2.6%. Restricting the fit to the range from 1.5 eV to 5.0 eV results in an excellent fit to $L(\omega_j)$, reducing the rms fractional error σ_0 to 1.0%. These values of σ_0 are

better than those obtainable with the CPPB model but more than twice as large as those which can be found with the new model proposed here. The values found for the energies and line widths of the peaks are shown in Table I for both fitting ranges.

Because the HO energies E_k are not simply related to the CP energies E_j , this model is not simply related to the band structure. Moreover, the phenomenological line widths Γ_k in this model bear no relationship to true line widths and are much broader; thus they contain no useful information about sample quality. Recently, Terry¹⁹ has substantially improved the HO model by allowing the contribution of each oscillator to have an arbitrary phase angle. He used nine oscillators, and thus 36 free parameters, with the resultant $\epsilon(\omega)$ depending nonlinearly on 27 of the 36, but did obtain an excellent fit to $L(\omega_j)$. However, even this improved version of the HO model still suffers from the basic weaknesses discussed above.

The CPPB model was originally developed by Cardona³⁷ and Aspnes³⁸ and has been used by many authors to investigate the optical properties of solids. This model gives an accurate representation of the derivatives of $\epsilon(\omega)$ of order higher than first order, because the critical point structure is greatly enhanced in those derivatives. Thus, its use enables one to determine the CP energies and line widths quite accurately. However, this

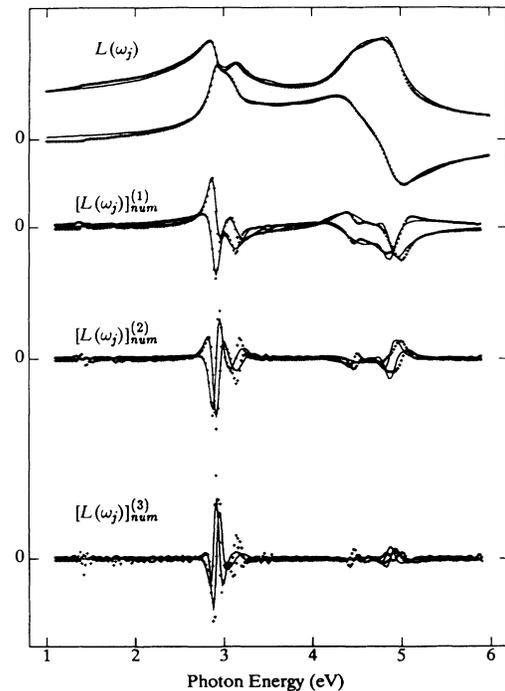


FIG. 1. Fit to $L(\omega_j)$ obtained from the seven-harmonic-oscillator model. The bullets and the plus signs are the real and imaginary part of $L(\omega_j)$, respectively. The solid lines in $L(\omega_j)$ show the fits obtained with the HO model. The first three numerical derivatives of $L(\omega_j)$ and of the fit to $L(\omega_j)$ also are shown. It is clear that the derivatives of the fit do not provide a good representation for the derivatives of $L(\omega_j)$.

TABLE I. The values in eV of the energies and line widths of the peaks in the HO model for the dielectric function of GaAs for fittings to $L(\omega_j)$ over the ranges (a) $1.0 \text{ eV} \leq \hbar\omega \leq 6.0 \text{ eV}$ and (b) $1.5 \text{ eV} \leq \hbar\omega \leq 5.0 \text{ eV}$.

Energy range	Parameter	Peak index						
		1	2	3	4	5	6	7
(a)	E_j	2.916	3.127	3.484	3.836	4.598	4.871	7.587
(b)	E_j	2.915	3.117	3.485	3.896	4.488	4.845	9.28
(a)	Γ_j	0.084	0.231	0.265	0.321	0.462	0.154	2.377
(b)	Γ_j	0.072	0.248	0.292	0.610	0.294	0.310	0.463

model gives only a very poor representation of the dielectric function itself and is not suitable for the description of the dielectric function or its first derivative with respect to photon energy, temperature or pressure.

Within a CP model, $\epsilon(\omega)$ in Eq. (2.9) usually is given by the approximate formula

$$\epsilon(\omega) = 1 - \frac{8\pi e^2}{m^2\omega^2} \sum_j P_j^2 \int dE \frac{J_j(E)}{\hbar\omega - E + i\Gamma_j}. \quad (3.3)$$

Here, the sum over c and v in Eq. (2.9) is replaced by a sum over the critical points j , which implicitly contains a sum over c and v . In this model $P_{cv}(E)$ in Eq. (2.9) is considered to be constant near any critical point and is replaced by the constant P_j . The second term in the second set of large parentheses in Eq. (2.9), being much smaller in magnitude than the first term near any critical point, is usually neglected. Finally, the factor E^{-2} in Eq. (2.9) is replaced by $(\hbar\omega)^{-2}$. This is justified only if the spectral range of the contribution of each CP is limited to the region $|\hbar\omega - E_j| \ll E_j$ near the CP. This approximation leads to a serious analytic error, a divergence in $\epsilon(\omega)$ as $\omega \rightarrow 0$, as well as quantitative errors. Also, note that the usual neglect of the second term in the second set of large parentheses in Eq. (2.9), although it is not so serious as the previous approximation, does prevent $\epsilon_2(\omega)$ from going to zero as $\omega \rightarrow 0$, as it should.

In the immediate vicinity of any three-dimensional (3D) CP of type M_j , $E_{cv}(\mathbf{k})$ is a parabolic function of \mathbf{k} . Thus, each 3D CP is characterized by a square-root singularity, the M_0 and M_2 two-dimensional (2D) critical points are characterized by a discontinuity in $J_{cv}(E)$, and each one-dimensional CP is characterized by an inverse square-root singularity. Excluding the only remaining type of CP, the 2D M_1 CP, which does not occur, the derivatives of $\epsilon(\omega)$ can be represented within the CPPB approximation by the equation

$$\frac{d^n}{d\omega^n} [\omega^2 \epsilon(\omega)] = \sum_j C_j \frac{e^{i\theta_j}}{(\hbar\omega - E_j + i\Gamma_j)^{1+n-0.5d_j}}, \quad (3.4)$$

where C_j is a constant, and d_j is the dimensionality of the CP. In the neglect of many-body effects, the phase angle θ_j is given by the equation³⁹

$$\theta_j = \frac{\pi}{2}(j - d_j). \quad (3.5)$$

However, usually θ_j is treated as an adjustable param-

eter. The parametrization of the value of θ_j (Ref. 39) is justified in the literature as representing the effect of excitons near the CP energy. However, this justification is at best questionable, because θ_j is found to be very near its theoretical value for the E_0 and $E_0 + \Delta_0$ critical points, at which one might expect excitons, but is found to be very different from the value expected theoretically for $d = 2$ and $d = 3$ at the E_1 and $E_1 + \Delta_1$ critical points, near which one does not expect excitons.

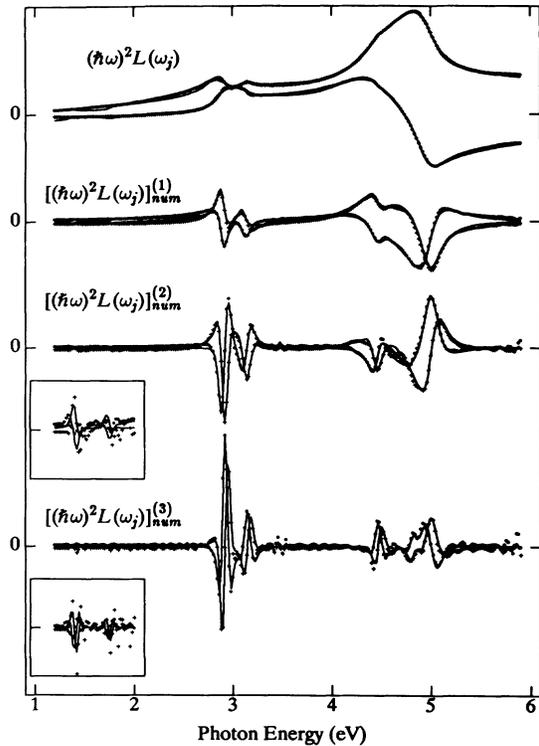


FIG. 2. Separate fits of the CPPB model to $(\hbar\omega_j)^2 L(\omega_j)$ and to its first three numerical derivatives. The bullets and the plus signs show the real and imaginary parts of $[(\hbar\omega_j)^2 L(\omega_j)]_{\text{num}}^{(n)}$, respectively. The solid lines show the resultant fits. The insets show the fits to the second and third numerical derivatives with the scale expanded by a factor of 20 and 10, respectively. The glitches in the second and third derivatives of $L(\omega_j)$ seen at 1.37, 1.5, and 3.5 eV in our figures result from the joining of different sets of data or of data from regions having different spacings in energy; they have no physical significance.

In Sec. VI we propose a different, more physically reasonable explanation for the anomalous values found for the θ_j at the E_1 and $E_1 + \Delta_1$ critical points and many higher-energy critical points.

Equation (3.4) is known as the standard analytical representation^{37,38} and has been widely used³⁵ to fit the line shape of the derivatives of $\epsilon(\omega)$. Rigorously speaking, it is not valid for $n = 0$. However, it is conventionally used even for $n = 0$, ignoring the contributions to $\epsilon(\omega)$ associated with any CP, j , for very large $|\hbar\omega - E_j|$, and making the change

$$\frac{1}{(\hbar\omega - E_j + i\Gamma_j)^{1+n-0.5d_j}} \rightarrow \ln(\hbar\omega - E_j + i\Gamma_j)$$

for $n = 0$ for those critical points with $d_j = 2$.

Figure 2 shows the results of independently fitting $(\hbar\omega)^2 L(\omega_j)$ and each of its first three numerical derivatives, $[L(\omega_j)]_{\text{num}}^{(n)}$, by $(\hbar\omega)^2 \epsilon(\omega_j)$ of Eq. (3.4) and its numerical derivatives, $[(\hbar\omega)^2 \epsilon(\omega_j)]_{\text{num}}^{(n)}$, within the standard analytical representation, Eq. (3.4). This figure shows that the CPPB model provides a good, but not excellent, representation for the experimentally measured dielectric function, particularly below E_1 for $L_1(\omega_j)$, the real part of $L(\omega_j)$, and provides an excellent fit to its second and third derivatives. However, as is shown by Fig. 3, fitting the imaginary part of $(\hbar\omega)^2 L(\omega_j)$, $(\hbar\omega)^2 L_2(\omega_j)$, and its first three derivatives simultaneously makes the fits to $(\hbar\omega)^2 L_2(\omega_j)$ and its first two derivatives poor and the fit to $[(\hbar\omega)^2 L_2(\omega_j)]_{\text{num}}^{(3)}$ only fairly good. In these fits seven critical points were used between 1.0 eV and 6.0 eV, and, as is customary,³⁵ the values of the θ_j were left free. The values chosen for the dimensionalities d_j , the values found for the parameters E_j and Γ_j for each critical point, and the rms fractional errors σ_n and σ defined by

$$\sigma_n^2 = \frac{\sum_j \{[\epsilon_2(\omega_j) - L_2(\omega_j)]_{\text{num}}^{(n)}\}^2}{\sum_j \{[L_2(\omega_j)]_{\text{num}}^{(n)}\}^2} \quad (3.6)$$

and

$$\sigma^2 = \frac{1}{4} \sum_{n=0}^3 \sigma_n^2 \quad (3.7)$$

are shown in Table II. Thus, these fits contained 28 adjustable parameters, with $\epsilon(\omega)$ depending nonlinearly on 21 of the 28. The dependence on n of the values found for the E_j and Γ_j clearly indicates the existence of shortcomings in the CPPB model.

