PHYSICAL REVIEW B VOLUME 6, NUMBER 12 15 DECEMBER 1972

Energy-Band Theory of the Second-Order Nonlinear Optical Susceptibility of Crystals of Zinc-Blende Symmetry

D. E. Aspnes
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 21 May 1971)

The zero-frequency limit of the second-order nonlinear optical susceptibility $\chi^{(2)}_{ijk}$ is calculated within the framework of energy-band theory. Both vector and scalar potential representations of the electric field are investigated. The apparent divergences of standard finite-frequency expressions in the limit of zero frequency are shown to vanish. The two- and three-band contributions in the scalar potential representation, describing the electric field effect on the coherent (phase factor) and cell-periodic parts of the Bloch function, respectively, are shown to combine as required by gauge invariance to yield the result obtained in the vector potential representation, where only three-band terms contribute. In crystals of zinc-blende symmetry, virtual electronic processes involving one valence and two conduction bands, originating from the Γ_{15} valence and Γ_1 and Γ_{15} conduction states at $\vec{k}=0$, dominate virtual-hole processes involving one conduction and two valence bands. The sign of $\chi^{(2)}_{123}$ is related to the Brillouin-zone average of the ordering with respect to energy of the lowest conduction bands. The magnitude of $\chi^{(2)}_{123}$ is related to the average over the Brillouin zone of the inverse fifth power of the local energy separation $E_{cv}(\vec{k})$ between the valence and lowest conduction bands, in contrast to the inverse third power average of the same quantity known for the linear susceptibility $\chi^{(1)}$. This enhances the contribution to $\chi^{(2)}_{123}$ from small-gap regions of the band structure in a manner similar to that also found for $\chi^{(3)}$. The value of $\chi^{(2)}_{123}$ calculated for zinc-blende crystals in both the constant energy gap and the parabolic critical-point models is consistently lower than that observed experimentally. Extended calculations using more realistic analytic approximations to the local energy-band structure suggest that this discrepancy is due in part to band nonparabolicity effects with additional contributions from enhancement of oscillator strength from the electron-hole Coulomb interaction.

I. INTRODUCTION AND SUMMARY

The second-order nonlinear optical susceptibility $\chi^{(2)}_{ijk}(\omega_1, \omega_2)$ is a third-rank tensor whose dispersion properties have been developed within the framework of both localized wave functions[1-5] and band theory.[6-13] Although its zero-frequency limit $\chi^{(2)}_{ijk}$ in the bonding orbital picture[14-16] now appears to be well understood, the corresponding limit in terms of energy-band theory, the complement of bond theory, is not clear. Standard finite-frequency expressions[1,7] appear to diverge strongly as the zero-frequency limit is approached, and approximate nondivergent expressions,[11,17,18] with the exception of the bond charge model of Levine,[15] restrict $\chi^{(2)}_{ijk}$ to positive values in contrast to experimental results.[19] In addition, the relationship among various two-band and three-band expressions obtained in scalar potential or vector potential representations of the electric field perturbation is confusing. Phillips and Van Vechten[17] obtained an expression, later modified by Kleinman,[18] based primarily on the two-band contribution of the scalar potential representation, whereas Bell[11] obtained an expression using only the three-band term of this representation. Yet in the vector potential representation, only terms of three-band form contribute.

The objective of this paper is to obtain an explicit expression for $\chi^{(2)}_{123}$ for zinc-blende crystals in both scalar potential and vector potential representations of the field perturbation in order to investigate these questions and to examine $\chi^{(2)}_{123}$ in terms of band theory. We use the one-electron and dipole approximations, taking Bloch functions as basis states and expressing $\chi^{(2)}_{123}$ in terms of momentum matrix elements and energy eigenvalues. Energy bands are assumed to be either filled or empty, and only the electronic contribution to $\chi^{(2)}_{123}$ is treated. The exact zero-frequency limit for $\chi^{(2)}_{ijk}$ in the one-electron approximation is obtained in Sec. II from the general expressions obtained by Butcher and McLean.[7] Following this, we restrict our attention to $\chi^{(2)}_{123}$ and obtain this quantity in the scalar potential representation. The results show the following: (i) The divergences which arise in taking the zero-frequency limit of the general expression $\chi^{(2)}_{ijk}(\omega_1, \omega_2)$ vanish identically owing to symmetry, leaving a well-behaved expression. (ii) The two- and three-band contributions in the scalar potential representation have opposite signs, but the two-band contribution is always larger. (iii) The apparent contradiction that a two-band term arises in the scalar potential representation but not in the vector potential representation is one of form only—the two-band term (representing the electric field effect on the coherent or phase part of the Bloch functions) simply expresses in closed form a summation over a third band by introducing a k-space gradient operator.

(iv) Three-band processes are of two complementary types—virtual-electron transitions from a valence to two conduction bands, or virtual-hole transitions from a conduction to two valence bands, both of which may be treated within the same formalism although typically only the virtual-electron process is significant. (v) The sign of $\chi^{(2)}_{123}$ is related to the Brillouin-zone average of the ordering with respect to energy of the lowest conduction bands. (vi) The contribution to $\chi^{(2)}_{123}$ at any given wave vector $\vec{k}$ is nearly proportional to the inverse fifth power of the valence- to lowest conduction-band energy separation, except for regions of very small separation when the contribution is nearly proportional to the inverse fourth power of this energy separation. Point (v) shows that $\chi^{(2)}_{123}$ can have either sign, and because $\chi^{(2)}_{123}$ is a three-band quantity (unlike $\chi^{(1)}_{ij}$ or $\chi^{(3)}_{ijkl}$), it contains information about the average relative ordering of the conduction bands. Point (vi) shows that the contribution to $\chi^{(2)}_{123}$ from small band-gap regions is enhanced with respect to the contributions from the same region to the linear susceptibility, where the band average is over $[E_{cv}(\vec{k})]^{-3}$. {Note, however, that since the average of $[E_{cv}(\vec{k})]^{-5}$ must be taken over the entire Brillouin zone, there is no simple power law relating $\chi^{(2)}_{123}$ to the fundamental direct gap or to any other feature in linear optical spectra.} Point (vi) also shows that $\chi^{(2)}_{123}$ is relatively insensitive to features of the higher conduction band and thus behaves in many respects as a two-band function.

In Sec. III, we apply this expression to crystals of zinc-blende symmetry using the constant matrix element approximation. We are interested here only in general properties, so $\chi^{(2)}_{123}$ will not be evaluated from detailed band-structure calculations, but rather will be evaluated approximately in the constant energy gap model,[1,8,11,17,18] extended to include explicitly the contributions of high-symmetry regions by means of parabolic approximations to the actual band structure.[20–22] Here, the dominant contribution to $\chi^{(2)}_{123}$ comes from virtual electron transitions involving the upper valence and two lower conduction bands, arising from the sp^3 states of Γ_{15} (valence), Γ_1 (conduction), and Γ_{15} (conduction) symmetry at $\vec{k}=0$. If the band transforming as Γ_1 lies on the average between the other two, as is the usual case, then $\chi^{(2)}_{123}$ is positive. These simple models enable $\chi^{(2)}_{123}$ to be calculated with no adjustable parameters, but the theoretical values obtained are consistently smaller than experimentally measured values, as has also been observed for $\chi^{(3)}_{1111}$.[22] We show for GaAs that this discrepancy is reduced by extending the simple parabolic models to include band nonparabolicity and Coulomb effects into the density of states. Whether nonparabolicity and Coulomb effects would account for the entire discrepancy can only be decided in a detailed calculation of $\chi^{(2)}_{123}$ using an accurate energy-band structure. Other possible corrections, and the connections to previous models, are discussed in Sec. IV.

