Resonant Nonlinear Optical Susceptibility: Electroreflectance in the Low-Field Limit

D. E. Aspnes and J. E. Rowe
Bell Telephone Laboratories, Murray Hill, New Jersey 07974
(Received 5 October 1971)

A theory of the electric field effect on the dielectric function of solids is developed by means of a perturbation treatment similar to that used to describe nonlinear optical phenomena. The field-induced change in the dielectric function is given directly as a Brillouin-zone integral over a fourth-power resonant denominator. It is shown that electric field modulation results in spectra nearly proportional to the third derivative of the unperturbed dielectric function for common experimental conditions. The perturbation treatment is related to the standard (high-field) Franz-Keldysh theories. The differences between the two approaches is discussed in physical terms. Simplified expressions for parabolic critical points are developed and used to discuss experimental criteria for the validity of the perturbation theory. Application to band-structure analysis is discussed. In particular, critical-point parameters (energy, broadening, etc.) enter nearly independently, and the mathematical form of the fourth-rank tensorial line shape is simple enough to permit the calculation of electroreflectance spectra from existing band-structure calculations.

I. INTRODUCTION

A dc electric field can be considered as a coherent superposition of photons of zero frequency. From this viewpoint, it is expected that the classical Franz-Keldysh theory,[1-3] which describes exactly the effect of a uniform electric field on unscattered crystal electrons in the one-band approximation, should be related to perturbation theory and therefore, the electroreflectance (ER) effect should be related to nonlinear optics.[4-6] This hypothesis is confirmed experimentally: Quantitative ER spectra[7-11] in the limit of a small dc field $\mathcal{E}$ are known to scale quadratically in $\mathcal{E}$ and linearly in light intensity, providing direct evidence that ER in the low-field large-broadening limit involves two dc "photons" and a single ac photon, characteristic behavior of a third-order susceptibility. In addition, ER theory[12-13] and spectra[14] have been used to calculate nonlinear optical coefficients to reasonable accuracy. The sharpness of ER spectra, and their connection to critical points in the joint density of states, indicate that this is a resonant susceptibility analogous to that encountered in Raman scattering. It is the objective of this paper to develop the perturbation description of ER[15] and to establish theoretically the connection of the Franz-Keldysh theory to that of nonlinear optics.

The outline of the paper is as follows. In Sec. II, we discuss the perturbation theoretic treatment of the electric field, using a standard time-dependent interaction formalism where the field is turned on adiabatically and the wave function evolves with time. We relate the resulting expression to the convolution-integral (strong-field) formalism in Sec. III by performing an asymptotic expansion of the kernel function which is valid in the limit of large broadening. The range of validity, relation to the unperturbed dielectric function, and connection to experiment are discussed in Sec. IV. The physical meaning of and the difference between various approaches to the electric field problem is given in the summary in Sec. V.

II. PERTURBATION THEORY

In this section we derive by means of standard time-dependent perturbation theory the dominant resonant term in the third-order nonlinear optical susceptibility caused by the uniform electric field. We work in the one-electron approximation, where the Bloch functions $\varphi_{n\vec{k}}(\vec{r})$ describing electron states in the crystal satisfy the time-dependent Schrödinger equation

$$H_0 \varphi_{n\vec{k}} e^{-i\omega_{n\vec{k}} t} = i\hbar \frac{\partial}{\partial t} \varphi_{n\vec{k}} e^{-i\omega_{n\vec{k}} t} \quad , \tag{2.1}$$

where H_0 is the unperturbed Hamiltonian and $E_{n\vec{k}} = \hbar\omega_{n\vec{k}}$ is the energy of the state $\varphi_{n\vec{k}}$. The spatially uniform dc field can be represented by a time-dependent perturbation term

$$H_1 = +e\vec{\mathcal{E}}_1 \cdot \vec{x} e^{\eta t} \quad , \tag{2.2}$$

and the photon field by the perturbation terms

$$H_+ + H_- = +e\vec{\mathcal{E}}_2 \cdot \vec{x}(e^{-i\omega_+ t} + e^{i\omega_- t}) \quad , \tag{2.3}$$

where

$$E_\pm = \hbar\omega_\pm = \hbar\omega \pm i\Gamma \quad . \tag{2.4}$$

The amplitude vectors $\vec{\mathcal{E}}_1$ and $\vec{\mathcal{E}}_2$ are real, and the turn-on parameters $\eta > 0$ and $\Gamma > 0$ enable the evolution of the one-electron wave functions to be calculated starting with the unperturbed crystal at $t \to -\infty$. We work in the scalar-potential gauge,

which is the standard representation of the dc term in the high-field Franz-Keldysh theory.[1-3] It is equally possible (but algebraically much more difficult) to work in the vector-potential gauge. The dipole approximation is used throughout.[16]

The time-dependent wave function $\psi_{n\vec{k}}(\vec{r}, t)$, describing the evolution of the one-electron state $\varphi_{n\vec{k}}(\vec{r})e^{-i\omega_{n\vec{k}}t}$, is defined by

$$(H_0+H_1+H_+ +H_-)\psi_{n\vec{k}}(\vec{r}, t)$$
$$= (H_0+e\vec{\mathcal{E}}_1\cdot\vec{x}e^{\eta t}+e\vec{\mathcal{E}}_2\cdot\vec{x}e^{-i\omega_+ t}+e\vec{\mathcal{E}}_2\cdot\vec{x}e^{i\omega_- t})\psi_{n\vec{k}}(\vec{r}, t)$$
$$= i\hbar\frac{\partial}{\partial t}\psi_{n\vec{k}}(\vec{r}, t) \ ,$$
$$\lim_{t\to-\infty}\psi_{n\vec{k}}(\vec{r}, t)=\varphi_{n\vec{k}}(\vec{r})e^{-i\omega_{n\vec{k}}t} \ . \qquad (2.5)$$

The wave function $\psi_{n\vec{k}}(\vec{r}, t)$ can be evaluated as a power series in the perturbation terms by standard techniques.[17] We find, to third order in the perturbations H_1, H_+, and H_-,

$$\psi_{n\vec{k}}(\vec{r}, t)=\varphi_{n\vec{k}}(\vec{r})e^{-i\omega_{n\vec{k}}t}+(G^1H_1+G^+H_++G^-H_-)\varphi_{n\vec{k}}(\vec{r})e^{-i\omega_{n\vec{k}}t}$$
$$+[G^{11}H_1G^1H_1+G^{1+}H_1G^+H_++(\text{seven remaining second-order terms})]\varphi_{n\vec{k}}(\vec{r})e^{-i\omega_{n\vec{k}}t}$$
$$+[G^{111}H_1G^{11}H_1G^1H_1+G^{11+}H_1G^{1+}H_1G^+H_+$$
$$+(25\text{ remaining third-order terms})]$$
$$\times\varphi_{n\vec{k}}(\vec{r})e^{-i\omega_{n\vec{k}}t} \ . \qquad (2.6)$$