The model of Adachi²⁰ combines features of the HO and CPPB models and thus has some of the virtues and some of the faults of each of those models. Adachi treated the E_0 , $E_0 + \Delta_0$, and $E_1(L)$ critical points within a CPPB model and mimicked the effect of the next three critical points, which are broader and close in energy, by a single HO transition. He explicitly included the effect of indirect transitions below $E_1(L)$, but neglected the $E_1(L) + \Delta_1(L)$ critical point. By also setting $\Gamma = 0$

at the E_0 and $E_0 + \Delta_0$ critical points and choosing $C_j(E_0 + \Delta_0) = 0.5C_j(E_0)$, he was able to reduce the number of free parameters in his model to 14, less than the number in either the HO or the CPPB model. However, he was forced to pay a high price for this reduction in the number of free parameters. First, as is shown in Fig. 4, his model does not provide a good fit to either $\epsilon(\omega)$ or its derivatives. Also, his model violates the full CP structure of $J_{cv}(E)$ imposed by topological considerations, thus requiring the adoption of artificial cutoffs in $J_{cv}(E)$. Furthermore, in order to obtain fits which are not very poor in the energy region between $E_0 + \Delta_0$ and E_1 , he was forced to assume a contribution to $\epsilon_2(\omega)$ from indirect transitions much larger than that from direct transitions and 2-3 orders of magnitude larger than is physically correct. Finally, he treated line broadening in a manner which violates the analytic properties of $\epsilon(\omega)$, as has been discussed in Sec. II above.

More recently,²¹ Adachi has obtained improved fits to $L(\omega)$ by introducing excitonic effects at the E_1 CP, increasing the number of parameters in his model to

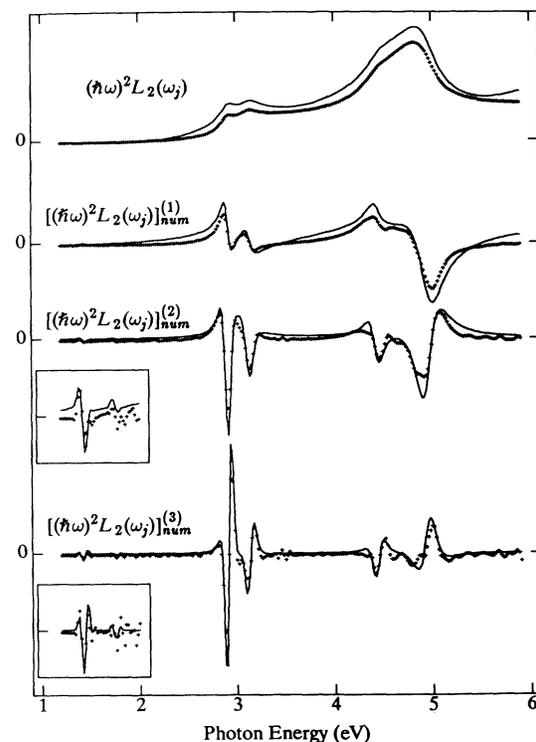


FIG. 3. Simultaneous fit of the CPPB model to $(\hbar\omega)^2 L(\omega_j)$ and its first three numerical derivatives. The bullets and the plus signs are the real and imaginary parts of $[(\hbar\omega)^2 L(\omega_j)]_{\text{num}}^{(n)}$, respectively. The solid lines show the resultant fit. The insets show the fits to the second and third numerical derivatives with the scale expanded by a factor of 17 and 8, respectively. The glitches in the second and third derivatives of $L(\omega_j)$ seen at 1.37, 1.5, and 3.5 eV in our figures result from the joining of different sets of data or of data from regions having different spacings in energy; they have no physical significance.

TABLE II. The dimensionalities assigned each CP and the values of the CP energies and line widths of GaAs obtained by fitting spectral data with the CPPB model given by Eq. (3.4). The values are given in units of eV. The first four columns of values were found by fitting the numerical n th derivatives of the spectral data with respect to photon energy with the numerical n th derivatives of the CPPB model for $\epsilon(\omega)$, for $n = 0, \dots, 3$. The values in the last column were found by simultaneously fitting the data and its first three numerical derivatives. Values for $E_0(\Gamma) + \Delta_0(\Gamma)$, $\Gamma[E_0(\Gamma)]$, and $\Gamma[E_0(\Gamma) + \Delta_0(\Gamma)]$ could not be found by fitting and were fixed at the values 1.745 eV, 0.005 eV, and 0.005 eV, respectively, which are not listed below.

Critical point	d_j	Order of the derivative, n				Simultaneous fit
		0	1	2	3	
$E_0(\Gamma)$	3	1.348	1.348	1.404	1.403	1.429
$E_1(\Lambda)$	2	2.806	2.896	2.911	2.918	2.914
$E_1(\Lambda) + \Delta_1(\Lambda)$	2	3.105	3.115	3.131	3.153	3.132
$E'_0(\Delta)$	2	4.447	4.462	4.452	4.472	4.431
$E_2(X)$	2	4.195	4.524	4.803	4.812	4.738
$E_2(\Sigma)$	2	4.939	4.974	5.006	4.996	4.951
$\Gamma[E_1(\Lambda)]$	2	0.079	0.036	0.036	0.042	0.038
$\Gamma[E_1(\Lambda) + \Delta_1(\Lambda)]$	2	0.002	0.044	0.063	0.061	0.063
$\Gamma[E'_0(\Delta)]$	2	0.004	0.073	0.079	0.067	0.085
$\Gamma[E_2(X)]$	2	0.311	0.426	0.335	0.068	2.375
$\Gamma[E_2(\Sigma)]$	2	0.010	0.135	0.139	0.133	0.141
rms fractional error		4.2%	7.0%	12.5%	15.2%	18.2%

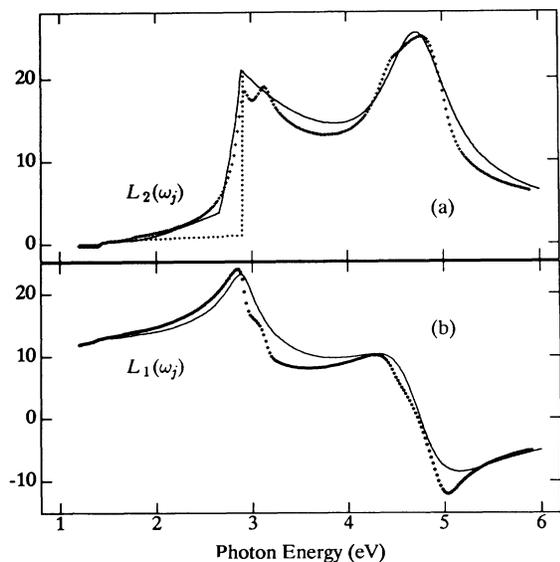


FIG. 4. Fit to $L(\omega_j)$ obtained from the model of Adachi. In (a) the plus signs show $L_2(\omega_j)$. The dotted line is obtained assuming $E_1(\Lambda)$ to be a 2D M_0 CP and excluding the contribution from indirect transitions. The solid line shows the final fit, in which the region below E_1 is improved by assuming $E_1(\Lambda)$ to be a 3D M_1 CP and by including a large contribution from indirect transitions, which is unphysical. In (b) the dots show $L_1(\omega_j)$, and the solid line shows the parametrized analytical form for $\epsilon_1(\omega_j)$ obtained by Adachi from the Kramers-Kronig relationship.

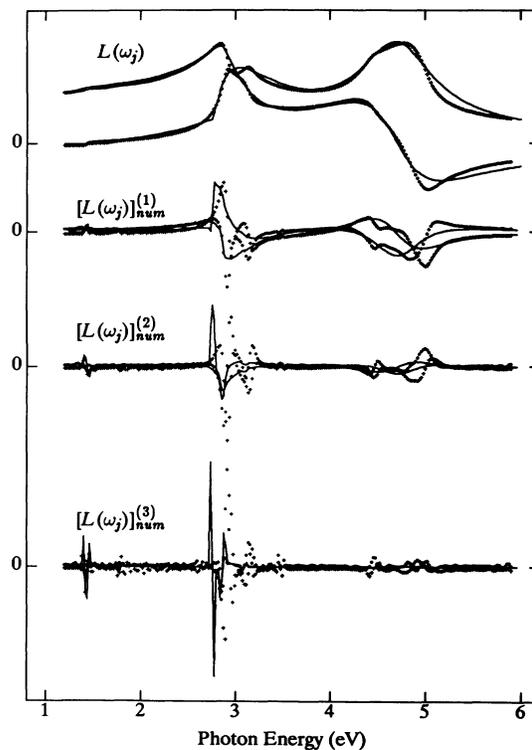


FIG. 5. Fit to $L(\omega_j)$ obtained from the improved model of Jenkins. The bullets and the plus signs show the real and imaginary parts of $L(\omega_j)$, respectively. The first three numerical derivatives of $L(\omega_j)$ and of the fit to $L(\omega_j)$ also are shown. It is clear that the derivatives of the fit do not provide a good representation for the derivatives of $L(\omega_j)$.

19. However, this would appear highly unphysical, especially for the fitting of room-temperature data; the room-temperature lifetime of an excitonic resonance at E_1 should be far too short for the resonance to significantly affect $\epsilon(\omega)$. Also recently, Jenkins²² has proposed an improved version of the model of Adachi which contains 20 parameters and in which the matrix elements $P(E)$ in Eqs. (2.6)–(2.11) decay exponentially with increasing energy away from any critical point. This modified model yields excellent fits to $L(\omega)$ without introducing excitonic effects at E_1 , but still does not yield good fits to the derivatives of $L(\omega)$, as is shown in Fig. 5. Furthermore, it still contains most of the unphysical features of the original model of Adachi.

The above brief discussion of the existing models for the optical dielectric function of semiconductors (the HO, CPPB, and composite models) clearly shows that each of these models is incomplete. It also provides a clear motivation for the development of a new model and a list of tests to be applied to any new model.

IV. THE NEW MODEL FOR THE DIELECTRIC FUNCTION

The model for $\epsilon(\omega)$ proposed here meets the need for a model which is capable of simultaneously accurately fitting $\epsilon(\omega)$ and its first three derivatives. As has been pointed out in Sec. III, no existing models can do that. Furthermore, unlike all existing models, it correctly describes the full analytic behavior of $\epsilon(\omega)$, with no unphysical features.

The development of this model contains two basic steps. The first is to find a simple functional representation for $W_{cv}(E) = P_{cv}(E)^2 J_{cv}(E)$. This representation must fully satisfy the CP behavior of $W_{cv}(E)$ with no artificial cutoffs, it must be capable of accurately mimicking $W_{cv}(E)$ for any zinc-blende semiconductor, and it must allow the integral in Eq. (2.9), which gives Lorentzian line broadening, to be performed analytically. Finally, it must allow one to find a good analytic approximation to the integral in Eq. (2.10), which gives Gaussian line broadening. Given such an analytic representation for $W_{cv}(E)$, the second step is to analytically perform the integrals in Eqs. (2.9) and (2.10), even if only approximately for the case of Eq. (2.10). We find a final expression for $\epsilon(\omega)$ which combines the results for Lorentzian and Gaussian broadening, with a parameter which gives the Gaussian fraction of the broadening at each critical point.

We develop our model for $\epsilon(\omega)$ starting from the fundamental equations (2.9) and (2.10). First, we consider the analytic structure of $W_{cv}(E)$. Because $P_{cv}(E)$ is a slowly varying function of E with no singularities, we need only consider the analytic structure of $J_{cv}(E)$. In general, in three dimensions there are four possible types of critical points. They are designated M_0, \dots, M_3 , with the subscript denoting the number of negative eigenvalues of the effective-mass tensor at a given critical point. Each type has a square-root singularity.