II. THEORY

The zero-frequency limit $\chi^{(2)}_{ijk}$ of $\chi^{(2)}_{ijk}(\omega_1, \omega_2)$ can be calculated directly from perturbation theory, where the electric field may be represented either as a vector potential or as a scalar potential with a time dependence $e^{\eta t}$, $\eta \to 0^+$. Alternatively, existing expressions[7] for $\chi^{(2)}_{ijk}(\omega_1, \omega_2)$ obtained in the vector potential representation may be evaluated using the complex frequency $\omega_1 = \omega_2 = i\eta$, $\eta \to 0^+$. The turn-on parameter η is needed for investigating the behavior of various orders of divergence as $\eta \to 0$.

A. Vector Potential Representation

We evaluate explicitly the zero-frequency limit of the expression obtained by Butcher and McLean[7] to describe the second-order conductivity relating the crystal current component $J_\mu(t)$ to two applied fields represented in component form as $\mathcal{E}_{1\alpha}e^{-i\omega_1 t}$ and $\mathcal{E}_{2\beta}e^{-i\omega_2 t}$, respectively. In their paper, the results are summarized in Eq. (25) and the discussion preceding Eq. (37). The two fields become identical in the zero-frequency limit, where $\omega_1 = \omega_2 = i\eta$, and are represented by the vector potential

$$\vec{A}(t) = -(c/\eta)e\vec{\mathcal{E}}e^{\eta t}. \tag{2.1}$$

The crystal polarization $\vec{P}(t)$ is the time integral of the current. In the one-electron dipole approximation, following the component notation of Butcher and McLean,[7]

$$\chi^{(2)}_{\mu\alpha\beta}(\eta) = -\frac{e^3}{4m^3\eta^3 V}\sum_{vnn'\vec{k}}\left(\frac{P^\mu_{vn}(P^\alpha_{nn'}P^\beta_{n'v}+P^\beta_{nn'}P^\alpha_{n'v})}{(E_{nv}-i2\hbar\eta)(E_{n'v}-i\hbar\eta)}\right.$$
$$+\frac{P^\mu_{nn'}(P^\alpha_{n'v}P^\beta_{vn}+P^\beta_{n'v}P^\alpha_{vn})}{(E_{n'v}-i\hbar\eta)(E_{nv}+i\hbar\eta)}$$
$$\left.+\frac{P^\mu_{n'v}(P^\alpha_{vn}P^\beta_{nn'}+P^\beta_{vn}P^\alpha_{nn'})}{(E_{n'v}+i2\hbar\eta)(E_{nv}+i\hbar\eta)}\right). \tag{2.2}$$

In terms of the one-electron Bloch wave functions $|n, \vec{k}\rangle$ we see

$$P^\epsilon_{n'n} = \hat{\epsilon}\cdot\langle n'\vec{k}\,|\,\vec{P}\,|\,n\vec{k}\rangle \quad \text{(momentum matrix element)}, \tag{2.3a}$$
$$E_{n'n} = E_{n'}(\vec{k}) - E_n(\vec{k}) \quad \text{(interband energy)}. \tag{2.3b}$$

The only restriction on Eq. (2.2) is that the valence state $|v, \vec{k}\rangle$ shall be occupied. If all three states $|v, \vec{k}\rangle$, $|n, \vec{k}\rangle$, and $|n'\vec{k}\rangle$ are occupied, their total contribution is identically zero. Therefore, at least one of the indices n, n' must correspond to an empty conduction band, but otherwise n and n' can range over all bands in the crystal (including v). It follows directly that the only two possible

processes, one involving a virtual hole (conduction-valence-valence or cvv), and the other a virtual electron (valence-conduction-conduction or vcc), can be expressed in complementary terms in the same formalism. The third-power divergence of the prefactor as $\eta \to 0$ in Eq. (2.2) is characteristic of the vector potential representation of the electric field.

Let the ground state of the system consist of filled bands and let no magnetic fields be present. Then, by time-reversal invariance, for any point $\vec{k}$ in the Brillouin zone and for any band pair (n', n) having an interband separation energy $E_{n'n}(\vec{k})$, there must exist a second band pair (m', m) such that[23]

$$E_{m'm}(\vec{k}) = E_{n'n}(-\vec{k}) , \tag{2.4a}$$

$$\vec{P}_{m'm}(\vec{k}) = -\vec{P}_{nn'}(-\vec{k}) . \tag{2.4b}$$

When this symmetry is included explicitly in Eq. (2.2) by manipulating the dummy indices n and n', we obtain

$$\chi^{(2)}_{\mu\alpha\beta}(\eta) = -\frac{ie^3\hbar}{4\eta^2 m^3 V} \sum_{\substack{vnn' \\ \vec{k}}} \left(\frac{(2E_{n'v} + E_{nv}) P^{\mu}_{vn} (P^{\alpha}_{nn'} P^{\beta}_{n'v} + P^{\beta}_{nn'} P^{\alpha}_{n'v})}{(E^2_{n'v} + \hbar^2\eta^2)(E^2_{nv} + 4\hbar^2\eta^2)} - \frac{(E_{n'v} - E_{nv}) P^{\mu}_{nn'} (P^{\alpha}_{n'v} P^{\beta}_{vn} + P^{\beta}_{n'v} P^{\alpha}_{vn})}{(E^2_{n'v} + \hbar^2\eta^2)(E^2_{nv} + \hbar^2\eta^2)} \right.$$

$$\left. - \frac{(E_{n'v} + 2E_{nv}) P^{\mu}_{n'v} (P^{\alpha}_{vn} P^{\beta}_{nn'} + P^{\beta}_{vn} P^{\alpha}_{nn'})}{(E^2_{n'v} + 4\hbar^2\eta^2)(E^2_{nv} + \hbar^2\eta^2)} \right) . \tag{2.5}$$

Note that the prefactor now diverges only as η^2. A short calculation shows that $\chi^{(2)}_{\mu\alpha\beta}$ is real, and that the one-band $(n = n' = v)$ term vanishes identically regardless of crystal symmetry. For crystals of zinc-blende symmetry where there is only one nonvanishing component $\chi^{(2)}_{xyz} = \chi^{(2)}_{123}$, both the field and the polarization are parallel in the $\langle 111 \rangle$ bonding, or ξ, direction.[11] In this case we may take $\mu = \alpha = \beta = \xi$, and the two-band $(n' = n \neq v;\ n' \neq n = v;\ n \neq n' = v)$ terms also vanish identically. Therefore, only the three-band $(n' \neq n,\ n' \neq v,\ n \neq v)$ terms contribute in zinc-blende crystals when the electric field is represented as a vector potential, a consequence of the dipole approximation and the fact that *both* ω_1 and ω_2 approach zero in the general expression.