For brevity, only two each of the nine second-order and 27 third-order terms are written; the complete sets are obtained by taking all possible ordered pairs and triples, respectively, of the symbol set 1, +, −, and forming products of the corresponding perturbations and Green's functions $G^{ij\cdots}$ as indicated by the terms shown. Here,

$$G^{ij\cdots}=\frac{1}{E_{n\vec{k}}+S_i+S_j+\cdots-H_0} \ , \qquad (2.7a)$$

where the $i, j, \ldots$ represent a particular combination of symbols (1, +, −), and if

$$i=\begin{Bmatrix}1\\+\\-\end{Bmatrix} \ , \quad \text{then} \quad S_i=\begin{Bmatrix}i\hbar\eta\\E_+\\-E_-\end{Bmatrix} \ , \quad \text{etc.} \qquad (2.7b)$$

The dielectric tensor ϵ_{ij} of the crystal can be calculated from the polarization $\vec{P}=\{P_i\}$ according to

$$D_i=\epsilon_{ij}\mathcal{E}_j=\delta_{ij}\mathcal{E}_j+4\pi P_i \ , \qquad (2.8)$$

where

$$\vec{P}=-\frac{e}{V}\sum_{v\vec{k}}\langle\psi_{v\vec{k}}(t)|\vec{x}|\psi_{v\vec{k}}(t)\rangle \qquad (2.9)$$

and $\psi_{v\vec{k}}(\vec{r}, t)$ is given by Eqs. (2.7). The subscript v for the band index denotes a filled valence band.

The exact evaluation of Eq. (2.9) using Eq. (2.6) is a formidable task, since it contains six first-order terms, 27 second-order terms, and 108 third-order terms, and contains the complete microscopic description of the entire range of nonlinear optical effects, such as harmonic generation, optical rectification, etc., through third order.[5] But the terms describing electroreflectance can be readily obtained by making use of the following observations:

(i) The induced polarization must have the time dependence of the photon field, since the frequency of the reflected beam is the same as that of the incident beam. We choose the time dependence to be that of $H_+\sim e^{-i\omega_+ t}$, in accordance with the usual conventions. This reduces the number of first-, second-, and third-order terms to be evaluated to 2, 6, and 24, respectively.

(ii) Of the remaining terms with the proper frequency dependence, there exists only one term in each order where each resonant factor in the denominator comes into resonance at the same frequency (barring "accidental" resonances due to specific interband energy separations). This dominant term for the nth order contains one photon interaction H_+ and $(n-1)$ dc field interactions H_1. Through third order, the contributions to $\Delta\vec{P}$, the field-induced correction to the polarization, will have the explicit form

$$\Delta\vec{P}=-\frac{e}{V}\sum_v\langle v\vec{k}|\vec{x}G^{1+}H_1G^+H_+|v\vec{k}\rangle$$
$$-\frac{e}{V}\sum_{v\vec{k}}\langle v\vec{k}|\vec{x}G^{11+}H_1G^{1+}H_1G^+H_+|v\vec{k}\rangle \ . \qquad (2.10)$$

All other terms within a given order will have at least one nonresonant denominator replacing a resonant denominator, and so will be at least of order $(\Gamma/E_\pm)$ smaller in magnitude. We note, however, that this assumption is valid only for $\hbar\omega$ near or above the fundamental absorption threshold, so that allowed valence and conduction states separated by the photon energy can always be found.

We now evaluate Eq. (3.10) through third order by inserting complete sets of wave functions $|n'\vec{k}'\rangle$ between operators, leading to terms of form

$$\langle n'\vec{k}'|G^{ij\cdots}|n\vec{k}\rangle=\delta_{n'n}\delta_{\vec{k}'\vec{k}}(E_{n\vec{k}}-E_{n'\vec{k}'}+S_i+S_j+\cdots)^{-1} , \qquad (2.11a)$$

$$\langle n'\vec{k}'|\vec{x}|n\vec{k}\rangle=\begin{cases}\dfrac{-i\hbar\vec{P}_{n'n}(\vec{k})}{mE_{n'n}(\vec{k})}\delta_{\vec{k}'\vec{k}} & \text{if } n'\neq n \ , \qquad (2.11b)\\ +i\delta_{\vec{k}'\vec{k}}\nabla_{\vec{k}} & \text{if } n'=n \ , \qquad (2.11c)\end{cases}$$

where $\vec{P}_{n'n}(\vec{k})=-i\hbar\langle n\vec{k}|\nabla_{\vec{x}}|n\vec{k}\rangle$ is the momentum matrix element. Clearly, all intermediate sums over $\vec{k}$ vanish because each operator is diagonal

in $\vec{k}$. Second, in order for each denominator to resonate at the same frequency, all intermediate states must in addition correspond to the same band. Therefore, the only interband terms involve the polarization operator $\vec{x}$ and the photon interaction H_+, and the dc interaction H_1 appears only in intraband matrix elements. The approximation of neglecting interband matrix elements of $\vec{x}$, used throughout in the Franz-Keldysh theory, is seen to arise naturally out of the procedure for isolating the dominant term in the perturbation approximation.

With these simplifications and through the use of Eqs. (2.11), Eq. (2.10) becomes approximately

$$\Delta\vec{P} = \frac{-ie^3\hbar^2}{m^2} e^{\eta t} \sum_{v,c,\vec{k}} \frac{\vec{P}_{vc}}{E_{cv}} \frac{1}{E_+ + i\hbar\eta - E_{cv}(\vec{k})} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left(\frac{1}{E_+ - E_{cv}(\vec{k})} \frac{\vec{\mathcal{E}}_2 \cdot \vec{P}_{cv}}{E_{cv}} \right)$$
$$+ \frac{e^4\hbar^2}{m^2} e^{2\eta t} \sum_{v,c,\vec{k}} \frac{\vec{P}_{vc}}{E_{cv}} \frac{1}{E_+ + i2\hbar\eta - E_{cv}(\vec{k})} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left[\frac{1}{E_+ + i\hbar\eta - E_{cv}(\vec{k})} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left(\frac{1}{E_+ - E_{cv}(\vec{k})} \frac{\vec{\mathcal{E}}_2 \cdot \vec{P}_{cv}}{E_{cv}} \right) \right] , \tag{2.12}$$