For zinc-blende semiconductors the conduction and valence bands are very nearly parallel for long distances along symmetry lines through many of the M_1 and M_2

critical points in the optical region of energies. This means that one eigenvalue of the inverse effective mass is essentially zero at those critical points. For that reason, many authors have treated 3D M_1 and M_2 critical points as 2D M_0 and M_2 points, respectively. However, even if two bands are exactly parallel for a long distance along some symmetry line, producing a 2D critical point, *there must also exist a 3D critical point of the same energy at that point in \mathbf{k} space where the two bands become non-parallel*. This situation is shown schematically in Fig. 6.

We have found that an excellent fit to $[L(\omega_j)]_{\text{num}}^{(n)}$ can be found using the theoretical values given for the phase angles θ_j in Eq. (3.5) if one assumes a superposition of 2D and 3D critical points at the 3D M_1 and M_2 CP energies. On the other hand, if one assumes only 2D or only 3D critical points at those energies one must treat the θ_j as free parameters in order to obtain even a rather good fit.³⁹ The assumption of only 2D critical points at those energies gives a somewhat better fit than the assumption of only 3D critical points. These results are physically reasonable, given the situation shown in Fig. 6.

The assumption of both a 2D and a 3D CP at many 3D M_1 and M_2 critical points obviates the need for artificial cutoffs in $J_{cv}(E)$. Because it gives a discontinuity in $J_{cv}(E)$ as well as a square-root singularity at 3D M_1 and M_2 critical points, it allows the total $J_{cv}(E)$ for any given pair of conduction and valence bands to be written as the sum of independent contributions from the three following types of energy ranges: (1) the range of energies between any pair of connected 3D M_0 and M_1 points (type I), (2) the range between any pair of connected 3D M_1 and M_2 points (type II), and (3) the range between any pair of connected 3D M_2 and M_3 points (type III). The resultant $J_{cv}(E)$ is shown schematically in Fig. 7 for the simple case in which there is only one region of each type. For a zinc-blende semiconductor, one needs to consider at least three pairs of bands—the heavy-hole, light-hole, and split-off valence bands, each paired with the lowest conduction band. For each of these pairs of bands one must have at least one M_0 , two M_1 , two M_2 , and one M_3 critical point, with some degeneracies or near degeneracies. The critical points and their connections are discussed in the next section for the case of GaAs; that case is representative of most zinc-blende semicon-

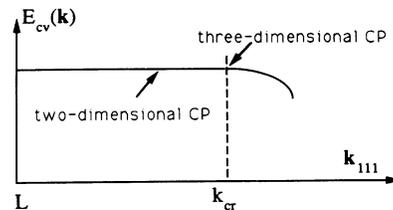


FIG. 6. Schematic figure showing 2D and 3D critical points at the same energy, as for instance at $E_1(\Lambda)$ along the Λ line. $E_{cv}(\mathbf{k})$ is very nearly constant as one moves away from L , until one reaches k_{cr} , at which point $E_{cv}(\mathbf{k})$ begins to decrease approximately as $(k - k_{cr})^2$, giving a 3D CP at k_{cr} .

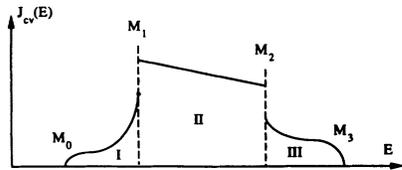


FIG. 7. Schematic drawing of $J_{cv}(E)$ for the simple case in which there is only one region of each type (simple cubic structure), but with both 3D and 2D critical points at E_1 and E_2 .

ductors, at least over the optical range of energies.

The contribution to $W_{cv}(E)$ from each of the three types of ranges above can be written as the product of a function which gives the correct CP analytic properties and a slowly varying analytic function of E which can be well approximated by a low-order polynomial. Specifically, those contributions can be taken to have the following forms:

$$W_I(E) = \sqrt{E - E_0} \times [p_I(E) - q_I(E) \sqrt{E_1 - E}], \quad (4.1)$$

$$W_{II}(E) = p_{II}(E), \quad (4.2)$$

$$W_{III}(E) = \sqrt{E_3 - E} \times [p_{III}(E) - q_{III}(E) \sqrt{E - E_2}]. \quad (4.3)$$

Here, p_I , p_{II} , p_{III} , q_I , and q_{III} are analytic functions of E which are well approximated over the optical range of energies by low-order polynomials, $p_\nu(E) = \sum_n p_{n,\nu} E^n$, $q_\nu(E) = \sum_n q_{n,\nu} E^n$ for $\nu=I,II,III$. The total $W_{cv}(E)$ then has the form

$$W_{cv}(E) = \sum_I W_I(E) + \sum_{II} W_{II}(E) + \sum_{III} W_{III}(E), \quad (4.4)$$

where the sums over I, II, and III go over all connected pairs of M_0 and M_1 critical points, connected pairs of M_1 and M_2 critical points, and connected pairs of M_2 and M_3 critical points, respectively.

Having found an analytically correct model for $W_{cv}(E)$, it remains only to perform the integration in Eq. (2.9) or Eq. (2.10) after substituting Eq. (4.4) into those equations. For the case of Lorentzian line broadening, Eq. (2.9), the integration can be performed analytically in each range upon approximating the slowly varying analytic functions $p_\alpha(E)$ and $q_\alpha(E)$ by polynomials of degree of five or less. The integrations are carried out in the Appendix, giving the result

$$\epsilon(\omega) = 1 - \frac{8\pi\hbar^2 e^2}{m^2} \sum_{c,v} \sum_n \left(\sum_I (p_n H_n - q_n F_n)_I + \sum_{II} (p_n G_n)_{II} + \sum_{III} (p_n K_n - q_n F_n)_{III} \right)_{cv}, \quad (4.5)$$

where the functions H_n , F_n , G_n , and K_n are defined in the Appendix and the summations over I, II, and III are defined as in Eq. (4.4).

The case of Gaussian line broadening is not as simple because the integral in Eq. (2.10) cannot be performed analytically. However, we have been able to find a functional representation which accurately mimics the numerical integrals given by substituting Eq. (4.4) into Eq. (2.10) with the $p_\alpha(E)$ and $q_\alpha(E)$ given by low-order polynomials in E . The substitution of the quantity

$$D_j = \Gamma_j \exp \left[-\alpha_j \left(\frac{\hbar\omega - E_j}{\Gamma_j} \right)^2 \right] \quad (4.6)$$

for Γ_j in Eqs. (A9)–(A14), for $j = i$ or f , leads to analytic functions H_n , F_n , G_n , and K_n which accurately mimic the numerical results given by Eqs. (A1)–(A4) for the Gaussian case, for appropriate values of α_j . Since $\alpha_j = 0$ gives the exact result for the Lorentzian case, varying the value of α_j interpolates between the cases of Lorentzian broadening and Gaussian broadening. The value of α_j which most closely mimics the exact results of Gaussian broadening is not exactly the same for the four functions H_n , F_n , G_n , and K_n and depends slightly on the value of n , but is approximately 0.2 in all cases. Therefore, we use the same value for α_j for all of the H_n , F_n , G_n , and K_n which occur at the CP j .

This completes the specification of our model. This model overcomes the deficiencies embodied in previous models, (i) yielding excellent fits to $\epsilon(\omega)$ and its derivatives with respect to ω , (ii) fully incorporating the analytic form of $J_{cv}(E)$ between any pair of valence and conduction bands, as well as the values of E_j and Γ_j , and (iii) allowing one to go beyond the approximation of Lorentzian broadening. If the α_j are treated as free parameters for the more important critical points and the $p_\nu(E)$ and $q_\nu(E)$ are approximated by first- or second-order polynomials, the model yields excellent simultaneous fits to $L(\omega_j)$ and its first three derivatives, as is shown in the next section.

V. APPLICATION TO GaAs

In applying the model developed in Sec. IV above to GaAs, fits were made to the same set of spectral data, $L(\omega_j)$, as was fit in Sec. III using previously published models. The data, which covers the range from 1.0 to 6.0 eV, was determined as follows. Values for both the refractive index n and the extinction coefficient k between 1.5 eV and 6 eV were taken from the *Handbook of Optical*

*Constants of Solids*⁴⁰ from the spectroscopic ellipsometry data of Theeten and co-workers.⁴¹ This is the same data as that later tabulated by Aspnes and Studna.⁴² The data was given in steps of 20 meV up to 3.5 eV and in steps of 40 meV from 3.5 eV to 6 eV. Values for k were also taken from Ref. 40 between 1.37 eV and 1.5 eV in steps of 5 meV and 10 meV, but from the transmission measurements by Casey, Sell, and Wecht.⁴³ Values for n between 1.0 eV and 1.5 eV in steps of 20 meV were calculated from the oscillator formula, with parameters determined by fitting the measured refractive index by Pikhin and Yas'kov⁴⁴ and Marple.⁴⁵ All the spectral data for n and k were reorganized in steps of 20 meV, with the spline method used where interpolation was needed. The values of $L(\omega_j)$ were then calculated from the formula $L(\omega_j) = [n(\omega) + ik(\omega)]^2$ between 1.0 eV and 6 eV, assuming $k = 0$ between 1.0 eV and 1.37 eV. The glitches in the second and third derivatives of $L(\omega_j)$ seen at 1.37, 1.5, and 3.5 eV in our figures result from the joining of different sets of data or of data from regions having different spacings in energy; they have no physical significance.

As is discussed in Sec. IV above, the first step in applying our model to fit the spectral data is the determination of the critical-point structure of the joint density of states, $J_{cv}(E)$, over a range in energy which includes the range over which the data is to be fit. A band structure⁴⁶ for GaAs showing the first three valence bands and first two conduction bands is given in Fig. 8. The first three valence bands are designated by $v1$, $v2$, and $v3$ and the first two conduction bands are designated by $c1$ and $c2$. Using Schottky-barrier electroreflectance at 4.2 K, Aspnes and Studna⁴⁷ identified thirteen critical points in $J_{cv}(E)$ lying below 5.6 eV and having nonzero matrix elements for optical transitions. From their identification of E_2 , $E_2 + \Delta_2$, and $E_2 + \Delta'_2$ transitions at 4.94, 5.01, and 5.34 eV, respectively, one would also expect

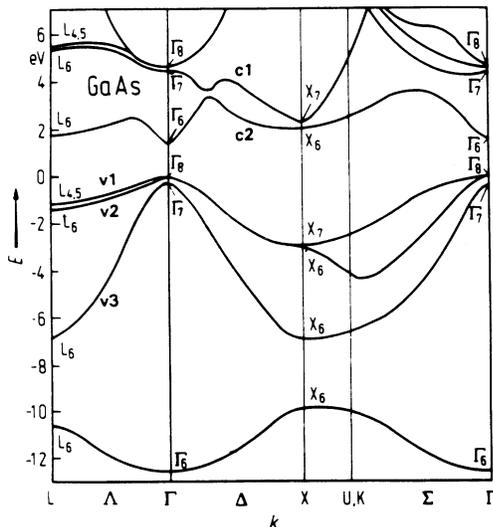


FIG. 8. Band structure of GaAs from Ref. 46 showing the bands we label $v1$, $v2$, $v3$, $c1$, and $c2$.

an $E_2 + \Delta_2 + \Delta'_2$ transition at 5.41 eV, which was not seen. Using numerically differentiated SE data obtained at 22 K, Logothetidis *et al.*⁴⁸ observed the same critical-point transitions, differing only in their identification of one transition. They assigned the transition labeled as $E_2 + \Delta'_2$ in Ref. 47 to a general point P in the ΓXUL symmetry plane; that assignment does not require any additional critical point.