When $E^2_{n'v},\ E^2_{nv} \gg \hbar^2\eta^2 \to 0$ in Eq. (2.5), the energy denominators may be expanded, giving a quadratically diverging term

$$\chi^{(2)}_{\mu\alpha\beta}(\eta \to 0) = -\frac{ie^3\hbar}{4\eta^2 m^3 V} \sum_{\substack{vnn' \\ \vec{k}}}{}' E^{-2}_{n'v} E^{-2}_{nv} [(2E_{n'v} + E_{nv}) P^{\mu}_{vn} (P^{\alpha}_{nn'} P^{\beta}_{n'v} + P^{\beta}_{nn'} P^{\alpha}_{n'v}) - (E_{n'v} - E_{nv}) P^{\mu}_{nn'} (P^{\alpha}_{n'v} P^{\beta}_{vn} + P^{\beta}_{n'v} P^{\alpha}_{vn})$$

$$- (E_{n'v} + 2E_{nv}) P^{\mu}_{n'v} (P^{\alpha}_{vn} P^{\beta}_{nn'} + P^{\beta}_{vn} P^{\alpha}_{nn'})] \tag{2.6}$$

and a well-behaved term

$$\chi^{(2)}_{\mu\alpha\beta} = +\frac{ie^3\hbar^3}{4m^3 V} \sum_{\substack{vnn' \\ \vec{k}}}{}' E^{-4}_{n'v} E^{-4}_{nv} [(2E_{n'v} + E_{nv})(4E^2_{n'v} + E_{nv}) P^{\mu}_{vn} (P^{\alpha}_{nn'} P^{\beta}_{n'v} + P^{\beta}_{nn'} P^{\alpha}_{n'v})$$

$$- (E_{n'v} - E_{nv})(E^2_{n'v} + E^2_{nv}) P^{\mu}_{nn'} (P^{\alpha}_{n'v} P^{\beta}_{vn} + P^{\beta}_{n'v} P^{\alpha}_{vn}) - (E_{n'v} + 2E_{nv})(E^2_{n'v} + 4E_{nv}) P^{\mu}_{n'v} (P^{\alpha}_{vn} P^{\beta}_{nn'} + P^{\beta}_{vn} P^{\alpha}_{nn'})] . \tag{2.7}$$

The primes on the summation signs indicate $n' \neq n$, $n' \neq v$, and $n \neq v$. The quadratically diverging term is clearly nonphysical unless the bracketed part either vanishes point by point or in summation. This term vanishes point by point for zinc-blende crystals as can be seen by setting $\mu = \alpha = \beta = \xi$. Presumably, both this term and the two-band contribution from Eq. (2.5) should always either cancel each other, or else vanish in summation over $\vec{k}$, but we have not found a way to prove this for an arbitrary crystal class.

The well-behaved term, Eq. (2.7), is the correct zero-frequency limit of $\chi^{(2)}_{\mu\alpha\beta}$. Its form may be greatly simplified for crystals of zinc-blende symmetry, by evaluating it in the bonding direction ξ. For sign consistency, let ξ represent the *positive* $\langle 111 \rangle$ orientation of the electric field [electric field pointing *from* the metal atom at $-\frac{1}{8}a_0(1, 1, 1)$ *to* the anion at $+\frac{1}{8}a_0(1, 1, 1)$ in the unit cube in standard coordinates[14]]; then $\mu = \alpha = \beta = \xi$, and $\chi^{(2)}_{123}$ is given by

$$\chi^{(2)}_{123} = \tfrac{1}{2}\sqrt{3}\, \chi^{(2)}_{\xi\xi\xi} , \tag{2.8}$$

where[24]

$$\chi^{(2)}_{\xi\xi\xi} = +\frac{3ie^3\hbar^3}{2m^3 V} \sum_{\substack{n'nv \\ \vec{k}}}{}' P^{\xi}_{vn} P^{\xi}_{nn'} P^{\xi}_{n'v} \times \left(\frac{E_{n'n}}{E^4_{n'v} E^4_{nv}} (2E^2_{n'v} + 3E_{n'v}E_{nv} + 2E^2_{nv}) \right) . \tag{2.9}$$

Equation (2.9) is different from the explicit band-structure expressions previously obtained by Butcher and McLean[7] and by Bell,[11] which is not obvious here but which will be shown in Sec. II B.[25]

B. Scalar Potential Representation

In this section, we consider for simplicity only the term $\chi^{(2)}_{\xi\xi\xi}$, and rederive Eq. (2.9) in the Coulomb gauge where the electric field of Sec. II A is represented by a scalar potential $-\vec{\mathcal{E}}\cdot\vec{x}e^{\eta t}$, $\eta\to 0^+$. By the invariance of the Hamiltonian to gauge transformations, it is clear that the final result must be the same.[26] Nevertheless, the scalar potential representation leads directly to nonvanishing contributions from both two-band and three-band terms in contrast to the results of Sec. II A where only three-band terms remain. This apparent contradiction is due to the intraband transformation properties of the dipole matrix element $\langle n'k'|\vec{x}|nk\rangle$ between Bloch states, and has resulted in confusion in the correct expression for $\chi^{(2)}_{\xi\xi\xi}$ in the zero-frequency limit in the Bloch function representation. We show here why the previous zero-frequency expressions obtained by Butcher and McLean[7] and by Bell[11] are not complete, and do not agree with the correct expression given by Eq. (2.9).

The one-electron wave function $\psi_{n\vec{k}}(t)$ evolves from the unperturbed Bloch eigenfunction $|nk\rangle e^{-i\omega_{n\vec{k}}t}$ of the Hamiltonian H_0 under the action of the perturbation term

$$H' = +e\vec{\mathcal{E}}_0\cdot\vec{x}e^{\eta t}, \quad \eta\to 0^+ \tag{2.10}$$

according to

$$\psi_{n\vec{k}}(t) = [1 + G^0_{n\vec{k}}H' + G^{00}_{n\vec{k}}H'G^0_{n\vec{k}}H' + \cdots]|nk\rangle e^{-i\omega_{n\vec{k}}t}, \tag{2.11}$$

where

$$G^0_{n\vec{k}} = 1/(E_{n\vec{k}} + i\hbar\eta - H_0), \tag{2.12a}$$

$$G^{00}_{n\vec{k}} = 1/(E_{n\vec{k}} + 2i\hbar\eta - H_0). \tag{2.12b}$$

The second-order susceptibility can be calculated by evaluating the polarization operator directly to second order in $\vec{\mathcal{E}}_0$:

$$\vec{P} = \frac{-e}{V}\sum_{n\vec{k}}\langle\psi_{n\vec{k}}|\vec{x}|\psi_{n\vec{k}}\rangle = \underline{\chi}^{(1)}\cdot\vec{\mathcal{E}} + \underline{\chi}^{(2)}:\vec{\mathcal{E}}\vec{\mathcal{E}} + \cdots. \tag{2.13}$$

We find

$$\begin{aligned}\chi^{(2)}_{\xi\xi\xi}(\eta) = -\frac{e^3}{V}\sum_{\vec{k}_0,\vec{k},\vec{k}',\vec{k}''}{}^{v,n,n'} &\langle v\vec{k}_0|x^\xi|n\vec{k}'\rangle \\ &\times\langle n\vec{k}'|x^\xi|n'\vec{k}''\rangle\langle n'\vec{k}''|x^\xi|v\vec{k}\rangle \\ &\times[(E_{n\vec{k}'} - E_{v\vec{k}} - i2\hbar\eta)^{-1}(E_{n'\vec{k}''} - E_{v\vec{k}} - i\hbar\eta)^{-1} \\ &+ (E_{n\vec{k}'} - E_{v\vec{k}_0} + i\hbar\eta)^{-1}(E_{n'\vec{k}''} - E_{v\vec{k}} - i\hbar\eta)^{-1} \\ &+ (E_{n\vec{k}'} - E_{v\vec{k}_0} + i\hbar\eta)^{-1}(E_{n'\vec{k}''} - E_{v\vec{k}_0} + i2\hbar\eta)^{-1}].\end{aligned} \tag{2.14}$$

Note that the prefactor is well behaved in the limit $\eta\to 0$, in contrast to the result obtained in the vector potential representation.