where $\vec{\mathcal{E}} = \vec{\mathcal{E}}_1$ represents the dc field. Since the photon turn-on parameter Γ appears in each denominator, Eq. (2.12) is well behaved in the limit $\eta \to 0$. This limit simply represents the physical fact that the dc field has been turned on long before the arrival of the photon. Accordingly, we set $\eta = 0$ and identify Γ as the phenomenological broadening parameter, representing excited-state lifetimes in the usual manner.[18] With this interpretation, Γ is properly a function of $\vec{k}$, but varies sufficiently slowly in the vicinity of a resonance so that it may be approximated by a constant. If we define

$$\frac{1}{W_{cv}} = \frac{1}{E_+ - E_{cv}(\vec{k})} \tag{2.13}$$

and retain only the component of polarization $\vec{P}$ in the direction of $\hat{\epsilon} = \vec{\mathcal{E}}_2/|\vec{\mathcal{E}}_2|$, the unit polarization vector of the photon, then by Eqs. (2.8), (2.12), and (2.13), per band pair (c, v):

$$\Delta\epsilon_{cv}(\omega, \Gamma, \vec{\mathcal{E}}) \simeq -\frac{4\pi e^3\hbar^2}{m^2V} \sum_{\vec{k}} \frac{(\hat{\epsilon} \cdot \vec{P}^*_{cv})}{E_{cv}} \frac{1}{W} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left(\frac{1}{W} \frac{(\hat{\epsilon} \cdot \vec{P}_{cv})}{E_{cv}} \right)$$
$$+ \frac{4\pi e^4\hbar^2}{m^2V} \sum_{\vec{k}} \frac{(\hat{\epsilon} \cdot \vec{P}^*_{cv})}{E_{cv}} \frac{1}{W} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left[\frac{1}{W} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left(\frac{1}{W} \frac{(\hat{\epsilon} \cdot \vec{P}_{cv})}{E_{cv}} \right) \right] . \tag{2.14}$$

Equation (2.14) gives the dominant second- and third-order field-induced corrections to the dielectric function, varying linearly and quadratically, respectively, with the dc field $\vec{\mathcal{E}}$. If the crystal has inversion symmetry ($\vec{P}^*_{cv} = \vec{P}_{cv}$) or if $\vec{P}_{cv}$ is independent of $\vec{k}$, then the second-order term can be written in the form

$$\Delta\epsilon \sim \sum_{\vec{k}} F(\vec{k}) \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} F(\vec{k}) = \tfrac{1}{2} \sum_{\vec{k}} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} [F(\vec{k})^2] , \tag{2.15}$$

which vanishes in summation by the symmetry of the Brillouin zone. Therefore, the existence of a correction term linear in $\vec{\mathcal{E}}$ depends *entirely* on the $\vec{k}$ dependence of the momentum matrix element, which is typically small enough to be neglected even in the absence of inversion symmetry.[19] The same condition and assumptions appear in the derivation of the Franz-Keldysh theory,[1-3] so the second-order correction term will not be discussed further.

The third-order correction term can be simplified by noting that $1/W_{cv}(\vec{k})$ depends much more strongly on $\vec{k}$ than do either $\vec{P}_{cv}(\vec{k})$ or $E_{cv}(\vec{k})$, so that the gradient operators in Eq. (2.14) need only act on the factors $1/W$. Also, since the energy range of the resonant parts of $\Delta\epsilon$ is of order Γ, the nonresonant denominators E_{cv} can be replaced by $\hbar\omega$, introducing errors no larger than $\sim \Gamma/\hbar\omega$. We find

$$\Delta\epsilon_{cv}(\omega, \Gamma, \vec{\mathcal{E}}) \cong \frac{4\pi e^4 |\hat{\epsilon} \cdot \vec{P}_{cv}|^2}{m^2\omega^2V} \times \sum_{\vec{k}} \frac{1}{W_{cv}} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \left(\frac{1}{W_{cv}} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} \frac{1}{W_{cv}} \right) . \tag{2.16}$$

Using the identities

$$\vec{\mathcal{E}} \cdot \nabla_{\vec{k}} (1/W_{cv}) = (1/W^2_{cv})(\vec{\mathcal{E}} \cdot \nabla_{\vec{k}} E_{cv}) , \tag{2.17a}$$

$$0 = \sum_{\vec{k}} \vec{\mathcal{E}} \cdot \nabla_{\vec{k}} [(1/W^4_{cv})(\vec{\mathcal{E}} \cdot \nabla_{\vec{k}} E_{cv})]$$
$$= -4 \sum_{\vec{k}} (1/W^5_{cv})(\vec{\mathcal{E}} \cdot \nabla_{\vec{k}} E_{cv})^2 + \sum_{\vec{k}} (1/W^4_{cv})(\vec{\mathcal{E}} \cdot \nabla_{\vec{k}})^2 E_{cv} , \tag{2.17b}$$

we conclude

$$\Delta\epsilon(\omega, \Gamma, \vec{\mathcal{E}}) = \frac{\pi e^4 |\hat{\epsilon} \cdot P_{cv}|^2}{m^2\omega^2V} \sum_{\vec{k}} \frac{1}{W^4_{cv}} (\vec{\mathcal{E}} \cdot \nabla_{\vec{k}})^2 E_{cv}(\vec{k})$$
$$= \frac{2Q}{\pi\omega^2} \int d^3k \frac{(\hbar\Omega)^3}{[E_+ - E_{cv}(\vec{k})]^4} , \tag{2.18a}$$

where

$$Q = \frac{4\pi^2 e^2 |\hat{\epsilon}\cdot\vec{P}_{cv}|^2}{m^2}\left(\frac{2}{(2\pi)^3}\right), \qquad (2.18b)$$

$$(\hbar\Omega)^3 = \tfrac{1}{8}e^2(\vec{\mathcal{E}}\cdot\nabla_{\vec{k}})^2 E_{cv} = e^2\mathcal{E}^2\hbar^2/8\mu_{\parallel} \quad . \qquad (2.18c)$$

Equations (2.18) represent the main result of this paper, derived by means of perturbation theory. For reference, the linear dielectric constant is

$$\epsilon_{cv}(\omega,\Gamma) = \frac{-Q}{\pi\omega^2}\int d^3k\,\frac{1}{E_+ - E_{cv}(\vec{k})} \quad . \qquad (2.19)$$

If $\hbar\Omega$ is nearly $\vec{k}$ independent in a given energy range, we can write

$$\Delta\epsilon_{cv}(E,\Gamma,\vec{\mathcal{E}}) \equiv \frac{1}{3E^2}\left(\hbar\Omega\frac{\partial}{\partial E}\right)^3 E^2\epsilon_{cv}(E,\Gamma); \qquad (2.20)$$

thus the third-order correction is closely related to the third derivative of the unperturbed dielectric function.[20] Of particular importance is the fact that Eqs. (2.18) are obtained independently of details of the interband energy, and may be used either with calculated energy-band structures or simplified models with equal validity. As with any third-order susceptibility, the line shape can also be represented as a fourth-rank tensor, here taking the form $\chi_{iijj}(\omega, 0, 0)$, where the tensor indices i arise from the momentum matrix element in the prefactor Q, and the indices j represent the interband reduced mass in the characteristic energy factor $(\hbar\Omega)^3$.