In principle, we should include all thirteen of these critical points in our model. However, only seven critical points are clearly evident in the room-temperature data which we fit. Therefore, in order to reduce the number of parameters in our model, we include only seven critical points in our model. Those include the four lowest-energy points— $E_0(\Gamma)$, $E_0(\Gamma) + \Delta_0(\Gamma)$, $E_1(\Lambda)$, and $E_1(\Lambda) + \Delta_1(\Lambda)$. Three of the remaining nine critical points identified in Refs. 47 and 48 were selected as being most important, based on the following rules: (1) The contribution of an M_1 or M_2 CP to any derivative of $\epsilon(\omega)$ is more important than that of any connected M_0 or M_3 critical point, (2) the strength of the CP line shape is stronger at any CP E_j than at $E_j + \Delta_j$ or $E_j + \Delta'_j$, and (3) the strength of the CP line shape at any given symmetry point in the Brillouin zone decreases with increasing transition energy, due to a decrease in the matrix element for optical transitions. These rules follow both from empirical observation and from theoretical calculations. Based on these three rules and on the results of Refs. 46 and 47, we selected the critical points $E'_0(\Delta)$, $E_2(X)$, and $E_2(\Sigma)$ for inclusion in our model, in order of increasing energy. Table III shows the connections between the seven critical points selected and the type of each critical point.

Figure 9 is a schematic diagram showing the way in which the total $J(E) \equiv \sum_{cv} J_{cv}(E)$ or $W(E) \equiv \sum_{cv} W_{cv}(E)$ is constructed for GaAs using these critical points. Figures 9(a), 9(b), 9(c), and 9(d) show the CP structure of the contributions to $W_{cv}(E)$ which arise from the pairs of bands ($v1$, $c1$), ($v2$, $c1$), ($v3$, $c1$), and ($v1$, $c2$), respectively, considering only the seven critical points which we include in our model for GaAs. Because the $E'_0(\Gamma)$ CP is not visible in $L(\omega_j)$ or in any of its numerical derivatives and thus is not included in our model for the room-temperature dielectric function of GaAs, the solid curve in Fig. 9(c) is replaced in our model by

TABLE III. The interconnections between the seven critical points used in fitting $L(\omega_j)$ for GaAs and the type of each critical point.

Band pair	Critical point	3D type	Connected higher-energy CP's below 6 eV
$v1 \rightarrow c1$	$E_0(\Gamma)$	M_0	$E_1(\Lambda)$, $E_2(X)$
	$E_1(\Lambda)$	M_1	$E_2(\Sigma)$
	$E_2(X)$	M_1	$E_2(\Sigma)$
	$E_2(\Sigma)$	M_2	
$v2 \rightarrow c1$	$E_0(\Gamma)$	M_0	$E_1(\Lambda) + \Delta_1(\Lambda)$
	$E_1(\Lambda) + \Delta_1(\Lambda)$	M_1	
$v3 \rightarrow c1$	$E_0(\Gamma) + \Delta_0(\Gamma)$	M_0	
$v1 \rightarrow c2$	$E'_0(\Delta)$	M_1	

the dashed curve, which shows no $E'_0(\Gamma)$ CP. Figure 9(e) shows the form of the resultant total $W(E)$, which is split into six independent segments. The first of these contains an infinite discontinuity in slope at $E_0(\Gamma) + \Delta_0(\Gamma)$; all of the others are continuous. The first segment is the sum of one curve (I) between $E_0(\Gamma)$ and $E_1(\Lambda)$ having the form of Eq. (4.1) and a second curve (I') between $E_0(\Gamma) + \Delta_0(\Gamma)$ having the same form, but with $q_{I'}(E)$ set to zero. The second, third, and fourth segments all have

the form

$$W_\nu(E) = p_\nu(E) - q_\nu(E) \sqrt{E_\nu - E}, \quad (5.1)$$

with $E_\nu = E_1(\Lambda) + \Delta_1(\Lambda)$, $E'_0(\Delta)$, and $E_2(X)$, respectively. The fifth segment has the form of Eq. (4.2), and the sixth has the form

$$W_{VI}(E) = p_{VI}(E) - q_{VI}(E) \sqrt{E - E_2(\Sigma)}. \quad (5.2)$$

The above analysis leads to the equation

$$\begin{aligned} \epsilon(\omega) = 1 - \frac{8\pi\hbar^2 e^2}{m^2} \sum_n [(p_n H_n - q_n F_n)_{I} + (p_n H_n)_{I'} + (p_n G_n - q_n K_n)_{II} + (p_n G_n - q_n K_n)_{III} \\ + (p_n G_n - q_n K_n)_{IV} + (p_n G_n)_{V} + (p_n G_n - q_n H_n)_{VI}] + \sum_n b_n E^n, \end{aligned} \quad (5.3)$$

where the last term gives the contribution to $\epsilon_1(\omega)$ from the critical points above 6 eV. A fitting to $L_2(\omega_j)$ and/or any or all of its first three numerical derivatives suffices to determine all of the free parameters in this equation except for the b_n 's. It is more appropriate to use $L_2(\omega_j)$ to determine the parameters than to use $L_1(\omega_j)$ because $L_2(\omega_j)$ is much more directly related to $W(E)$. Thus, we determine from $L_2(\omega_j)$ all of the CP parameters and the coefficients in the polynomials $p(E)$ and $q(E)$ in each energy range.

Fittings to the n th derivative are done by minimizing

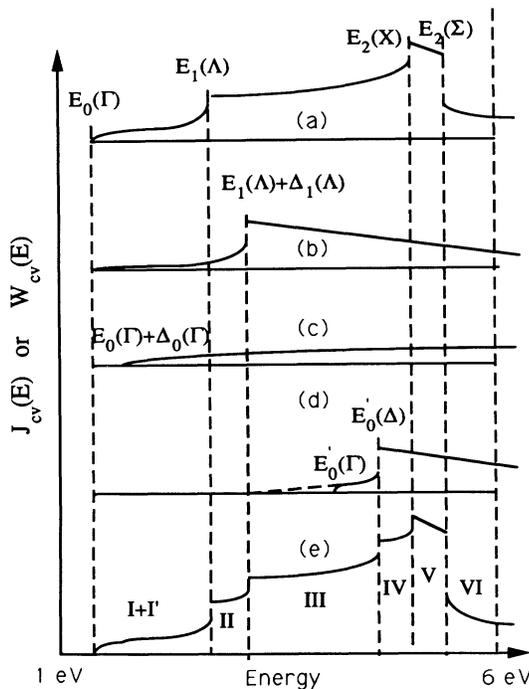


FIG. 9. Schematic drawing of the contributions of the pairs of bands (a) (v_1, c_1), (b) (v_2, c_1), (c) (v_3, c_1), and (d) (v_1, c_2) to (e) the total $J(E)$ or $W(E)$. The roman numerals label the different regions in energy which are separated by discontinuities in $J(E)$ or $W(E)$.

the rms fractional deviation σ_n defined in Eq. (3.6) with respect to the free parameters in our model for $\epsilon_2(\omega)$, with uniform weighting of all points in energy. In performing simultaneous fits to $L_2(\omega_j)$ and its first three numerical derivatives, we minimize the overall rms fractional deviation σ defined in Eq. (3.7). The fittings are performed using the method described in Ref. 7. In performing simultaneous fits to $L_2(\omega_j)$ and its first three derivatives, faster convergence is obtained if the CP parameters are first obtained from a fit to $[L_2(\omega_j)]_{\text{num}}^{(3)}$, with all of the polynomials reduced to constants, before performing the full simultaneous fits.

It follows from Eq. (5.3) that the specification of $\epsilon_2(\omega)$ within our model requires the specification of twelve polynomials in energy as well as the specification of the CP parameters at the seven critical points we have chosen to include in the model. If each polynomial were replaced by a constant and if all of the α_j were set equal to zero (Lorentzian broadening) or 0.2 (approximately Gaussian broadening) or some intermediate value, this would leave 26 free parameters for the fitting of $L_2(\omega_j)$ and its derivatives. Allowing $p(E)$ to become a quadratic function of E and $q(E)$ to become linear in E in region I, but suppressing the unimportant 3D contribution at $E_2(\Sigma)$ only increases the number of free parameters to 28. This is the same number of parameters as was used in the CPPB model fits to $\epsilon(\omega)$ shown in Figs. 2 and 3 and is 8 fewer than were used in the HO model as improved by Terry.¹⁹ In this version of our model, $\epsilon(\omega)$ is a nonlinear function of only 14 of the parameters as compared to 14, 27, 21, and 13 of the parameters, respectively, in the HO models of Erman *et al.*¹⁸ and of Terry,¹⁹ the CPPB model, and the model of Jenkins.²²

Fits obtained from our model using only this minimum number of parameters, with all the α_j 's set to zero, are shown in Figs. 10 and 11. The values found for the parameters E_j and Γ_j for each critical point and the rms fractional errors σ_n and σ are shown in Table IV. All of the critical point energies except $E_2(X)$ and all of the line widths except $\Gamma(E_0)$ and possibly $\Gamma(E_0 + \Delta_0)$ are determined with confidence. The uncertainty of $E_2(X)$ is due to the fact that its structure is located between two dominant critical-point structures and its strength in the

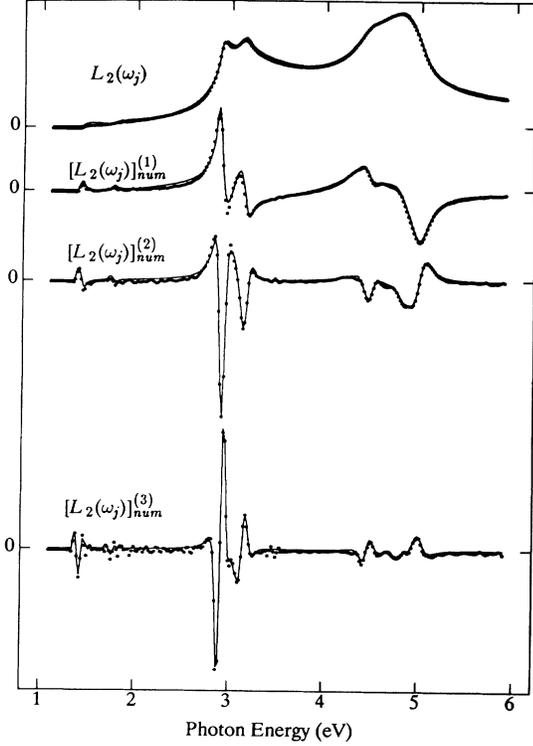


FIG. 10. Separate fits of our model to $L_2(\omega_j)$ and its first three numerical derivatives, assuming only Lorentzian broadening for each CP and using 28 parameters, the same number as in the fits of Figs. 2 and 3. The dots show the imaginary parts of $[(\hbar\omega_j)^2 L_2(\omega_j)]_{\text{num}}^{(n)}$. The solid lines show the resultant fits.

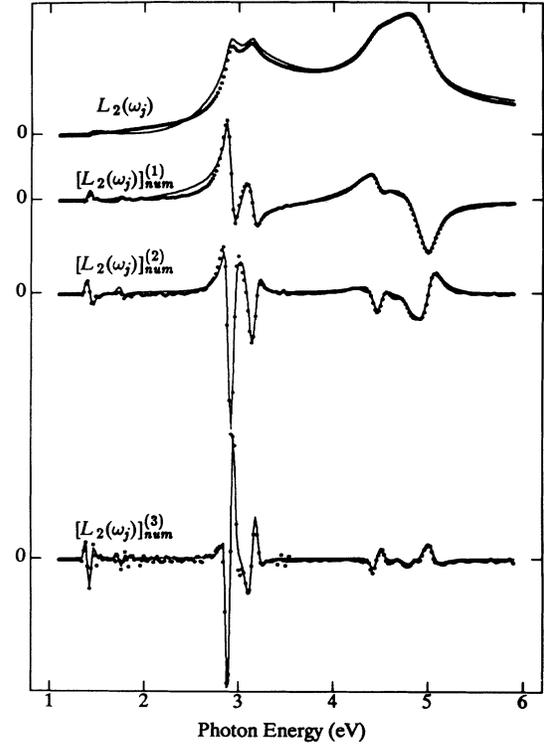


FIG. 11. Simultaneous fit of our model to $L_2(\omega_j)$ and its first three numerical derivatives, assuming only Lorentzian broadening for each CP and using 28 parameters, the same number as in the fits of Figs. 2 and 3.