The form of Eq. (2.14) has been influenced by the fact that matrix elements of $\vec{x}$ in the Bloch function representation are not well defined, since Bloch functions are eigenfunctions of the momentum operator and extend in principle over all space. The meaning of interband Bloch matrix elements of $\vec{x}$ can be interpreted by taking the expectation value of the commutator identity $[\vec{x}, H] = (i\hbar/m)\vec{P}$. Intraband matrix elements can be determined either by direct evaluation from the Bloch form $e^{i\vec{k}\cdot\vec{x}}u_n(\vec{k},\vec{x})$, or, alternatively, by using the definition of the group velocity $\vec{P}_{nn}(\vec{k}) = (\hbar/m)\nabla_{\vec{k}}E_n(\vec{k})$ with the above commutator. One finds[27]

$$\langle n'k'|\vec{x}|nk\rangle = \begin{cases}\dfrac{i\vec{P}_{n'n}(\vec{k})}{m\omega_{n'n}(\vec{k})}\delta_{\vec{k}'\vec{k}}, & n'\neq n \quad \text{(inter)} \quad (2.15a)\\ +i\delta_{\vec{k}'\vec{k}}[\nabla_{\vec{k}}f(\vec{k}',\vec{k})], & n'=n \quad \text{(intra)} \quad (2.15b)\end{cases}$$

where $f(\vec{k}',\vec{k})$ is the function of $\vec{k}$ and $\vec{k}'$ multiplying the matrix element, and the δ function is evaluated following differentiation. Equation (2.15b) requires that the wave vectors $\vec{k}$ belonging to the (same) states $\langle\psi_{nk}|$ and $|\psi_{nk}\rangle$ appearing in Eq. (2.13) be individually distinguished. We do this by adding a subscript 0 to the wave vector of the left-hand state. The final δ function between $\vec{k}$ and $\vec{k}_0$ which results is eliminated by the extra sum over $\vec{k}_0$ in Eq. (2.14).

Using Eqs. (2.15), we obtain from Eq. (2.14) the following contributions, listed in order of increasing number of contributing bands.

a. One-band term. We find

$$\chi^{(2)}_{\xi\xi\xi}{}_1 = -\frac{e^3}{4\hbar^3\eta^3 V}\sum_{n\vec{k}}\frac{\partial^3}{\partial k_\xi^3}E_{n\vec{k}}. \tag{2.16}$$

This is the low-frequency limit obtained by Butcher and McLean.[7] This term diverges as the third power of η if the sum is finite, but since the bands are either filled or empty, the summation over $\vec{k}$ vanishes identically. The analogous result in the linear susceptibility is well known, where it leads to the Thomas–Reiche–Kuhn sum rule in the Bloch representation.[28] Since this term vanishes identically for filled bands, it does not represent the true zero-frequency limit except in nonequilibrium situations which cannot be treated in perturbation theory but require the use of the Boltzmann equation. The sum of all third-power divergent terms which would have arisen in the vector-potential representation if time-reversal invariance had not been used in Sec. II A, could have been transformed into Eq. (2.16) through the use of sum rules. Therefore, the zero-frequency limit of $\chi^{(2)}_{\xi\xi\xi}$ obtained by Butcher and McLean is not complete.

b. Two-band terms. The three possible combinations of two-band terms ($n'\neq n=v$, $n'=n\neq v$, $n\neq n'=v$) can be transformed by straightforward

but tedious calculations into the following form:

$$\chi^{(2)}_{\xi\xi\xi_2}(\eta) = -\frac{3e^3\hbar}{2\eta m^2}\sum_{\substack{n\neq v\\ \vec{k}}}\frac{\partial}{\partial k_\xi}\left(\frac{P^\xi_{vn}P^\xi_{nv}}{E^3_{nv}}\right) - i\frac{3e^3\hbar^2}{2m^2}\sum_{\substack{n\neq v\\ \vec{k}}}E^{-2}_{nv}\left[\frac{P^\xi_{vn}}{E_{nv}}\frac{\partial}{\partial k_\xi}\left(\frac{P^\xi_{nv}}{E_{nv}}\right) - \left(\frac{\partial}{\partial k_\xi}\frac{P^\xi_{vn}}{E_{nv}}\right)\frac{P^\xi_{nv}}{E_{nv}}\right]. \quad (2.17)$$

Both terms are real. The first term diverges as η^{-1} but for filled bands the summation vanishes. The second term is well defined in the limit $\eta \to 0$ and *does not vanish*: It is the second-order analog of the term which describes the perturbation limit of the Franz–Keldysh effect in third order,[29] and represents the contribution of the phase factor or long-range coherency part of the Bloch function in the scalar potential representation of the electric field.

It is surprising in view of Eq. (2.9) that a two-band term should contribute. This apparent contradiction can be resolved by explicitly evaluating the gradient operator by means of $\vec{k}\cdot\vec{p}$ perturbation theory. We find[27]

$$\frac{\partial}{\partial k_\xi}P^\xi_{nv} = -\frac{P^\xi_{nv}}{E_{nv}}\frac{\partial E_{nv}}{\partial k_\xi} - \frac{2\hbar}{m}\sum_{n'}{}'P^\xi_{nn'}P^\xi_{n'v}\left(\frac{1}{E_{n'n}}+\frac{1}{E_{n'v}}\right). \quad (2.18)$$

It follows directly from Eq. (2.18) that the expanded form of Eq. (2.17) is

$$\chi^{(2)}_{\xi\xi\xi_2} = +\frac{3ie^3\hbar^3}{2m^3}\sum_{\substack{n',v,v\\ \vec{k}}}{}'P^\xi_{vn}P^\xi_{nn'}P^\xi_{n'v}\left[E^{-1}_{n'n}(E^{-4}_{n'v}+E^{-4}_{nv}) + E^{-1}_{nv}E^{-1}_{n'v}(E^{-3}_{nv}-E^{-3}_{n'v})\right]. \quad (2.19)$$

This is a three-band expression. The operator ∇_k simply represents in closed form a summation over a complete set of intermediate states, and the contradiction is resolved. This is a "new" term in the sense that it has not been included previously in Bloch function representations of $\chi^{(2)}$.

c. Three-band term. The only three-band term in the scalar potential representation is independent of η to lowest order and can be written

$$\chi^{(2)}_{\xi\xi\xi_3} = +\frac{3ie^3\hbar^3}{2m^3V}\sum_{\substack{n',n,v\\ \vec{k}}}{}'P^\xi_{vn}P^\xi_{nn'}P^\xi_{n'n}\left[E^{-1}_{n'n}(-2E^{-2}_{n'v}E^{-2}_{nv})\right]. \quad (2.20)$$

This term is also real, and this is the term considered by Bell.[11] Its sign is opposite that of Eq. (2.19). Note that the term in parentheses in Eq. (2.20) completes the square of the first term in parentheses in Eq. (2.19), so that the "two-band" contribution is always the larger of the two and determines the sign of the complete expression. A quick calculation shows that the sum of Eqs. (2.19) and (2.20) is equal to Eq. (2.9), demonstrating gauge invariance explicitly.