III. RELATION TO CONVOLUTION FORMALISM

In this section we derive Eq. (2.18a) as a weak-field large-broadening limit of the convolution-integral formalism,[3,21] the strong-field limit of the Franz-Keldysh effect. In the presence of lifetime broadening, represented by the same phenomenological broadening parameter Γ used in Sec. II, the convolution integral giving the linear dielectric function of a crystal in the presence of a dc field $\vec{\mathcal{E}}$ in the one-electron approximation can be written as the Fourier transform of a time-dependent polarization current[22]:

$$\epsilon_{cv}(E,\Gamma,\vec{\mathcal{E}}) = 1 + \frac{i4\pi e^2}{m^2\omega^2\hbar}\int_{B.Z.} d^3k \int_0^\infty dt\,\{\hat{\epsilon}\cdot\vec{P}_{vc}[\vec{k} - \tfrac{1}{2}(t/\hbar)e\vec{\mathcal{E}}]\}$$
$$\times\{\hat{\epsilon}\cdot\vec{P}_{cv}[\vec{k} + \tfrac{1}{2}(t/\hbar)e\vec{\mathcal{E}}]\}e^{iE_+ t/\hbar}e^{-i\int_{-t/2}^{t/2}(dt'/\hbar)E_{cv}(\vec{k} - e\vec{\mathcal{E}}t'/\hbar)} \quad . \qquad (3.1)$$

The unit polarization vector $\hat{\epsilon}$ represents the polarization of the incident light and the wave vector integral is over the Brillouin zone (B. Z.). In standard developments, $\vec{P}_{cv}$ is taken as $\vec{k}$ independent (thereby eliminating the second-order correction term linear in field discussed in Sec. II), and the exponent is expanded in the weak-field approximation

$$\int_{-t/2}^{t/2}(dt'/\hbar)E_{cv}(\vec{k} - eEt'/\hbar) \cong (t/\hbar)E_{cv} + \tfrac{1}{3}t^3\Omega^3 \quad , \qquad (3.2)$$

where $\hbar\Omega$ is defined in Eq. (2.18c). With these approximations and the use of Eq. (2.18b), Eq. (3.1) becomes

$$\epsilon_{cv}(E,\Gamma,\vec{\mathcal{E}}) = \frac{iQ}{\pi\omega^2}\int_{B.Z.} d^3k \int_0^\infty \frac{dt}{\hbar}\exp\{-\tfrac{1}{3}it^3\Omega^3 + it[E_+ - E_{cv}(\vec{k})]/\hbar\} \quad , \qquad (3.3a)$$

$$\epsilon_{cv}(E,\Gamma,\vec{\mathcal{E}}) = \frac{Q}{\omega^2}\int_{B.Z.} d^3k\left[\frac{1}{\hbar\Omega}\,\mathrm{Gi}\left(\frac{E_{cv}(\vec{k}) - E_+}{\hbar\Omega}\right) + \frac{i}{|\hbar\Omega|}\,\mathrm{Ai}\left(\frac{E_{cv}(\vec{k}) - E_+}{\hbar\Omega}\right)\right] . \qquad (3.3b)$$

The unbroadened convolution integral[3] is recovered by setting $\Gamma = 0$ in Eq. (3.3b).

The perturbation result is obtained by taking Γ so large that the integrand in Eq. (3.3a) cuts off before the term $-it^3\Omega^3/3$ changes appreciably, a condition expressed by $|\hbar\Omega| \ll \Gamma$, i.e., the field is sufficiently weak so the magnitude of the characteristic electro-optic energy $|\hbar\Omega|$ is significantly less than the broadening energy Γ. Then, writing

$$e^{-it^3\Omega^3/3} = 1 - it^3\Omega^3/3 - \cdots \quad , \qquad (3.4)$$

we convert Eq. (3.3a) into an asymptotically convergent integral, the first two terms of which are, by explicit integration,

$$\epsilon_{cv}(\omega,\Gamma,\vec{\mathcal{E}}) \sim \frac{-Q}{\pi\omega^2}\int_{B.Z.} d^3k\,\frac{1}{E_+ - E_{cv}(\vec{k})} + \frac{2Q}{\pi\omega^2}\int_{B.Z.} d^3k\,\frac{(\hbar\Omega)^3}{[E_+ - E_{cv}(\vec{k})]^4} \quad , \qquad (3.5)$$

which are identical to the results obtained in Sec. II.

An alternative approach is to use explicitly the asymptotic expansions

$$\mathrm{Ai}(x) \sim 0, \qquad (3.6a)$$

$$\mathrm{Gi}(x) \sim \frac{1}{\pi x}\left(1 - \frac{2}{x^3} + \cdots\right) , \qquad (3.6b)$$

in Eq. (3.3b), which leads directly to Eq. (3.5). Since the function $\mathrm{Ai}(x)$ has an essential singularity at infinity (converges to zero more rapidly than an inverse power of x), it contributes nothing to the asymptotic expansion, which comes entirely from $\mathrm{Gi}(x)$. A similar situation occurs in the theory of Zener tunneling,[23] a direct electric field tunneling effect which also cannot be described in any finite order of perturbation theory. In regard to the Franz-Keldysh theory, this result demonstrates that the term $\mathrm{Ai}(x)$, describing the imaginary part of the field-dependent dielectric function in the absence of broadening, can be obtained only by summing the perturbation series exactly to infinite order, as is done in the derivation of Eq. (3.1).[3] Therefore, Eq. (3.5) *must* fail as $\Gamma \to 0$, since it is completely independent of the term $\mathrm{Ai}(x)$ in Eq. (3.1) in this limit. We discuss this more fully in Sec. IV.

By combining the results of Secs. I and II, we estimate the effect of the $\vec{k}$ dependence of the momentum matrix element in crystals lacking inversion symmetry, and show this is small. We expand the matrix element in a power series in t in Eq. (3.1). This expansion is inherently asymptotic and, in the absence of broadening, resulting integrals for parabolic-model band structures fail to converge (evaluation for exact band structures presumably remains finite but is too difficult to perform). Since multiplication of the integrand by t is operationally equivalent to $\partial/\partial E$, the $\vec{k}$-dependent term of the matrix element results in a W_{cv}^{-2} resonance compared to the W_{cv}^{-4} resonance of the ER term in Eq. (3.5). Therefore, matrix element effects are inherently of order Γ^2/E_g^2 smaller in magnitude. Even though it is always possible to find a range of field small enough so that the linear (matrix element) term dominates the quadratic (ER) term, it is unlikely that such linear effects related to the $\vec{k}$ dependence of the matrix element will be observed.