TABLE IV. The dimensionalities assigned each CP and the values of the CP energies and line widths of GaAs obtained by fitting spectral data with the new model given by Eq. (5.3). The values are given in units of eV. The first four columns of values were found by fitting the numerical n th derivatives of the spectral data with respect to photon energy with the numerical n th derivatives of the new model for $\epsilon(\omega)$, for $n = 0, \dots, 3$. The values in the last column were found by simultaneously fitting the data and its first three numerical derivatives. Values for $E_0(\Gamma) + \Delta_0(\Gamma)$, $\Gamma[E_0(\Gamma)]$, and $\Gamma[E_0(\Gamma) + \Delta_0(\Gamma)]$ were determined by fitting $[L_2(\omega_j)]_{\text{num}}^{(3)}$ and were fixed at those values 1.745 eV, 0.005 eV and 0.009 eV, respectively, for the other fittings.

Critical point	d_j	Order of the derivative, n				Simultaneous fit
		0	1	2	3	
$E_0(\Gamma)$	3	1.423	1.414	1.408	1.411	1.412
$E_1(\Lambda)$	2	2.894	2.894	2.916	2.928	2.910
$E_1(\Lambda) + \Delta_1(\Lambda)$	2	3.175	3.175	3.176	3.175	3.170
$E'_0(\Delta)$	2	4.476	4.476	4.478	4.472	4.483
$E_2(X)$	2	4.714	4.742	4.794	4.751	4.769
$E_2(\Sigma)$	2	4.994	4.994	5.004	4.996	5.000
$\Gamma[E_1(\Lambda)]$	2	0.031	0.032	0.031	0.042	0.029
$\Gamma[E_1(\Lambda) + \Delta_1(\Lambda)]$	2	0.065	0.066	0.079	0.079	0.073
$\Gamma[E'_0(\Delta)]$	2	0.087	0.084	0.084	0.089	0.086
$\Gamma[E_2(X)]$	2	0.136	0.129	0.106	0.114	0.120
$\Gamma[E_2(\Sigma)]$	2	0.131	0.131	0.119	0.121	0.129
rms fractional error		11%	9.8%	9.4%	11.2%	13.7%

derivative spectra is much smaller than that of the other two. The line width $\Gamma(E_0)$ cannot be determined from these fittings because it is much less than the spacing between successive data points in $L_2(\omega_j)$; therefore, we have fixed it at the estimated value of 5 meV. $\Gamma(E_0 + \Delta_0)$ is also less than the spacing between data points, but is large enough to be fairly well determined. The fits shown in Figs. 10 and 11 are better than those shown in Figs. 1–3 obtained from the HO and CPPB models and are vastly superior to the 20-parameter fit of Jenkins shown in Fig. 5. The superiority of the fits obtained using our model as opposed to those obtained using the CPPB model with the same number of free parameters is confirmed by a comparison of Tables II and IV. The use of our model yields smaller rms fractional errors in the fits and gives values for the CP energies and line widths which are more nearly independent of n . Furthermore, even this simplest version of our model preserves the correct analytic structure of $\epsilon(\omega)$ for all of the critical points considered, unlike the other models. Finally, our model permits the fits shown in Figs. 10 and 11 to be substantially improved with almost no increase in the difficulty of fitting. This can be done by allowing more of the $p(E)$ and $q(E)$ to be low-order polynomials in E , rather than being approximated by constants. This leads to no increase in the number of parameters with respect to which $\epsilon(\omega)$ is nonlinear.

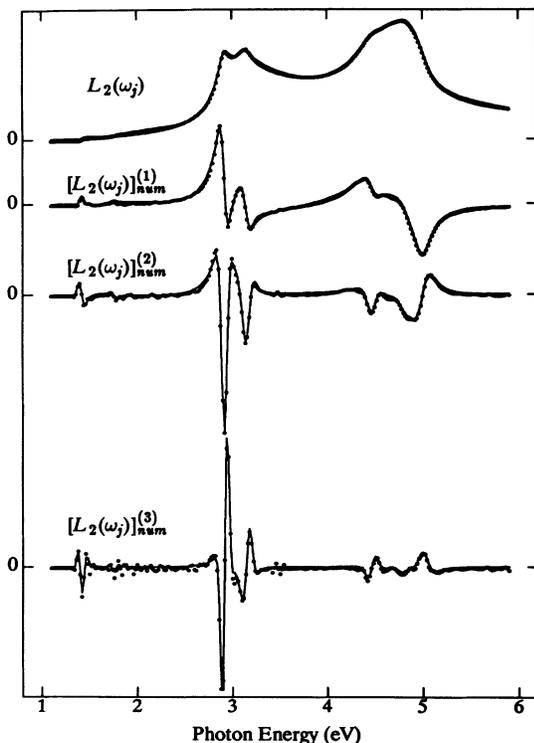


FIG. 12. Simultaneous fit of our model to $L_2(\omega_j)$ and its first three numerical derivatives, similar to that of Fig. 11, but with five more free parameters with respect to which $\epsilon(\omega)$ varies linearly. The dots show $L_2(\omega_j)$ and its first three numerical derivatives; the solid lines show the fit.

Figure 12 shows the improved simultaneous fit to $L_2(\omega_j)$ and its first three derivatives obtained by allowing five more free parameters in the determination of the p_n 's and q_n 's. In region I $q(E)$ was allowed to become a quadratic function of E and in regions I', II, III, and IV the $p(E)$'s were allowed to become linear functions of E . This reduced the overall rms fractional error in our fit, σ , from 13.8% to 8.6%. The contribution to σ from noise in the derivatives of $L_2(\omega_j)$ is approximately 5%, so this is an excellent fit. Almost all of the small remaining inaccuracy in the fit above 2 eV can be eliminated by further increasing the number of free parameters in the polynomials $p(E)$ and $q(E)$, keeping the Lorentzian approximation.

Our simultaneous fit to $L_2(\omega_j)$ and its first three derivatives can be somewhat further improved by allowing the parameters α_j to become nonzero, i.e., by allowing an admixture of Gaussian and Lorentzian line broadening. Figure 13 shows the result of a simultaneous fit to $L_2(\omega_j)$ and its first three derivatives in which α_j was treated as a free parameter at the four lowest-energy critical points. This added four parameters to the 33 used in the fit shown in Fig. 12. The values found for the α_j 's were as follows: $\alpha(E_0) = 0.18$, $\alpha(E_0 + \Delta_0) = 0.15$, $\alpha(E_1) = 0.002$, $\alpha(E_1 + \Delta_1) = 0.02$. The values of the line widths Γ_j were found to increase slightly with increasing values of the α_j . The final values found for the

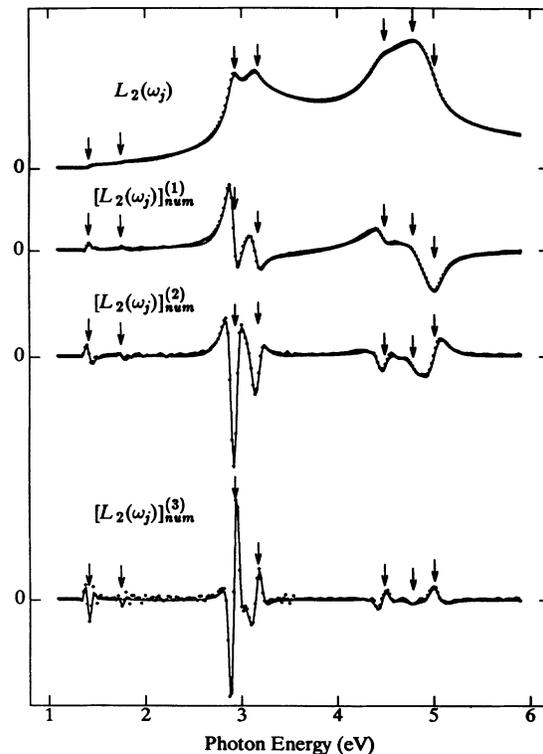


FIG. 13. The simultaneous fit of our model to $L_2(\omega_j)$ and its first three numerical derivatives, allowing a mixture of Lorentzian and Gaussian broadening for the four lowest-energy critical points. The arrows show the position of the seven critical points used in the fit.

TABLE V. Values of the CP energies E_j , line widths Γ_j , and Gaussian broadening factors α_j , determined through simultaneously fitting $L_2(\omega_j)$ and its three derivatives. The zero values were fixed to reduce the number of free parameters.

Critical point	$E_0(\Gamma)$	$E_0(\Gamma) + \Delta_0(\Gamma)$	$E_1(L)$	$E_1(L) + \Delta_1(L)$	$E'_0(\Delta)$	$E_1(X)$	$E_1(K)$
E_j (eV)	1.411	1.748	2.930	3.171	4.478	4.811	5.003
Γ_j (meV)	5.0	8.8	43.9	77.5	76.2	111.3	125.8
α_j	0.18	0.15	0.002	0.018	0	0	0

CP parameters E_j , Γ_j , and α_j are given in Table V. The final values found for the p_n and q_n in each region are given in Table VI. In this fit σ was reduced to 6.5 %. Taking the values of the CP parameters E_j , Γ_j , and α_j from this fit and allowing only the 19 coefficients in the polynomials $p(E)$ and $q(E)$ to vary, we found an excellent fit to $L_2(\omega_j)$ with an rms fractional error, σ_0 , of only 0.4%. The fit is not shown because the curves for $L_2(\omega_j)$ and $\epsilon_2(\omega_j)$ fall exactly on top of one another. Figure 14 shows the separate contributions to $L_2(\omega_j)$ of each of the seven different regions we consider.

Unfortunately, because $\epsilon(\omega)$ is rather insensitive to the values of the α_j 's, their values from our simultaneous fits to $L_2(\omega_j)$ and its derivatives should not be trusted. Better values could be obtained by fitting only the second and third derivatives of $L_2(\omega_j)$; however, even those values are rather uncertain, and those fits are difficult because of the highly nonlinear dependence of $L_2(\omega_j)$ on the α_j 's. The best determination of the α_j 's using this model would be to fit absorption data for $\hbar\omega < E_0$, with all parameters except the α_j 's determined from our simultaneous fittings of $L_2(\omega_j)$ and its derivatives. That procedure leads to a value of approximately 0.2 for $\alpha(E_0)$ (completely Gaussian broadening) and an approximate lower limit of 0.1 for α at $E_0 + \Delta_0$, E_1 and $E_1 + \Delta_1$, lending support to our previous determination⁷ that the line broadening in GaAs at room temperature is primarily Gaussian. However, the best determinations of the Gaussian versus Lorentzian character can be made from fits to the second and third derivatives of $L(\omega_j)$ within the CPPB model, within which Gaussian broadening can be simulated more easily and more accurately.²³

The effect of allowing the α_j 's to be nonzero is seen most easily at energies $\hbar\omega \leq E_0$. Figure 15 shows a blowup of the range from 1.2 eV to 1.7 eV for three fits to $L_2(\omega_j)$. Each fit was made over the entire range from 1 eV to 6 eV, and only the 19 coefficients in the polynomials

$p(E)$ and $q(E)$ were allowed to vary. In the first fit (the dotted line) the CP parameters were those obtained in the Lorentzian fit shown in Fig. 12. The second (the dashed line) was the best overall fit to $L_2(\omega_j)$ alone with $\sigma_0 = 0.4$ %. The third (the solid line) was similar to the second, but with $\alpha_j=0.2$ at E_0 and $E_0 + \Delta_0$ and $\alpha_j=0.1$ at all higher critical points. This figure clearly shows that one must assume at least partially Gaussian broadening to obtain a reasonable fit to $L_2(\omega_j)$ for $\hbar\omega \leq E_0$. Very close to E_0 , it also shows the effect of excitons, which were not included in our model.