The relative contributions of the "two-band" and "three-band" terms to $\chi^{(2)}_{\xi\xi\xi}$ are shown in Fig. 1. The curves plotted are the square-bracketed parts of Eq. (2.19), Eq. (2.20) times -1, and Eq. (2.9), each multiplied by the smaller gap (either $E_{n'v}$ or E_{nv}) to the fifth power. The sign of the curves in Fig. 1 assumes that $E_{n'v}$, $E_{nv} > 0$ (vcc or virtual-electron process) and that E_{nv} is the smaller gap; if $E_{n'v}$ were smaller, then each expression would reverse its sign, and the sign of the ordinate scale in Fig. 1 would also have to be reversed. The point of Fig. 1 is to demonstrate that neither a two-band nor a three-band model will be adequate enough to completely describe $\chi^{(2)}_{\xi\xi\xi}$ in the scalar potential representation, although there exist certain ranges where each is approximately equal in magnitude to the exact result (if one is willing to ignore the fact that since two- and three-band contributions have opposite sign, then one of the two must have the wrong sign). It is also interesting to note that for the range which is physically interesting in most crystals, $1.25 < E_{n'v}/E_{nv} < 2.3$, the total is equal to 1 within 10%. In this approximation Eq. (2.9) becomes

$$\chi^{(2)}_{\xi\xi\xi} \cong +\frac{3ie^3\hbar^3}{2m^3V}\sum_{\substack{n',n,v\\ \vec{k}}}{}'\frac{P^\xi_{vn}P^\xi_{nn'}P^\xi_{n'v}}{E^5_{nv}}, \quad 1.25 < \frac{E_{n'v}}{E_{nv}} < 2.3; \quad (2.21)$$

i.e., we have the remarkable result that the exact value of the larger gap is relatively unimportant (except for lying within the given limits), and that

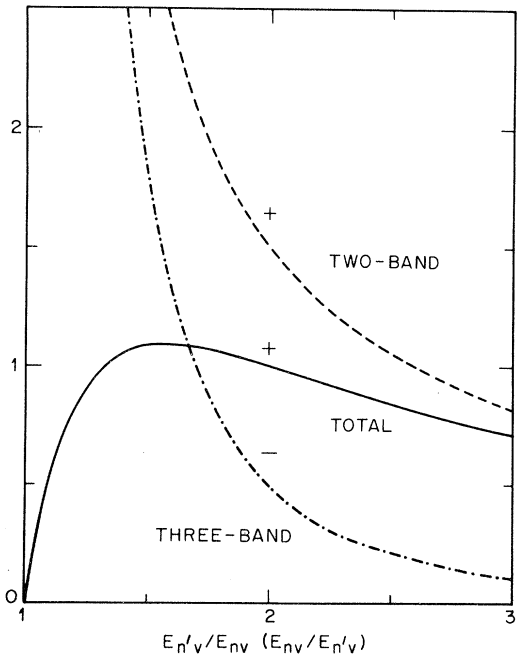

FIG. 1. Relative contributions of the two- and three-band parts of $\chi^{(2)}_{\xi\xi\xi}$ to the total expression in the scalar potential representation.

the contribution to $\chi^{(2)}_{\xi\xi\xi}$ is proportional to the inverse *fifth* power of the smaller gap at any given wave vector $\vec{k}$ of the Brillouin zone. We note that if $E_{n'v} < E_{nv}$, then $E_{n'v}$ and E_{nv} must be interchanged in Eq. (2.21), and the entire expression multiplied by -1. Thus the *magnitude* of $\chi^{(2)}$ is a measure of the band average of the inverse fifth power of the energy gap between valence and lowest conduction bands, and its *sign* depends on the relative average ordering of the two conduction bands n and n'. We note, however, that these relationships are strictly valid only if the momentum matrix elements are independent of $\vec{k}$, which is a good approximation in small-gap regions of the Brillouin zone. In Sec. III, we show that the band n corresponds to the singly degenerate conduction band in crystals of zinc-blende symmetry, and since this is normally the lowest-lying conduction band, $\chi^{(2)}$ is positive in agreement with experiment.

The above discussion assumes a vcc virtual-electron process, for which Eqs. (2.9) and (2.20) apply directly with the condition that v is a filled band and both n and n' represent empty bands. Here, both orderings of intermediate states (n-n' and n'-n) give identical contributions since both matrix element product and energy terms change sign in Eq. (2.9) with interchange of n and n'. For cvv virtual-hole processes where v and $n = v'$ are filled and n' is empty, then the sum over both valence bands for a given ordering results in the exactly the same expressions, but with a minus sign:

$$\chi^{(2)}_{\xi\xi\xi} = -\frac{3ie^3\hbar^3}{2m^3V} \sum_{n,v,v,\vec{k}}{}' \frac{P^\xi_{nv}P^\xi_{vv'}P^\xi_{v'n}}{E^4_{nv}E^4_{nv'}} \times E_{v'v}(2E^2_{nv} + 3E_{nv}E_{nv'} + 2E_{nv'}) \quad (2.22a)$$

$$\cong -\frac{3ie^3\hbar}{2m^3V} \sum_{n,v,v,\vec{k}}{}' \frac{P^\xi_{nv}P^\xi_{vv'}P^\xi_{v'n}}{E^5_{nv'}}, \quad 1.25 < \frac{E_{nv}}{E_{nv'}} < 2.3 \quad (2.22b)$$

where E_{nv} is the larger gap, for each conduction band n. Thus, the above discussion concerning signs, representations of the electric field, and two- and three-band contributions, also applies to virtual-hole processes.

We note that Eq. (2.9) can also be obtained directly from the third-order correction to the energy, given by stationary-state perturbation theory[30]:

$$E^{(3)} = \frac{e^3}{\hbar^2 V} \mathcal{E}^\xi\mathcal{E}^\xi\mathcal{E}^\xi \sum_{n',n,v,\vec{k}} \Big(\langle v\vec{k}|x^\xi|n\vec{k}\rangle \frac{1}{E_{nv}} \langle n\vec{k}|x^\xi|n'\vec{k}\rangle \frac{1}{E_{n'v}} \langle n'\vec{k}|x^\xi|v\vec{k}\rangle - \delta_{n'n}\langle v\vec{k}|x^\xi|n\vec{k}\rangle \frac{1}{E^2_{nv}} \langle n\vec{k}|x^\xi|v\vec{k}\rangle\langle v\vec{k}|x^\xi|v\vec{k}\rangle \Big)$$

$$= -\tfrac{1}{3}\chi^{(2)}_{\xi\xi\xi}\mathcal{E}^\xi\mathcal{E}^\xi\mathcal{E}^\xi, \quad (2.23)$$

where $n \neq v$ and $n' \neq v$. The matrix elements are evaluated with the help of Eqs. (2.15), using the convention that the operators ∇_k from intraband matrix elements operate rightwards only. Thus, the second term vanishes identically in summation over the Brillouin zone. By using the identity

$$\sum_{\vec{k}} \nabla_{\vec{k}} \left(\frac{P^\xi_{vn}}{E_{nv}} \frac{P^\xi_{nv}}{E_{nv}} \frac{1}{E^2_{nv}} \right) \equiv 0$$

$$= \sum_{\vec{k}} \frac{1}{E^2_{nv}} \left[\frac{P^\xi_{vn}}{E_{nv}} \nabla_{\vec{k}} \left(\frac{P^\xi_{nv}}{E_{nv}} \right) + \left(\nabla_{\vec{k}} \frac{P^\xi_{vn}}{E_{nv}} \right) \frac{P^\xi_{nv}}{E_{nv}} \right] - 2\sum_{\vec{k}} \frac{P^\xi_{vn}P^\xi_{nv}}{E^5_{nv}} (\nabla_{\vec{k}} E_{nv}) \quad (2.24)$$

the two-band ($n' = n$) contribution from the first term is seen to be equal to the two-band contribution given in Eq. (2.17). Since the three-band ($n' \neq n$) term in Eq. (2.23) is identical to Eq. (2.20), the result follows. This method is suggested for obtaining the band formulation of $\chi^{(3)}$.

III. APPLICATION: CRYSTALS OF ZINC-BLENDE SYMMETRY

Equations (2.9), (2.21), and (2.22) represent the main result of Sec. II, simplified to describe crystals of zinc-blende symmetry. In this section, we apply them to various crystals of zinc-blende symmetry, calculating sp^3-band contributions to $\chi^{(2)}_{123}$ in a constant energy-gap model,[1,8,11,17,18] and also calculating the relative contributions from the regions Γ and Λ in the Brillouin zone by means of approximate analytic expressions for the energy in these high-symmetry regions.[20–22] The calculated sign is found to be in agreement with experiment, but the theory consistently underestimates the observed values of $\chi^{(2)}_{123}$. Possible reasons for this are discussed in this section and also in Sec. IV.