IV. DISCUSSION

A. Range of Validity

We have shown in Sec. III that the strong-field or Franz-Keldysh limit, given by Eq. (3.3b), can only be obtained by summing the perturbation expansion, given by Eq. (2.5), to all orders in the perturbation. It is therefore of interest to investigate the range of validity of Eq. (3.5). We do this by comparing directly in Fig. 1 line shapes for $\Delta\epsilon_1$ and $\Delta\epsilon_2$, calculated from Eqs. (3.3b) and (3.5) for the special case of a three-dimensional parabolic M_0 critical point defined by the three-dimensional interband energy

$$E_{cv}(\vec{k}) = E_g + \hbar^2\vec{k}^2/2\mu . \qquad (4.1)$$

Here, Eqs. (3.3) and (3.5) can be expressed as analytic closed-form field-induced changes in the dielectric function[24]:

$$\Delta\epsilon(\omega, \Gamma, \vec{\mathcal{E}}) = (2\pi i/\omega^2)\, QD^3(\hbar\theta)^{1/2}\{2\pi[e^{i\pi/3}\,\mathrm{Ai}'(z)\,\mathrm{Ai}'(w) + w\,\mathrm{Ai}(z)\,\mathrm{Ai}(w)] - z^{1/2}\} \qquad (4.2a)$$

$$\sim (2\pi i/\omega^2)\, QD^3(\hbar\theta)^{1/2}\{-\tfrac{1}{32}\, z^{-5/2}\} , \qquad (4.2b)$$

where the quantities not defined in Eqs. (2.18) are given by[25]

$$D = (2\mu/\hbar^2)^{1/2} , \qquad (4.2c)$$

$$(\hbar\theta)^3 = 4(\hbar\Omega)^3 , \qquad (4.2d)$$

$$z = e^{2\pi i/3}\, w = (E_+ - E_g)/\hbar\theta . \qquad (4.2e)$$

In Fig. 1, the real and imaginary projections of 10^3 times the bracketed parts of Eq. (4.2a), the convolution expression, and Eq. (4.2b), the perturbation approximation, are plotted for various values of $\Gamma/\hbar\theta$

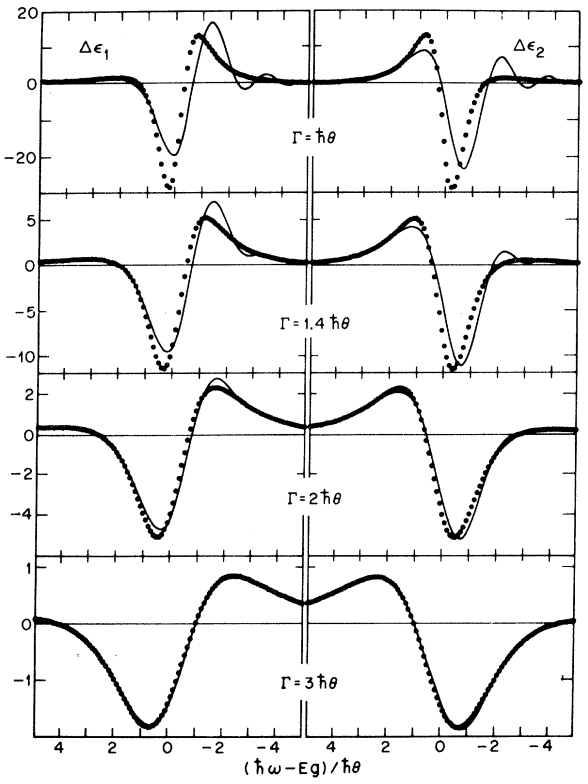

FIG. 1. Comparison of the perturbation (dotted line) and broadened-exact (solid line) forms of the electric-field-induced change in the dielectric function. $\Delta\epsilon = \Delta\epsilon_1 + i\Delta\epsilon_2$, as a function of the relative broadening $\Gamma/\hbar\theta = 2^{-2/3}\Gamma/\hbar\Omega$ for a three-dimensional M_0 critical point. Functions plotted: solid curves, $10^3\times$ {term in brackets in Eq. (4.2a)}; dotted curves, $10^3\times$ {term in brackets in Eq. (4.2b)}.

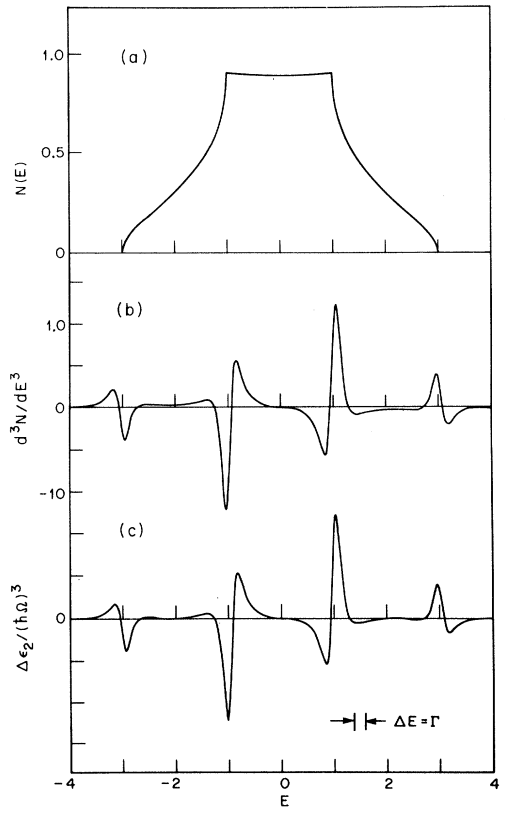

FIG. 2. Comparison of (a) the joint density of states $n(E)$ of a three-dimensional simple-cubic lattice, (b) its third derivative $(d/dE)^3 n(E)$, and (c) a composite modulation spectrum synthesized from M_0, M_1, M_2, and M_3 three-dimensional critical points from Table I. The quantity $n(E)$ is calculated from the integral representation of Ref. 29, and the energy units chosen place the M_0, M_1, M_2, and M_3 critical points at -3, -1, 1, and 3, respectively. In (b) and (c), $\Gamma = 0.2$ in the same units, as shown in the figure. In (c), each critical-point contribution extends one unit to either side of its respective critical-point energy, and the magnitudes are scaled in a 1 : 3 : 3 : 1 ratio to simulate the critical-point degeneracies in $n(E)$. For comparison purposes, the sign contributed by the prefactor $(\hbar\Omega)^3$ in Table I was neglected in (c): The M_0, M_1, and M_2 line shapes correspond to the field oriented in a positive-mass direction and the M_3 line shape is inverted.