After the CP parameters and the p_n 's and q_n 's are determined by fitting $L_2(\omega_j)$ and its derivatives, $\epsilon_1(\omega)$ become a function only of the coefficients in the polynomial $b = \sum_n b_n E^n$ in Eq. (5.3). Figure 16 shows the simultaneous fit to $L_1(\omega_j)$ and its first three derivatives obtained using the parameter values from the fit to $L_2(\omega_j)$ shown in Fig. 13 and allowing b to be a quadratic function of E . The fit has an overall rms fractional error $\sigma=8.7\%$, 2.2% greater than for the equivalent fit to $L_2(\omega_j)$ and its first three derivatives. A large part of the increased error in fitting occurs near E_0 , where $L_1(\omega_j)$ does not show the proper CP behavior. This is because $n(\omega_j)$, and hence to a large extent $L_1(\omega_j)$, was determined from a HO model in the absence of experimental data. Thus, we believe that the large error in fitting the derivatives of $L_1(\omega_j)$ near E_0 arises from errors in $L_1(\omega_j)$ itself, not from any problem with our model.

Figure 17 compares $L_1(\omega_j)$ with the functions $\epsilon_1(\omega)$ obtained from our best fit of $\epsilon_2(\omega)$ to $L_2(\omega_j)$ for two cases: (1) the case in which $b = 0$, a zero-parameter calculation of $\epsilon_1(\omega)$, and (2) the case in which $b(E)$ is allowed to be a linear function of E , a two-parameter fit to $L_1(\omega_j)$. The first case gives an rms fractional error $\sigma_0 = 1.5\%$; the second gives $\sigma_0 = 1.15\%$. Allowing $b(E)$ to become a cubic function of E decreases σ_0 only an additional 0.1%.

TABLE VI. Values of the coefficients p_n and q_n determined in each region through simultaneously fitting $L_2(\omega_j)$ and its three derivatives, which are multiplied by $8\pi\hbar^2 e^2/m^2$. The zero values were fixed to reduce the number of free parameters.

Region	p_0	p_1	p_2	q_0	q_1	q_2
I	13 795.3	-7820.7	1077.2	8077.0	-3243.4	239.9
I'	21.2	-11.8	0.0	0.0	0.0	0.0
II	2 427.2	-711.6	0.0	593.2	0.0	0.0
III	903.8	-157.6	0.0	299.9	0.0	0.0
IV	270.4	0.0	0.0	210.7	0.0	0.0
V	258.9	0.0	0.0	0.0	0.0	0.0
VI	88.4	0.0	0.0	17.2	0.0	0.0

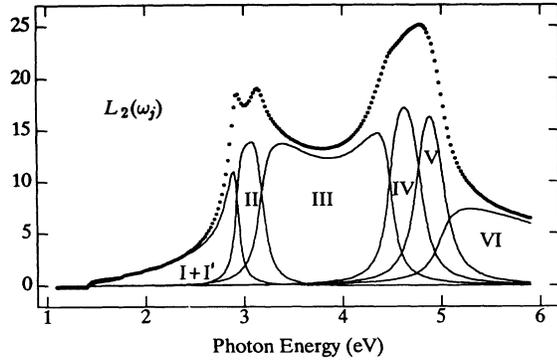


FIG. 14. Contribution of each of our six regions to $\epsilon_2(\omega)$. The dotted line shows $L_2(\omega_j)$, which is indistinguishable from the total $\epsilon_2(\omega)$.

VI. DISCUSSION AND CONCLUSIONS

We have shown that the new model proposed here is more generally valid than the HO model as used by Erman *et al.*¹⁸ or as improved by Terry,¹⁹ the CPPB model^{37,38} or the composite model of Adachi²⁰ as improved by Jenkins.²² First, our model is applicable over the entire range of photon energies, below and above the lowest band gaps. Second, it gives the correct analytic structure of $J_{cv}(E)$ or $W_{cv}(E)$ at every critical point considered, unlike other models, and it exactly satisfies the Kramers-Kronig relations. Third, it allows one to go beyond the approximation of Lorentzian broadening, which is known to be incorrect for elements and compounds at room temperature. Also, it contains fewer parameters with respect to which $\epsilon(\omega)$ is nonlinear than

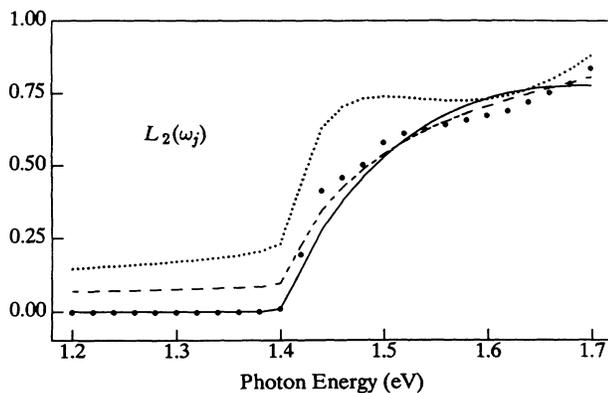


FIG. 15. Blowup of the region near E_0 for three different fits of our model to $L_2(\omega_j)$. Each fit was performed over the entire energy range from 1 eV to 6 eV. The bullets show $L_2(\omega_j)$. The dotted curve shows the Lorentzian fit of Fig. 11. The dashed curve shows the best fit to $L_2(\omega_j)$, using the CP parameters of the fit shown in Fig. 13. The solid curve shows the effect of choosing $\alpha_j = 0.1$ (50% Gaussian broadening) rather than $\alpha_j = 0$ at the three higher-energy critical points as well as at $E_1(\Lambda)$ and $E_1(\Lambda) + \Delta_1(\Lambda)$.

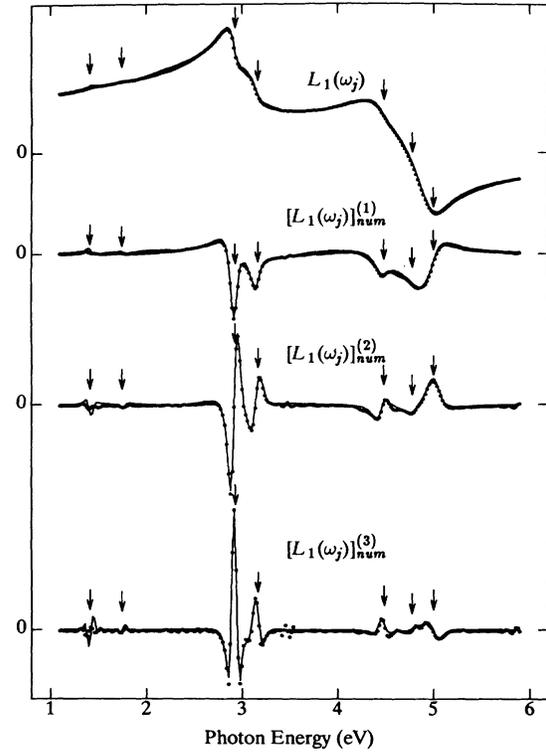


FIG. 16. Simultaneous fit of our model to $L_1(\omega_j)$ and its first three numerical derivatives. The parameters obtained in the simultaneous fit to $L_2(\omega_j)$ and its first three numerical derivatives were fixed; the only free parameters were b_0 , b_1 , and b_2 in Eq. (5.3). The arrows indicate the positions of the seven critical points used in the fitting.

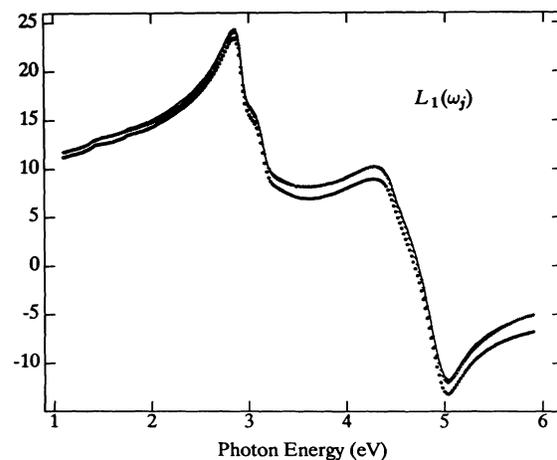


FIG. 17. Zero-parameter calculated $\epsilon_1(\omega)$ and two-parameter fit to $L_1(\omega_j)$. The plus signs show $L_1(\omega_j)$. The bullets show $\epsilon_1(\omega)$ as calculated from the best fit to $L_2(\omega_j)$ with no free parameters (all b_n 's equal to zero). The solid line shows $\epsilon_1(\omega)$ as calculated allowing b_0 and b_1 in Eq. (5.3) to be free parameters.

any model except the seven-peak HO model of Erman *et al.*, thus giving better convergence and less possibility of finding false minima in fittings. Finally, even within the Lorentzian approximation and with less than or approximately the same number of free parameters as in other models, it yields better fits to spectral data and gives more reliable values for CP parameters.

Not only is it the only model proposed thus far which yields simultaneous good fits to the dielectric function and its derivatives; it yields individual fits to the dielectric function and to its derivatives which are better than the best previous individual fits (although the fit to the dielectric function itself is not significantly better than the excellent fit obtained using the method of Terry.¹⁹

For all of these reasons, this model should provide an excellent interpolation scheme for predicting the dielectric function of semiconductor alloy series, given the measured dielectric function $L(\omega_j)$ at the end points of the alloy series and at one or two intermediate points.

It also should provide excellent predictions for $n(\omega)$ at all energies $\hbar\omega$ below the energy gap E_0 of any semiconductor, given $L(\omega_j)$ over a reasonable range in energy including E_0 . Finally, given $k(\omega_j)$ for a reasonable range of energies below E_0 , say 200–300 meV, it should provide a good prediction for $k(\omega)$ all the way down to $\omega = 0$. We have tested these expectations and have found them to be fulfilled; our results will be published elsewhere.

Having established the value of our model for $\epsilon(\omega)$, we must now ask if any new physics is manifest in the results of our fitting to $L(\omega_j)$ and its derivatives for GaAs. The answer is yes. Our model is unique in its ability to predict $W(E)$. We found to our surprise that the fitting of our model to $L(\omega_j)$ and its derivatives for GaAs yields *discontinuous drops* in $W(E)$ at the M_1 critical points $E_1(\Lambda)$, $E_1(\Lambda) + \Delta_1(\Lambda)$, and $E'_0(\Delta)$, as is shown in Fig. 18. We know from the analytic properties of $J_{cv}(E)$ that these discontinuities should be positive, rather than negative, at $E_1(\Lambda)$ and $E_1(\Lambda) + \Delta_1(\Lambda)$ and should be ≥ 0 at $E'_0(\Delta)$.⁴⁹ This is because these discontinuities arise from a two-dimensional M_0 critical point, as was discussed in Sec. IV and shown in Fig. 6. Thus the observed drops are puzzling. They remain essentially unchanged when one separately fits $L(\omega_j)$ of any of its first three derivatives. Therefore, they cannot result from a failure to include enough terms in the polynomials $p(E)$ and $q(E)$, because the fits to the second and third derivatives of $L(\omega_j)$ are essentially independent of those polynomials, so long as they are slowly varying functions. We find these drops to be essentially independent of the values of the α_j , which interpolate between Lorentzian and Gaussian broadening. These discontinuities can be made positive, as we know they must be, only by introducing empirical phase angles θ_j at those critical points, as is done within CPPB model. The values found for these phase angles must be independent of the form of broadening assumed and of all details of our model. They follow simply from the analytic properties of the critical points.