The tetrahedral bonds of zinc-blende crystals are formed primarily from the sp^3 orbitals of the constituent atoms, and give rise to energy bands which transform at $\vec{k} = 0$ as Γ_1 (valence v'), Γ_{15} (upper valence v), Γ_1 (conduction c), and Γ_{15} (upper conduction c'), in the usual order of increasing energy.[31] Flytzanis and Ducuing[14] have summarized the reasons why $\chi^{(2)}_{123}$ should be related mainly to these orbitals in the bond picture. Similar conclusions are obtained in the band picture by noting that remaining bands tend to be relatively far removed in energy from those generated from these orbitals. The three-band form of Eqs. (2.9) and (2.21) represents no limitation since the bonding and antibonding combinations of sp^3 orbitals always results in more than three bands in any crystal. The results discussed here are not expected to apply to crystals where d bands hybridize substantially with the sp^3 bands.

Since detailed band-structure calculations are not available over the entire Brillouin zone, approximations must be made. We use the standard assumption that the momentum matrix elements are constant over the entire Brillouin zone and equal to their value at Γ. Bell[11] has discussed the validity of this assumption and has pointed out that it is equivalent to the constant matrix element approximation in the theory of the linear dielectric constant. These matrix elements are calculated by constructing the zinc-blende wave functions from the basis wave functions of a fictitious homopolar semiconductor by means of a perturbing antisymmetric potential; the solution is straightforward and worked out in detail elsewhere.[11,32,33] The results can be expressed in terms of the zinc-blende matrix elements

$$\langle v|p|c\rangle = P_{vc} = +iAP, \tag{3.1a}$$

$$\langle v|p|c'\rangle = P_{vc'} = -iQ, \tag{3.1b}$$

$$\langle c|p|c'\rangle = P_{cc'} = +iAPV/W_c, \tag{3.1c}$$

where P and Q are the momentum matrix elements between the $\Gamma_{25'}$ valence band and the Γ_1 and Γ_{15} conduction bands, respectively, of the fictitious homopolar semiconductor; V is the matrix element of the antisymmetric potential between the $\Gamma_{25'}$ valence and Γ_{15} conduction bands, and if E is the $\Gamma_{25'} - \Gamma_{15}$ energy gap, then

$$W_c = \tfrac{1}{2}\{E + [E^2 + (2V)^2]^{1/2}\}, \tag{3.2a}$$

$$A = [1 + (V/W_c)^2]^{-1/2}. \tag{3.2b}$$

All quantities P, Q, and V are positive in standard coordinates [anion at $\frac{1}{8}a_0(1, 1, 1)$]. Mixing of the Γ_1 valence and $\Gamma_{2'}$ conduction band in the homopolar semiconductor is assumed negligible. Following Bell, we assume that P, Q, and E depend mainly on the lattice constant a_0 and are given for all materials by means of a suitable interpolation from their values in Si, Ge, and α-Sn:

$$P \cong (7.15a_B/a_0)\hbar/a_B, \tag{3.3a}$$

$$Q \cong (0.365 + 1.65a_B/a_0)\hbar/a_B, \tag{3.3b}$$

$$E \cong (9.02 - 0.55a_0/a_B)\ \text{eV}, \tag{3.3c}$$

where a_B is the Bohr radius. V is determined from Eq. (3.3c), and from the observed value of the E_0' transition energy of a particular crystal by means of the equation

$$V = +[(E_0')^2 - E^2]^{1/2}. \tag{3.4}$$

In the constant matrix element approximation, only three-band combinations of $\Gamma_{15} - \Gamma_{15} - \Gamma_1$ symmetry will contribute to $\chi^{(2)}_{123}$, which restricts attention to the virtual-electron process vcc′ and the virtual-hole process c′v′v. By Eqs. (3.1) and (2.9), either will give a positive contribution to $\chi^{(2)}_{123}$ only if the Γ_1 band lies on the average between the Γ_{15} bands. Therefore, the c′v′v virtual-hole process always results in a negative contribution, but since typically $E_{v'v} > E_{cv}$ the normally positive vcc′ virtual-electron process dominates by over an order of magnitude (except possibly for SiC, where $E_{cc'}$ is probably small[34]). We consider explicitly only the virtual-electron process vcc′. In this approximation, note that the momentum matrix elements are completely determined by experimental measurements.

The simplest approximation to the interband energies is to replace their detailed variation with constant average or effective gaps[1,8,11,17,18] defined so that the linear dielectric constant $\epsilon_1(0)$ is unchanged. For most zinc-blende crystals, the lowest conduction band is that arising from the Γ_1 state at $\vec{k} = 0$. In this crude model it is justifiable to neglect the remaining conduction bands and define the constant energy gap $\overline{E}$ by

$$\epsilon_1(0) - 1 = \frac{4\pi e^2\hbar^2}{m^2 V} \sum_{v,c,\vec{k}} \frac{2P^t_{vc}P^t_{cv}}{E_{cv}(\vec{k})^3} \tag{3.5a}$$

$$\cong \frac{5}{2}\,\frac{64\pi e^2\hbar^2 P^2}{3m^2 a_0^3 \overline{E}^3}. \tag{3.5b}$$

In Eq. (3.5b) we have used

$$P^t = (1/\sqrt{3})(P^x + P^y + P^z), \tag{3.6}$$

$$(1/V)\textstyle\sum_k 1 = 8/a_0^3. \tag{3.7}$$

The factor of $\frac{5}{2}$ arises from the assumption that the lowest valence band Γ_1 contributes nothing, and that the lowest of the three valence bands of Γ_{15} symmetry at $\vec{k} = 0$ is sufficiently lower in energy so its contribution is half of each of the other two. The factor of 2 for spin degeneracy has been included in Eq. (3.7). Values of a_0/a_B, $\epsilon_1(0)$, $\overline{E}$, and the average gap E_g calculated by Van Vechten[35] are given in Table I for each of the nine zinc-blende crystals for which $\chi^{(2)}_{123}$ is known. $\overline{E}$ and E_g do not differ by more than a few percent for any crystal in Table I (this is also true for the remaining VI, III-V, and II-VI materials listed by Van Vechten; the poorest agreement occurs for the four II-VI compounds listed in Table I, mainly because Van Vechten explicitly takes the d-shell contribution into account while we do not.) Average gaps can be defined in several different but physically reasonable ways[35–38]; our point here is that the average gap $\overline{E}$ obtained via momentum matrix elements, the k-space volume, and the restricted use of the sp^3 orbital bands agrees well with the average gap E_g, and therefore to interpret $\overline{E}$ as a dominant average gap appears reasonable within this model.

A straightforward calculation using Eq. (2.21) with the preceding model yields

$$\chi^{(2)}_{123} = +\frac{20e^3\hbar^3}{m^3 a_0^3 \overline{E}^5} A^2P^2Q\frac{V}{W_c} \tag{3.8a}$$

TABLE I. Comparison of the experimental and theoretical values for $\chi^{(2)}_{123}$ calculated in the average band model. The energy $\bar{E}$ is the average energy obtained from $\epsilon_1(0)$. E_{eff} is the average energy obtained from the experimental values of $\chi^{(2)}_{123}$.