$= 2^{-2/3}\Gamma/\hbar\Omega$ as a function of $(\hbar\omega - E_g)/\hbar\theta$. Of particular importance is the fact that the perturbation expression accurately reproduces the main structure centered about $\hbar\omega = E_g$. Subsidiary oscillations which appear in the exact formulation at small values of $\Gamma/|\hbar\Omega|$ are generated by higher-order resonant terms in the perturbation expansion, and are not contained in Eq. (4.2b). Figure 1 shows that Eqs. (4.2a) and Eq. (4.2b) are virtually identical for $\Gamma > 3|\hbar\Omega| \cong 2|\hbar\theta|$, and we take this as a working definition of the range of validity of the perturbation limit. For $\Gamma < 3|\hbar\Omega|$, the asymptotic limit will approximate the main features of the exact expression, but it becomes progressively less accurate as the broadening decreases or the field increases.

Whether or not the perturbation approximation is applicable to a specific experiment can be determined directly from the experimental line shape if we note that in this limit, the field enters only as a scaling prefactor $\vec{\mathcal{E}}^2$. From this square-law scaling property of Eq. (2.18b), it follows that limit is applicable whenever the spectral amplitude scales as $\vec{\mathcal{E}}^2$, and no change in the line shape (relative heights of line-shape extrema, or shifts in energy of line-shape features) is observed as the field changes. In practice, for typical values of momentum matrix elements ($P \sim 0.7\,\hbar/a_B$) and interband reduced masses ($\mu_T \sim 0.1 m_e$) the low-field limit is usually applicable whenever the rule of thumb $|\Delta R/R| < 10^{-4}$ is satisfied. Of particular importance to the experimentalist is the fact that, in this limit, experimental restrictions for obtaining "good" spectra are considerably relaxed: Since the line *shape* is field invariant, it is not necessary to use square-wave modulation nor even to modulate from flat band (unless quantitative conclusions are expected to be obtained from the amplitude of an experimental spectrum). Also, the appearance of the field *and* reduced mass as the scaling factor $(\vec{\mathcal{E}}^2/\mu)$ greatly simplifies the theoretical analysis of experimental spectra.

B. Relation to the Unperturbed Dielectric Function and Energy-Band Structure

Examination of Eq. (2.20) shows the field-induced change in the dielectric function in the perturbation limit is proportional to the third derivative of the unperturbed dielectric function. In particular, the imaginary part of this change is proportional to the third derivative of the density of states. This result shows electroreflectance is qualitatively different from other modulation techniques (e.g., piezoreflectance,[26] thermoreflectance,[27] wavelength derivative spectroscopy[28]) where the experimentally measured change in the dielectric function is given by the first derivative of the unperturbed dielectric function. This explains the generally sharper, better separated, and more richly structured spectra of electroreflectance, as compared to those obtained with other modulation techniques.

To illustrate this relationship, we compare in Fig. 2 the third derivative of the theoretical density-of-states curve of a three-dimensional simple-cubic lattice[29] with a composite curve of changes in the imaginary part of the dielectric constant, calculated from Eq. (3.5), for a sequence of parabolic three-dimensional M_0, M_1, M_2, and M_3 critical points. The functional forms of the changes

TABLE I. Field-induced changes in the complex dielectric function for one-, two-, and three-dimensional critical points in the perturbation or low-field limit. Branch cuts for one and three dimensions are defined for integer N as $\lim(E+i\Gamma-E_g)^{-N/2}=+E^{-N/2}$ as $E\to\infty$. l is the order of the critical point M_l, where $l=0$, 1, 2, 3 and $l\leq$ dimension. $D_i=(2|\mu_i|/\hbar^2)^{1/2}$ is the density of states for the ith symmetry axis. K_i is the k-space cutoff for the ith symmetry axis. $Q=(4\pi^2e^2/m^2)\,[2/(2\pi)^3]\,|\hat{\epsilon}\cdot\vec{P}_{cv}|^2$ and $(\hbar\Omega)^3=(\hbar^2e^2\mathcal{E}^2/8\mu_{\parallel})=\frac{1}{8}e^2(\vec{\mathcal{E}}\cdot\nabla_{\vec{k}})^2E_{cv}(\vec{k})$.

Dimension	$\Delta\epsilon(E,\Gamma,\vec{\mathcal{E}})\sim$
1	$\dfrac{5i^{l-1}D_xK_yK_zQ(\hbar\Omega)^3}{8\omega^2(E+i\Gamma-E_g)^{7/2}}$
2	$\dfrac{2i^{l-2}D_xD_yK_zQ(\hbar\Omega)^3}{3\omega^2(E+i\Gamma-E_g)^3}$
3	$\dfrac{\pi i^{l-3}D_xD_yD_zQ(\hbar\Omega)^3}{4\omega^2(E+i\Gamma-E_g)^{5/2}}$

in the dielectric function associated with each of these critical points are all obtainable from Eq. (4.2b); the general expression in practical units is given in Table I. The density of states $n(E)$ given in Fig. 2(a) is unbroadened. Its third derivative and the composite in Figs. 2(b) and 2(c), respectively, are both broadened by an amount $\Gamma=0.2$, in units of energy of the density-of-states curve (M_0 critical point at -3, M_1 at -1, etc.). The amplitude of the critical-point contributions in Fig. 2(c) have been scaled to agree with their counterparts in Fig. 2(b), and the sign of the prefactor $(\hbar\Omega)^3$ has been neglected as explained in the figure caption. The agreement between Figs. 2(b) and 2(c) is immediately apparent, and illustrates clearly the separation and independence of critical points to an ER spectrum in a nonparabolic band structure, and how well the simple line shapes calculated from parabolic approximations represent the nonparabolic band structure.[20]

The relative accuracy of simple parabolic line shapes in the three-dimensional case indicates that two- and one-dimensional simple parabolic line shapes, which approximate ER line shapes for continuum exciton E_0[8] and E_1 and $E_1+\Delta_1$ transitions[20,30] in semiconductors, and magnetoelectroflectance line shapes,[31] respectively, may also be useful. Accordingly, the functional forms of these line shapes are also given in Table I, and the line shapes associated with the changes in the real and imaginary parts of the dielectric functions, $\Delta\epsilon_1$ and $\Delta\epsilon_2$, for an M_0 critical point for each dimension are shown in Fig. 3. Since the line shape of $\Delta\epsilon=\Delta\epsilon_1+i\Delta\epsilon_2$ for any critical point of a given dimension can be represented as a linear combination of the two for the M_0 critical point, Fig. 3 is general. Note particularly the sharpening of the line shapes with decreasing dimension, a consequence of the corresponding sharpening of the density of states. The presence of two well-defined extrema on these curves is a general property, and these can be used to determine the critical point energy,[32] as well as the broadening and critical point type,[33] to high accuracy. Excitonic effects in experimental spectra act mainly to mix line shapes of adjacent critical points[34] and do not appreciably affect the above results.[35]

V. CONCLUSION

We have shown that perturbation theory and the classical strong-field theory both yield the same expression for the field-induced change in the dielectric function for weak fields where the broadening energy Γ is significantly larger than $|\hbar\Omega|$. Since the two derivations are based on substantially different assumptions, we conclude by examining their differences in more detail.