What can be the origin of such phase angles? It has been proposed³⁹ that they arise from excitonic effects, but one certainly does not expect excitons at room tem-

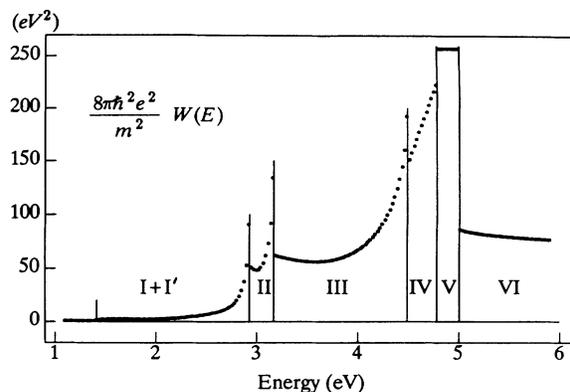


FIG. 18. The quantity $(8\pi\hbar^2 e^2/m^2)W(E)$ calculated from the fit shown in Fig. 13.

perature at the $E_1(\Lambda)$, $E_1(\Lambda) + \Delta_1(\Lambda)$, or $E'_0(\Delta)$ M_1 critical points. We also have found by numerical simulations that such phase angles cannot arise from thin (≤ 10 Å) surface overlayers or from surface roughness. Both of those effects strongly influence the measured $L(\omega_j)$ but have almost no effect on the phase of the measured $[L(\omega_j)]_{\text{num}}^{(3)}$ at the critical points. We suggest a very different explanation, which we propose to investigate in detail, namely, that the phase angles arise from the failure of SE to measure the true bulk dielectric function at energies far above E_0 . At those energies the penetration depth of light is small, so that SE probes only a thin layer of material near the surface. Both the presence of a built-in electric field in that region and the boundary condition that all electronic wave functions must vanish at a sample surface affect the reflection coefficients of light and hence the values of $L(\omega_j)$ calculated from SE data, especially when that light penetrates only a small distance into the sample under study. For semi-insulating or lightly doped samples the primary effect on $L(\omega_j)$ is to introduce an energy-dependent phase angle $\theta(E)$, intermixing the real and imaginary parts of the dielectric function.^{2-4,50,51} We have studied theoretically and experimentally the effect of built-in electric fields and of surface termination on the electroreflectance line shape⁵⁰ and have found that effect to be large. Previous authors have considered the effect of surface termination on the reflection coefficients, but obtained somewhat unphysical results because they ignored questions of wave-function coherence.^{2-4,51}

In principle one could use our model to determine a true bulk $\epsilon(\omega)$ in either of the two following ways: (1) One could introduce θ_j as a free parameter at each CP in the fitting of $L(\omega_j)$ and its derivatives, and then evaluate $\epsilon(\omega)$ by setting each θ_j equal to zero, keeping the values of all other parameters fixed. That would cancel out the effect of the surface termination of wave functions on the phase angles θ_j , replacing them by their physically correct value, zero. (2) One could find the ratio of $J_{cv}(E)$ just above each discontinuity to $J_{cv}(E)$ just below the discontinuity from a band-structure calculation and use those ratios as constraints on the choice of our parameters when finding the values of the θ_j 's, otherwise following

method (1) above. Method (2) should yield an excellent representation for the bulk $\epsilon(\omega)$, as the constraints introduced would make the determination of the fit values of the θ_j quite precise. However, because there is a strong interplay between the values of the discontinuities in $W_{cv}(E)$ and the values of the θ_j 's, it is not clear that our model is capable of yielding accurate values for the θ_j 's in the absence of any knowledge about the discontinuities in $W_{cv}(E)$. Thus, the accuracy of method (1) is questionable.

Finally, we ask how our model could be improved even further. In principle, it could be improved by introducing θ_j as a free parameter at each critical point. However, in practice, that improves the quality of our fits only marginally, and at the cost of introducing seven new parameters with respect to which $\epsilon(\omega)$ varies nonlinearly, which makes the convergence of our fits much slower and less stable. Therefore, we have elected not to include the θ_j as parameters in our model. Second, in principle it could be improved by including the effect of indirect transitions. However, above the direct band gap indirect transitions contribute only negligibly to $\epsilon(\omega)$. Therefore, the introduction of additional parameters to allow the inclusion of indirect transitions is justified only in fitting $k(\omega_j)$ or $L_2(\omega_j)$ below the direct band gap of an indirect-band-gap semiconductor, and not at all for a direct-band-gap semiconductor. However, our model could easily be improved through the addition of an excitonic contribution to $\epsilon(\omega)$ near $E_0(\Gamma)$. As is apparent from $L_2(\omega_j)$ in Fig. 15, that would significantly improve our fit near $E_0(\Gamma)$, although it would have no significant effect at other energies. That could be done easily using formulas from the literature.²¹ Finally, it may be possible to improve our treatment of non-Lorentzian line shapes.

In conclusion, we have proposed a model for the optical dielectric function $\epsilon(\omega)$, have shown in detail how to use it in the fitting of spectral data, and have shown that it is superior to previous models in several ways and inferior in none. Further, we have shown how the use of this model to fit SE data reveals a fundamental flaw in the usual interpretation of that data. Finally, we have considered the possible ways in which our model could be improved.

ACKNOWLEDGMENTS

The authors are deeply grateful to the Defense Advanced Research Projects Agency for its support under ONR Contract No. N00014-86-K-0070.

APPENDIX: INTEGRATION OF EQUATIONS (2.9) AND (2.10)

Our new model requires the integration of Eqs. (2.9) and (2.10) in closed form. Inspection of the functional forms of the joint density of states given in Eqs. (4.1)–(4.3) reveals that there are only four types of integrals to be performed, with the integrals bounded above and below by critical points. For convenience we designate the subscripts on the energies and line widths by i (initial) at each lower CP and by f (final) at each upper CP. Then

the basic four types of integrals to be performed are

$$H_n(\omega) = \int_{E_i}^{E_f} dE \frac{\sqrt{E - E_i}}{E^{2-n}} \mathcal{L}(\omega, E), \quad (\text{A1})$$

$$F_n(\omega) = \int_{E_i}^{E_f} dE \frac{\sqrt{E - E_i} \sqrt{E_f - E}}{E^{2-n}} \mathcal{L}(\omega, E), \quad (\text{A2})$$

$$G_n(\omega) = \int_{E_i}^{E_f} dE \frac{1}{E^{2-n}} \mathcal{L}(\omega, E), \quad (\text{A3})$$

and

$$K_n(\omega) = \int_{E_i}^{E_f} dE \frac{\sqrt{E_f - E}}{E^{2-n}} \mathcal{L}(\omega, E), \quad (\text{A4})$$

where

$$\mathcal{L}(\omega, E) = \Phi(\hbar\omega - E) - \Phi(\hbar\omega + E),$$

n is a small non-negative integer and $\Phi(\hbar\omega \pm E)$ is a broadening function. Its imaginary part becomes equal to $-\pi\delta(\hbar\omega \pm E)$ in the absence of line broadening. For the case of Lorentzian broadening $\Phi(\hbar\omega \pm E)$ assumes the form

$$\Phi_L(\hbar\omega \pm E) \equiv \frac{1}{\hbar\omega \pm E + i\Gamma(E)}. \quad (\text{A5})$$

For the case of Gaussian broadening it assumes the form

$$\Phi_G(\hbar\omega \pm E) \equiv -i \int_0^\infty ds \exp\{i[\hbar\omega \pm E + i\sigma^2(E)s]s\}. \quad (\text{A6})$$

We consider first the case of Lorentzian broadening, for which Eqs. (A1)–(A4) can be integrated analytically in closed form if we choose the simple form

$$\Gamma(E) = \gamma E + \beta \quad (\text{A7})$$

for $\Gamma(E)$, with

$$\gamma = \frac{\Gamma_f - \Gamma_i}{E_f - E_i}, \quad \beta = \frac{E_f \Gamma_i - E_i \Gamma_f}{E_f - E_i}. \quad (\text{A8})$$

Although not accurate far from the critical points, this simple form suffices because the integrals to be performed are relatively insensitive to the values of $\Gamma(E)$ far from the critical points. The substitution of Eq. (A7) into Eqs. (A5) and (A6) enables one to perform the integrals in Eqs. (A1)–(A4) analytically for $n \leq 5$. Here we show the results only up to $n = 3$, because we find those results sufficient for accurately modeling the dielectric function of semiconductors. Our results are written in such a way that the computer generated numerical values of all multiple-valued functions will be the correct values. Thus, for any complex energy $z = |z|e^{i\phi}$, $z^{0.5} = |z|^{0.5}e^{i\phi/2}$ and $\ln z = \ln|z| + i\phi$, with ϕ chosen to lie in the range from $-\pi$ to π .

For brevity we define the following quantities:

$$\Delta_j = \hbar\omega - E_j + i\Gamma_j, \quad j = i \text{ or } f, \quad (\text{A9})$$

$$\Omega_j = \hbar\omega + E_j + i\Gamma_j, \quad j = i \text{ or } f, \quad (\text{A10})$$

$$\Upsilon_1 \equiv 2 \ln[\sqrt{\Delta_i} + \sqrt{E_f - E_i - i(\Gamma_f - \Gamma_i)}] - \ln \Delta_f, \quad (A11)$$

$$\Upsilon_3 \equiv 2 \ln[\sqrt{\Delta_f} + i\sqrt{E_f - E_i - i(\Gamma_f - \Gamma_i)}] - \ln \Delta_i, \quad (A13)$$

$$\Upsilon_2 \equiv 2 \ln[\sqrt{\Omega_i} + i\sqrt{E_f - E_i + i(\Gamma_f - \Gamma_i)}] - \ln \Omega_f, \quad (A12)$$

$$\Upsilon_4 \equiv 2 \ln[\sqrt{\Omega_f} - \sqrt{E_f - E_i + i(\Gamma_f - \Gamma_i)}] - \ln \Omega_i, \quad (A14)$$

Then, the desired integrals are given by the following equations:

$$H_0(\omega) = \frac{\sqrt{1-i\gamma}\sqrt{\Delta_i} \Upsilon_1 - i\sqrt{1+i\gamma}\sqrt{\Omega_i} \Upsilon_2 - 4\sqrt{E_i} \tan^{-1} \sqrt{\frac{E_f}{E_i} - 1}}{(\hbar\omega + i\beta)^2}, \quad (A15)$$

$$H_1(\omega) = \frac{1}{\hbar\omega + i\beta} \left(\frac{\sqrt{\Delta_i}}{\sqrt{1-i\gamma}} \Upsilon_1 + i \frac{\sqrt{\Omega_i}}{\sqrt{1+i\gamma}} \Upsilon_2 \right), \quad (A16)$$

$$H_2(\omega) = \frac{\sqrt{\Delta_i}}{(1-i\gamma)\sqrt{1-i\gamma}} \Upsilon_1 - i \frac{\sqrt{\Omega_i}}{(1+i\gamma)\sqrt{1+i\gamma}} \Upsilon_2 - 4 \frac{\sqrt{E_f - E_i}}{1 + \gamma^2}, \quad (A17)$$

$$F_0(\omega) = \frac{\pi}{(\hbar\omega + i\beta)^2} \left[0.5 \left(\sqrt{\frac{E_i}{E_f}} \Delta_f + \sqrt{\frac{E_f}{E_i}} \Delta_i \right) - \sqrt{\Delta_i} \sqrt{\Delta_f} \right] - \frac{\pi}{(\hbar\omega + i\beta)^2} \left[0.5 \left(\sqrt{\frac{E_i}{E_f}} \Omega_f + \sqrt{\frac{E_f}{E_i}} \Omega_i \right) - \sqrt{\Omega_i} \sqrt{\Omega_f} \right], \quad (A18)$$