Crystal	a_0/a_B[a]	ϵ_0[a]	E_0 (eV)	E_g (eV)[a]	$\bar{E}$ (eV)	E_{eff} (eV)	$\chi^{(2)}_{123}$ (expt)[b] (10^{-8} esu)	$\chi^{(2)}_{123}$ (calc) (10^{-8} esu)
GaAs	10.684	10.9	4.44[c]	5.19	5.00	4.20	90	38
InSb	12.242	15.7	3.16[c]	3.74	3.49	2.79	330	110
GaP	10.300	9.1	4.78[c]	5.74	5.68	4.88	52	24
GaSb	11.561	14.4	3.27[c]	4.08	3.96	2.92	300	70
InAs	11.406	12.3	4.44[c]	4.58	4.29	3.41	200	64
ZnSe	10.710	5.9	7.8[d]	7.06	6.30	5.84	22	15
CdTe	12.246	7.2	5.3[c]	5.38	4.66	3.91	80	33
ZnS	10.222	5.2	8.4[e]	7.84	7.17	6.46	17	10
ZnTe	11.510	7.3	7.8[d]	5.76	5.14	4.21	73	29

[a]Reference 35.

[b]Data tabulated in Ref. 18 [see note added in proof at end of Sec. IV].

[c]M. Cardona, K. L. Shaklee, and F. H. Pollak, Phys. Rev. 154, 696 (1967).

[d]J. P. Walter, M. L. Cohen, Y. Petroff, and M. Balkanski, Phys. Rev. B 1, 2661 (1970).

[e]J. W. Baars, in *II-VI Semiconducting Compounds*, edited by D. G. Thomas (Benjamin, New York, 1967), p. 631.

$$= +\frac{3e\hbar}{2m\bar{E}^2}\chi^{(1)}_{tt}A^2Q\frac{V}{W_c}. \qquad (3.8b)$$

The same valence-band contribution assumed in Eq. (3.5) has been used in Eqs. (3.8). In standard coordinates $V>0$, so $\chi^{(2)}_{123}>0$. Values of $\chi^{(2)}_{123}$ calculated within this model using the values of E_0' listed in Table I are compared with experimental results on Table I. It is evident that the concept of an average gap cannot be extended within this model to calculate $\chi^{(2)}_{123}$. Obviously, the reason is due in part to the fact that an energy value which accurately represents (by its definition) the band average of $[E_{cv}(\vec{k})]^{-3}$ might not be a good representation of the band average of $[E_{cv}(\vec{k})]^{-5}$ since k-space regions having smaller interband separation should enter with greater weight in the latter average. This has been demonstrated[22] for $\chi^{(3)}$, where the effective average is $[E_{cv}(\vec{k})]^{-6}$. As an indication that this is probably correct (but may not be complete) we treat $\bar{E}=E_{eff}$ as an adjustable parameter to be calculated from Eq. (3.8a) using the experimentally determined value of $\chi^{(2)}_{123}$. The average energies E_{eff} so obtained are also listed in Table I. They follow closely the averages $\bar{E}$ but lie about 0.8 eV lower, as expected from the difference between the averages of smaller and larger inverse power. We point out that E_{eff} does not correlate with any particular feature of the optical spectrum (although in the II-VI compounds, it is quite nearly equal to the energy $E_1+\Delta_1$) and that to make any such identification would probably obscure the physics of the process. The true energy average is that defined by Eq. (2.9) and not $[E_{cv}(\vec{k})]^{-5}$; for very small band gap materials (InSb, InAs, GaSb) this function behaves more like $[E_{cv}(\vec{k})]^{-4}$, and thus the very strong enhancement of the Γ point contribution encountered in $\chi^{(3)}$ is not present to as great an extent here.

By making analytic approximations to the interband energy,[20-22] the contributions from the high-symmetry regions at Γ and Λ in the Brillouin zone can be treated more accurately to investigate possible enhancement of $\chi^{(2)}_{123}$ from the low-energy regions of the Brillouin zone. Only GaAs will be considered since its intermediate-valued band gaps are representative of most zinc-blende materials, and these gaps are large enough to allow the use of Eq. (2.21) which does not involve parameters (energy gaps and masses) of higher conduction bands. We have obtained very similar results for InAs, InSb, GaSb, GaP, and ZnTe (and GaAs) using the more exact form given by Eq. (2.9) when known or estimated values[11] of higher interband parameters are used.

Using the selection rules appropriate to the Γ region of the Brillouin zone, Eq. (2.21) can be written

$$\chi^{(2)}_{123\Gamma} = (0.0888\ \text{eV}^5)P^2Q\left(\frac{W_cV}{W_c^2+V^2}\right)a_B^3 \times\int_0^{K_\Gamma} k_\rho^2\,dk_\rho\,E_{cv}^{-5}(k_\rho), \qquad (3.9)$$

where P and Q are in atomic units ($a_B/\hbar$) and K_Γ is a cutoff taken to be one-fourth of the distance from Γ to X. This cutoff restricts the above integral to about 1.6% of the total volume of the Brillouin zone. Each valence-band contribution is to be calculated separately, and the three contributions from the light hole (lh), heavy hole (hh), and spin-orbit (so) interaction are considered only to the extent of modifying the interband energy gaps and reduced masses at Γ in the parabolic approximation

$$E_{cv}(k) = E_g + \hbar^2k_\rho^2/2\mu_{cv}. \qquad (3.10)$$

Values used in the computation, as well as the total

contribution to $\chi^{(2)}_{123}$ from this region, are given in Table II for GaAs. For comparison, we also give the total contribution to $\epsilon_1(0)$ calculated in this model.

A similar calculation can also be performed for the eight regions lying between Γ and L. Making use of selection rules and summing over intermediate conduction bands, we find for each of the two upper valence bands along Λ the contribution

$$\chi^{(2)}_{123_\Lambda} = (1.21\ \text{eV}^5)P^2Q\left(\frac{W_c V}{W_c^2+V^2}\right)\left(\frac{a_B}{a_0}\right)a_B^2 \times \int_0^{K_L} k_\rho\, dk_\rho\, E_{cv}^{-5}(k_\rho), \quad (3.11)$$

where P and Q are in atomic units, k_ρ is the component of $\vec{k}$ measured perpendicular to the $\langle 111\rangle$ axis (longitudinal symmetry is assumed), and k_L is a cutoff vector taken to be half the distance from L to K. The cutoff in the $\langle 111\rangle$ direction is taken as $\frac{3}{4}$ of the distance from Γ to L. Also shown in Table II are the relative contributions from the Λ directions evaluated in the simple parabolic approximation given by Eq. (3.10). The regions Γ and Λ together comprise less than 40% of the Brillouin zone, but contribute 64% of the total value of $\chi^{(2)}_{123}$ calculated in the average band approximation, suggesting low band-gap enhancement is significant.

A more representative estimate is obtained if one uses a nonparabolic approximation to the energy. One such approximation,

$$E_{cv}(k) = E_g(1+\hbar^2k_\rho^2/2\mu z E_g)^z \cong E_g + \hbar^2k_\rho^2/2\mu + O(k_\rho^4), \quad (3.12)$$

is a generalization of the two-band model[32] (for which $z=\frac{1}{2}$). This is the simplest nonparabolic form still retaining the parabolic variation for small k_ρ. The exponent z is an adjustable parameter; Fig. 2 shows that by choosing the proper value of z it is possible to achieve a very good fit to the actual energy bands over a much larger region in the Brillouin zone, in this case to the calculated variation in the Δ direction of GaAs obtained from $\vec{k}\cdot\vec{p}$ calculations.[31] Although nonparabolic refinements should have little effect on resonance behavior, as has been demonstrated in calculations of electroreflectance spectra,[39] it should have a substantial influence on average parameters such as long-wavelength susceptibilities. This is