The strong-field theory is a stationary-state or constant-energy approach which requires the nonphysical assumption that electron scattering is nonexistent (the electrons set up standing waves in the Brillouin zone) in order to impose a condition on the phase-coherent part of the Bloch functions to enable normalization of the wave functions of the

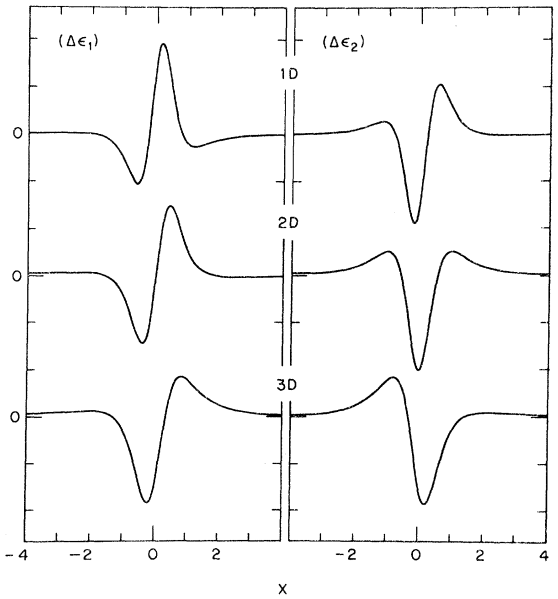

FIG. 3. Electric-field-modulated line-shape functions for one-, two-, and three-dimensional critical points. For each dimension $D=1$, 2, or 3, the curves are generated by the real and imaginary parts of the function $i^{-D}\times(X+i)^{-4+D/2}$, representing the line shapes of the real ($\Delta\epsilon_1$) and imaginary ($\Delta\epsilon_2$) field-induced change in the dielectric function for an M_0 transition of each dimension, as shown in Table I. Note the increase in width of the dominant structure with increasing dimensionality.

infinite potential resulting from the uniform field.[36] The perturbation approach avoids the difficulty of a stationary infinite potential by turning the potential on as a function of time and calculating the resulting evolution of the unperturbed wave functions. Thus the nonphysical requirement of zero scattering is avoided by perturbation theory where the broadening (turn-on) parameter plays a central part. For comparison, it is easy to show directly from the acceleration theorem[37] that the perturbation approach requires phase coherence of the electron only for a range

$$\frac{\Delta k}{K} = \frac{a}{2\pi}\,\Delta k = \frac{e\mathcal{E}a}{2\pi\Gamma}\ , \qquad (5.1)$$

where K is the width of the Brillouin zone, and a is the width of the unit cell in real space. This ratio is of order 0.01 for typical conditions involving higher interband transitions, indicating phase coherence over a physically reasonable range of 100 unit cells.[15]

A second difference is a direct consequence of the stationary-state vs time-dependent nature of the two derivations. The stationary-state wave functions of the strong-field theory represent exact diagonalization of the intraband part of the potential, and are linear combinations of *all* Bloch functions (of a single band). These new wave functions cannot be identified with any single Bloch state, whereas perturbation theory describes directly the time evolution of specific one-electron states. The inability to relate strong-field states to particular Bloch electrons was previously noted by Wannier and Van Dyke,[38] and has consequences for the description of the Stark ladder in ideal one-electron crystals in an electric field.[39] Stark steps do not appear in the perturbation approach.

The essential difference between our treatment and the high-frequency theory of Yacoby[40] is Yacoby's use of Houston functions[41] to describe the time evolution of the one-electron system. These representations of the wave function have a time-integrated (Floquet) phase factor, as contrasted to the perturbation phase factor increasing linearly in time, and yield the standard Franz-Keldysh result because they represent too accurate an approximation.

The mathematical simplicity of the perturbation formulation of electric field effects on the dielectric function of solids, and the relative independence of parameters determining line shape and magnitude, should simplify theoretical calculations and aid in determining band-structure parameters from experimental spectra. The factoring of the field and reduced mass as scaling parameters to a general line shape is a significant simplification which allows the energy gap, broadening parameter, exciton strength, and momentum matrix element to be extracted from the position, width, asymmetry, and magnitude, respectively, for a given resonance. This analysis will be treated in a later paper.[33]

---

[1]W. Franz, Z. Naturforsch. 13a, 484 (1958); L. V. Keldysh, Zh. Eksperim. i Teor. Fiz. 34, 1138 (1958) [Sov. Phys. JETP 7, 788 (1958)].

[2]K. Tharmalingam, Phys. Rev. 130, 2204 (1963); D. E. Aspnes, *ibid.* 147, 554 (1966).

[3]D. E. Aspnes, P. Handler, and D. F. Blossey, Phys. Rev. 166, 921 (1968), and references therein.

[4]J. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Phys. Rev. 127, 1918 (1962); S. S. Jha and N. Bloembergen, *ibid.* 171, 891 (1968), and references therein.

[5]P. N. Butcher and T. P. McLean, Proc. Phys. Soc. (London) 81, 219 (1963); 83, 579 (1964).

[6]F. N. H. Robinson, Bell System Tech. J. 46, 913 (1967).

[7]T. Nishino and Y. Hamakawa, J. Phys. Soc. Japan 26, 403 (1969); Y. Hamakawa, T. Nishino, and J. Yamaguchi, in *Proceedings of the Ninth International Conference on Semiconductor Physics, Moscow*, 1968, edited by S. M. Rykin and Yu. V. Shmartsev (Nauka, Leningrad, 1968), p.384.

[8]A. Frova and D. E. Aspnes, Phys. Rev. 182, 795 (1969); D. E. Aspnes and A. Frova, Phys. Rev. B 2, 1037 (1970); 3, 1511 (1971).