$$F_1(\omega) = \pi \frac{0.5(\Delta_i + \Delta_f) - \sqrt{\Delta_i} \sqrt{\Delta_f}}{(\hbar\omega + i\beta)(1-i\gamma)} + \pi \frac{0.5(\Omega_i + \Omega_f) - \sqrt{\Omega_i} \sqrt{\Omega_f}}{(\hbar\omega + i\beta)(1+i\gamma)}, \quad (A19)$$

$$F_2(\omega) = \pi \frac{0.5(\Delta_i + \Delta_f) - \sqrt{\Delta_i} \sqrt{\Delta_f}}{(1-i\gamma)^2} + \pi \frac{0.5(\Omega_i + \Omega_f) - \sqrt{\Omega_i} \sqrt{\Omega_f}}{(1+i\gamma)^2}, \quad (A20)$$

$$G_0(\omega) = \frac{(1-i\gamma)(\ln \Delta_i - \ln \Delta_f)}{(\hbar\omega + i\beta)^2} + \frac{(1+i\gamma)(\ln \Omega_i - \ln \Omega_f)}{(\hbar\omega + i\beta)^2} + \frac{2(\ln E_f - \ln E_i)}{(\hbar\omega + i\beta)^2}, \quad (A21)$$

$$G_1(\omega) = \frac{\ln \Delta_i - \ln \Delta_f}{\hbar\omega + i\beta} - \frac{\ln \Omega_i - \ln \Omega_f}{\hbar\omega + i\beta}, \quad (A22)$$

$$G_2(\omega) = \frac{\ln \Delta_i - \ln \Delta_f}{1-i\gamma} + \frac{\ln \Omega_i - \ln \Omega_f}{1+i\gamma}, \quad (A23)$$

$$K_0(\omega) = \frac{i\sqrt{1-i\gamma}\sqrt{\Delta_f} \Upsilon_3 + \sqrt{1+i\gamma}\sqrt{\Omega_f} \Upsilon_4 - 4\sqrt{E_f} \ln \left(\sqrt{\frac{E_f}{E_i}} - \sqrt{\frac{E_f}{E_i} - 1} \right)}{(\hbar\omega + i\beta)^2}, \quad (A24)$$

$$K_1(\omega) = \frac{1}{\hbar\omega + i\beta} \left(i \frac{\sqrt{\Delta_f}}{\sqrt{1-i\gamma}} \Upsilon_3 - \frac{\sqrt{\Omega_f}}{\sqrt{1+i\gamma}} \Upsilon_4 \right) \quad (A25)$$

and

$$K_2(\omega) = i \frac{\sqrt{\Delta_f}}{(1-i\gamma)\sqrt{1-i\gamma}} \Upsilon_3 + \frac{\sqrt{\Omega_f}}{(1+i\gamma)\sqrt{1+i\gamma}} \Upsilon_4 + 4 \frac{\sqrt{E_f - E_i}}{1 + \gamma^2}. \quad (A26)$$

For the case of Gaussian broadening, Eqs. (A1)–(A4) cannot be integrated analytically in closed form. To leave these equations to be integrated numerically in the fitting

of experimental data would increase the computer time by approximately 2 orders of magnitude. Therefore, we must find an analytic approximation to the integrals in those equations. From the form of

$$\text{Im}\{\Phi_G(\hbar\omega \pm E)\} = -\frac{\sqrt{\pi}}{2\sigma(E)} \exp\left(\frac{-(\hbar\omega \pm E)^2}{4\sigma^2(E)}\right), \quad (A27)$$

as compared to

$$\text{Im}\{\Phi_L(\hbar\omega \pm E)\} = -\frac{\Gamma(E)}{(\hbar\omega - E)^2 + \Gamma^2(E)}, \quad (\text{A28})$$

it is clear that one can approximately mimic the exact integrals given by Eqs. (A1)–(A4) in the Gaussian case by replacing Γ_j for $j = i$ or $j = f$ in Eqs. (A8)–(A14) by

$$D_j \equiv \Gamma_j \exp\left[-\alpha_j \left(\frac{\hbar\omega - E_j}{\Gamma_j}\right)^2\right]. \quad (\text{A29})$$

We have found that this approximation is satisfactory for the proper choice of the parameter α_j .

*Present address: Naval Research Laboratory, Washington, D.C. 20375-5000.

¹It should be pointed out that $\epsilon(\omega)$ is modified by the band bending present in the depletion region adjacent to any semiconductor surface or interface, becoming position dependent in the depletion region (Refs. 2–6), in which the band bending destroys translational invariance in one direction. That modification is significant for heavily doped semiconductors for energies $\hbar\omega$ near a critical-point energy. Also, $\epsilon(\omega)$ is modified by surface termination effects near the surface of a semiconductor (Refs. 3, 4, and 6). A knowledge of these effects, which are not considered in this paper, as well as a knowledge of the bulk $\epsilon(\omega)$, is important in the design of wave-guiding semiconductor devices.

²D. E. Aspnes and A. Frova, *Solid State Commun.* **7**, 155 (1969).

³R. Del Sole, *Solid State Commun.* **19**, 207 (1976).

⁴R. Del Sole and D. E. Aspnes, *Phys. Rev. B* **17**, 3310 (1977).

⁵R. A. Batchelor, A. C. Brown, and A. Hamnett, *Phys. Rev. B* **41**, 1401 (1990).

⁶D. Yang, J. W. Garland, P. M. Raccah, C. Coluzza, P. Frankl, M. Capizzi, F. Chambers, and G. Devane, *Appl. Phys. Lett.* **57**, 2829 (1990); *Physica B* **170**, 557 (1991).

⁷J. W. Garland, C. Kim, H. Abad, and P. M. Raccah, *Phys. Rev. B* **41**, 7602 (1990).

⁸P. M. Raccah, J. W. Garland, D. Yang, H. Abad, R. L. Strong, and M. C. McNeill, *J. Vac. Sci. Technol. A* **7**, 509 (1989).

⁹P. M. Raccah, J. W. Garland, Z. Zhang, U. Lee, D. Z. Xue, L. L. Abels, S. Ugur, and W. Wilinsky, *Phys. Rev. Lett.* **53**, 1958 (1984).

¹⁰P. M. Raccah, U. Lee, S. Ugur, D. Z. Xue, L. L. Abels, and J. W. Garland, *J. Vac. Sci. Technol. A* **3**, 138 (1985).

¹¹J. W. Garland and P. M. Raccah, *Proc. SPIE* **659**, 32 (1986).

¹²P. M. Amirtharaj, J. H. Dinan, J. J. Kennedy, P. R. Boyd, and O. J. Glembocki, *J. Vac. Sci. Technol. A* **4**, 2028 (1986).

¹³W. Franz, *Z. Naturforsch.* **13a**, 484 (1958).

¹⁴L. V. Keldysh, *Zh. Eksp. Teor. Fiz.* **34**, 1138 (1958) [*Sov. Phys. JETP* **74**, 788 (1958)].

¹⁵D. E. Aspnes and J. E. Rowe, *Phys. Rev. B* **5**, 4022 (1972).

¹⁶F. Wooten, *Optical Properties of Solids* (Academic, New York, 1972).

¹⁷D. E. Aspnes, in *Optical Properties of Solids*, edited by B. O. Seraphin (North-Holland, Amsterdam, 1976), pp. 821–826.

¹⁸M. Erman, J. B. Theeten, N. Vodjdani, and Y. Demay, *J. Vac. Sci. Technol. B* **1**, 328 (1983).

¹⁹F. L. Terry, Jr., *J. Appl. Phys.* **70**, 409 (1991).

²⁰S. Adachi, *Phys. Rev. B* **35**, 7454 (1987).

²¹S. Adachi, *Phys. Rev. B* **41**, 1003 (1990).

²²D. W. Jenkins, *J. Appl. Phys.* **68**, 1848 (1990).

²³J. W. Garland, H. Abad, M. Viccaro, and P. M. Raccah, *Appl. Phys. Lett.* **52**, 1176 (1988).

²⁴The smallness of $\epsilon_2(\omega)$ below E_0 , coupled with the rather large line widths observed at E_1 and higher-energy critical points, provides perhaps the clearest proof of the incorrect-

ness of the Lorentzian approximation.

²⁵D. Brust, *Phys. Rev.* **134**, A1337 (1964).

²⁶D. E. Aspnes, P. Handler, and D. F. Blosssey, *Phys. Rev.* **166**, 921 (1968).

²⁷F. Bassani and G. Pastori Parravicini, *Electronic States and Optical Transitions in Solids* (Pergamon, New York, 1975), Chap. 5.

²⁸J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **28**, 8 (1954).

²⁹P. Nozières and D. Pines, *Phys. Rev.* **109**, 741 (1958); **109**, 762 (1958); **109**, 1062 (1958); *Nuovo Cimento* **9**, 470 (1958); *Phys. Rev.* **111**, 442 (1958).

³⁰H. Ehrenreich and M. H. Cohen, *Phys. Rev.* **115**, 286 (1959).

³¹C. C. Kim, Ph.D. thesis, University of Illinois at Chicago, 1991.

³²S. Y. Toyozawa, *Prog. Theor. Phys.* **20**, 53 (1958).

³³S. F. Edwards and Y. B. Gulyaev, *Proc. Phys. Soc. London* **83**, 496 (1964).

³⁴T. Lukes and K. T. S. Somaratna, *J. Phys. C* **3**, 2044 (1970).

³⁵P. Lautenschlager, M. Garriga, L. Vina, and M. Cardona, *Phys. Rev. B* **36**, 4821 (1987).

³⁶D. E. Aspnes, *Surf. Sci.* **135**, 284 (1983).

³⁷M. Cardona, *Modulation Spectroscopy*, Suppl. 11 of *Solid State Physics* (Academic, New York, 1969).

³⁸D. E. Aspnes, in *Handbook on Semiconductors*, edited by M. Balkanski (North-Holland, Amsterdam, 1980), Vol. 2, p. 109.

³⁹S. M. Kelso, D. E. Aspnes, M. A. Pollack, and R. E. Nahory, *Phys. Rev. B* **26**, 6669 (1982).

⁴⁰E. D. Palik, in *Handbook of Optical Constants of Solids*, edited by E. D. Palik (Academic, New York, 1985), pp. 429–432.

⁴¹J. B. Theeten, D. E. Aspnes, and R. P. H. Chang, *J. Appl. Phys.* **49**, 6097 (1978); J. B. Theeten, D. E. Aspnes, and R. P. H. Chang (private communication).

⁴²D. E. Aspnes and A. A. Studna, *Phys. Rev. B* **27**, 985 (1983).

⁴³H. C. Casey, D. D. Sell, and K. W. Wecht, *J. Appl. Phys.* **46**, 250 (1975).

⁴⁴A. N. Pikhtin and A. D. Yas'kov, *Fiz. Tekh. Poluprovodn.* **12**, 1047 (1978) [*Sov. Phys. Semicond.* **12**, 622 (1978)].

⁴⁵D. T. F. Marple, *J. Appl. Phys.* **35**, 1241 (1964).

⁴⁶J. R. Chelikowski and M. L. Cohen, *Phys. Rev. B* **14**, 556 (1976).

⁴⁷D. E. Aspnes and A. A. Studna, *Phys. Rev. B* **7**, 4605 (1973).

⁴⁸S. Logothetidis, M. Alouani, M. Garriga, and M. Cardona, *Phys. Rev. B* **41**, 2959 (1990).

⁴⁹The ratio of the value of $W_{cv}(E)$ just above a discontinuity to its value just below the discontinuity is approximately equal to the corresponding ratio for $J_{cv}(E)$ because the matrix elements $P_{cv}(\mathbf{k})$ are slowly varying functions of \mathbf{k} for a single pair of bands.

⁵⁰J. Zhang, C. C. Kim, D. Yang, J. W. Garland and P. M. Raccah (unpublished).

⁵¹R. Del Sole, *J. Phys. C* **8**, 2971 (1975).