[9]H. Lange and E. Gutsche, Phys. Status Solidi 32, 293 (1969); E. Mohler, *ibid.* 38, 81 (1970).

[10]R. A. Forman, D. E. Aspnes, and M. Cardona, J. Phys. Chem. Solids 31, 227 (1970).

[11]D. S. Kyser and V. Rehn, Solid State Commun. 8, 1437 (1970).

[12]A. G. Aronov and G. E. Pikus, Fiz. Tverd. Tela 10, 825 (1968) [Sov. Phys. Solid State 10, 648 (1968)]; J. A. Van Vechten and D. E. Aspnes, Phys. Letters 30A, 346 (1969).

[13]J. A. Van Vechten, M. Cardona, D. E. Aspnes, and R. L. Martin, in *Proceedings of the Tenth International Conference on the Physics of Semiconductors* (U.S. AEC, Oak Ridge, Tenn., 1970), p. 82.

[14]D. E. Aspnes, Phys. Rev. Letters 26, 1429 (1971).

[15]D. E. Aspnes and J. E. Rowe, Solid State Commun. 8, 1145 (1970).

[16]The dipole approximation is rigorously correct for the dc field, and introduces negligible errors for the photon field.

[17]See, for instance, A. Messiah, *Quantum Mechanics* (Interscience, New York, 1962), pp. 724ff.

[18]H. R. Phillip and H. Ehrenreich, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1967), Vol. 3, pp. 101 ff.

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

[20]D. E. Aspnes and J. E. Rowe, in Ref. 13, p. 422.

[21]H. D. Rees, J. Phys. Chem. Solids 29, 143 (1968).

[22]D. E. Aspnes and N. Bottka, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, to be published), Vol. 9.

[23]E. O. Kane, J. Phys. Chem. Solids 12, 181 (1959); E. N. Adams, Phys. Rev. 107, 698 (1957); G. H. Wannier, *ibid.* 100, 1227 (1955).
[24]D. E. Aspnes, Phys. Rev. 153, 972 (1967).
[25]Throughout, the branch cut for half-integer powers of complex quantities is taken as the negative real axis and the branch is defined by

$$\lim_{E\to\infty}(E_+ - E_g)^{N/2} \to +(E-E_g)^{N/2}.$$

[26]M. Garfinkel, J. J. Tiemann, and W. E. Engeler, Phys. Rev. 148, 698 (1966).
[27]B. Batz, Solid State Commun. 4, 241 (1965).
[28]I. Balslev, Phys. Rev. 143, 636 (1966).
[29]J. Oitmaa, Solid State Commun. 9, 745 (1971).
[30]D. Brust, in *Methods of Computational Physics*, edited by B. Alder, S. Fernbach, and M. Rotenberg (Academic, New York, 1968), p. 60.
[31]S. H. Groves, C. R. Pidgeon, and J. Feinleib, Phys. Rev. Letters 17, 463 (1966).
[32]D. E. Aspnes and J. E. Rowe, Phys. Rev. Letters 27, 188 (1971).
[33]D. E. Aspnes and J. E. Rowe (unpublished).
[34]J. E. Rowe and D. E. Aspnes, Phys. Rev. Letters 25, 162 (1970).
[35]J. E. Rowe and D. E. Aspnes (unpublished).
[36]P. W. Argyres, Phys. Rev. 126, 1386 (1962).
[37]C. Kittel, *Quantum Theory of Solids* (Wiley, New York, 1963), p. 190.
[38]G. H. Wannier and J. P. Van Dyke, J. Math. Phys. 9, 899 (1968).
[39]G. H. Wannier, Phys. Rev. 117, 432 (1960).
[40]Y. Yacoby, Phys. Rev. 169, 610 (1968).
[41]W. V. Houston, Physics 57, 184 (1940).

PHYSICAL REVIEW B VOLUME 5, NUMBER 10 15 MAY 1972

Theory of the Piezoelectricity of Zinc-Blende-Type and Wurtzite-Type Crystals

Takehiko Hidaka
Electrotechnical Laboratory, MITI, 5-4-1, Mukodai, Tranashi-shi, Tokyo, Japan
(Received 27 January 1971)

A theory of piezoelectricity in zinc-blende-type and wurtzite-type crystals is given. The charge-transfer effect contributes mainly to the piezoelectric constant. The assumption of the second-nearest-neighbor electron interaction (π-electronic energy) in the crystal enables us to predict theoretically the charge-transfer value, so that the piezoelectric coefficients can be calculated. The π-electronic interaction energy is found to be $0.5 \sim 0.8$ times the σ-bonding (tetrahedral-bonding) energy in wurtzite crystals.

I. INTRODUCTION

Zinc-blende-type (ZB) and wurtzite-type (Wu) crystals are piezoelectric. The piezoelectric coefficients of these crystals have been obtained experimentally.[1,2] The most striking aspect of the experimental data for the piezoelectric constants of ZB and Wu crystals is that $\sqrt{3}\,e_{14}$(ZB), or its equivalent e_{33}(Wu), reverses sign on going from II-VI to III-V crystals. But no successful theory has been proposed which enables us to predict the piezoelectric constants or to understand the origin of the effect.

To explain these data, a simple theory of the rigid-ion model had been proposed.[1] In that theory it is assumed that the charges of anion and cation in the crystals under stress remain constant, and only the relative bond angle changes. Then, for ZB, the piezoelectric constant e_{14} is given by

$$e_{14} = \tfrac{3}{16} e^*/d^2 \quad , \tag{1}$$

where e^* is the effective charge of the lattice and d is the bond length from anion to cation. The piezoelectricity of the cubic ZnS crystal shows good agreement with the theory,[3] but this is not the case for all other crystals.[1]

Arlt and Quadflieg[2] introduced the charge redistribution effect (i.e., change in ionicity) for the explanation of experimental data on piezoelectricity. They gave the relation

$$\begin{aligned} e_{14} = {} & (\text{the displacement of ionic charge}) \\ & + (\text{the internal displacement of the electron cloud relative to the atomic nucleus}) \\ & + (\text{the effect due to the strain-induced change in ionicity}). \end{aligned} \tag{2}$$

The first and second terms in Eq. (2) are equal to Eq. (1), because the effective charge e^* includes the effect of the displacement of the electronic charge, i.e., electronic polarization.[4] The third term is the charge-transfer effect under the strain. They estimated the contribution of the change in ionicity such that the positively charged atoms lose their charges linearly in expanding the bond length; i.e., if the bond length is expanded a factor of 2, all atoms would become neutral.

Phillips and Van Vechten[5] showed that there is some correlation between the magnitude of the charge-transfer effect and the ionicity f_i of the crystal. They gave the relation