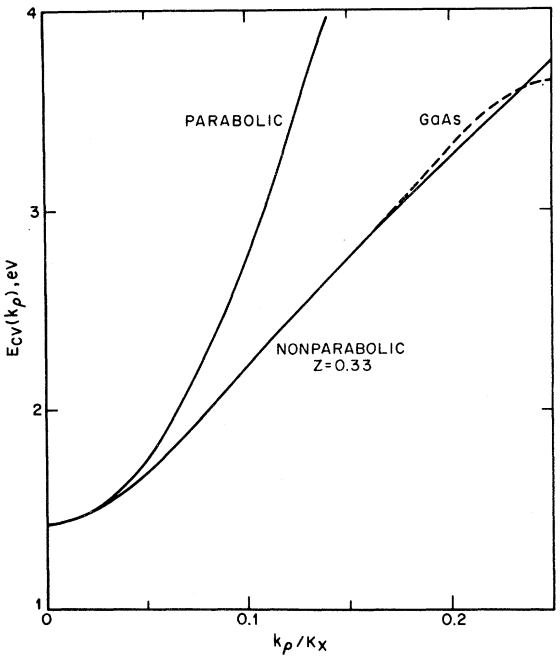

FIG. 2. Fit of the nonparabolic analytic expression defined by Eq. (3.12) to the actual interband energy variation of the light-hole–conduction-band pair in the Δ direction for GaAs calculated in Ref. 31. The abscissa is the fractional distance from Γ to X.

TABLE II. Contributions to $\epsilon_1(0)$ and $\chi^{(2)}_{123}$ from high-symmetry regions in GaAs for various analytic models of the interband energy. The results shown here are typical of the zinc-blende compounds listed in Table I.

		Γ, 1.6%			Γ		Λ, 38%	Λ	$\Lambda+\Gamma$
		lh	hh	so	Total	E_1	$E_1+\Delta_1$	Total	Total
Parabolic	μ_{cv}/m_e	0.035[a]	0.058[a]	0.046[a]	...	0.086[b]	0.086[b]	...	...
	E_g/eV	1.43[c]	1.43[c]	1.77[c]	...	2.93[c]	3.16[c]	...	...
	$\epsilon_1(0)$	0.11	0.20	0.11	0.42	1.9	1.6	3.5	3.9
	$\chi^{(2)}_{123}(10^{-8}$ esu)	1.7	3.5	1.2	6.4	9.8	7.4	17.2	24
Non-parabolic	z	0.33	0.40	0.50	...	0.3	0.3	...	...
	$\epsilon_1(0)$	0.33	0.41	0.18	0.92	2.7	2.3	5.0	5.9
	$\chi^{(2)}_{123}(10^{-8}$ esu)	4.8	6.9	1.9	13.6	15	11	26	40
Coulomb	$\epsilon_1(0)$	0.13	0.27	0.13	0.53	...	...	...	...
	$\chi^{(2)}_{123}(10^{-8}$ esu)	2.4	5.3	1.7	9.4	...	...	...	...

[a]Reference 11. [b]Reference 31. [c]Reference c of Table I.

supported by the calculated contributions to $\chi^{(2)}_{123}$ and $\epsilon_1(0)$ shown in Table II, obtained with Eq. (3.12) where z has been chosen by a best fit for each valence band. This more realistic approximation increases the calculated contribution to $\chi^{(2)}_{123}$ by nearly a factor of 2.

A final correction which should be included is the band-edge oscillator strength enhancement by the electron-hole Coulomb interaction. This correction is necessary (and sufficient) to explain *quantitatively* the observed absorption coefficient in GaAs,[40] Ge,[41] and GaP,[42] and the observed low-field electroreflectance spectra of Ge,[41] at the fundamental absorption edge. This correction cannot be explicitly included in the bond charge calculations but must be represented as a local field factor, unlike the band theory where explicit treatment is possible.

Coulomb effects are most easily treated by modifying Eqs. (2.9) and (2.21) into a density-of-states structure by replacing the k-space integration with

$$\int d^3k\, f(E(\vec{k})) \rightarrow \int n(E)\, dE\,. \tag{3.13}$$

For direct transitions we find

$$n(E) = \frac{\pi m^2 E^2}{e^2\hbar^2 |\hat{\epsilon}\cdot\vec{P}_{cv}|^2}\,\epsilon_2(E) \tag{3.14}$$

in terms of the imaginary part of the dielectric constant $\epsilon_2(E)$ and the interband matrix element $\vec{P}_{cv}$. From the known expressions for $\epsilon_2(E)$ in the simple parabolic and continuum exciton approximations

$$n(E) = 2\pi(2\mu/\hbar^2)^{3/2}(E-E_g)^{1/2}\,, \qquad E > E_g \text{ (simple parabolic)} \tag{3.15a}$$

$$n(E) = \frac{4\pi^2}{13.6\text{ eV}}\left(\frac{\mu}{m}\right)^2 \frac{1}{\epsilon_1(0)}\, a_B^{-3} \times\left\{1-\exp\left[-\frac{2\pi}{\epsilon_0}\left(\frac{(13.6\text{ eV})\mu}{m_e(E-E_g)}\right)^{1/2}\right]\right\}^{-1}\,, \qquad E > E_g \text{ (continuum exciton)}. \tag{3.15b}$$

The discrete exciton contribution has been omitted in Eq. (3.15b); it can be included approximately by extending E_g down to the lowest exciton bound-state energy. The use of the continuum exciton density of states introduces a screening factor $\epsilon_1(0)^{-1}$ into the calculation. Nonparabolicity effects are not included in Eqs. (3.15).

The values of $\chi^{(2)}_{123}$ and $\epsilon_1(0)$ calculated from Eqs. (2.21), (3.13), and (3.15b) in the vicinity of the Γ point in GaAs are also shown in Table II. It is evident that the Coulomb interaction has a much stronger effect on $\chi^{(2)}_{123}$ than on $\epsilon_1(0)$, which may be an indication of the true origin of local-field corrections in the bond picture. The 50% increase in $\chi^{(2)}_{123}$ calculated for GaAs is typical; we note in fact that if generally applicable and combined with nonparabolicity effects, the resulting value for $\chi^{(2)}_{123}$ would be in range of experimental uncertainty. We emphasize, however, that a detailed calculation from an accurate band structure would be required to be able to satisfactorily draw this conclusion, and that our treatment of the Coulomb interaction by means of a density-of-states argument neglects possible excited-state interactions and is therefore phenomenological.

IV. DISCUSSION

On the basis of approximate model band structures for crystals of zinc-blende symmetry, it was shown in Sec. III that the one-electron approximation to $\chi^{(2)}_{123}$ in the energy-band formalism consistently underestimated experimental values, but that a detailed consideration of nonparabolicity and electron-hole Coulomb interaction effects reduced the discrepancy. Previously proposed approximate expressions have generally obtained much better agreement without these refinements. It is, therefore, of interest to investigate these expressions in detail.

The approximate expression proposed by Bell[11] is

$$\chi^{(2)}_{123} = \frac{3e\hbar}{m}\,\chi^{(1)}\,\frac{QV}{(E_0')^3}\,, \tag{4.1}$$

which gives reasonably good agreement with experiment. However, if Eq. (4.1) was obtained from the interband part of the scalar potential representation as stated, then a factor of $E_{10}/E_0' \sim \frac{1}{2}$ is missing and the sign is wrong; if in fact it was obtained from Eq. (2) of Bell's paper (our starting point) then a factor of $\frac{1}{2}$ is again missing. In addition to this factor, better agreement is also obtained because the optical transition E_0' is assigned the function of the average energy instead of the somewhat larger energies $\bar{E}$ and E_g. We note, however, that these discrepancies do not affect Bell's dispersion results, which were calculated from the general finite-frequency expressions. The result of Phillips and Van Vechten (PVV),[17] later extended by Kleinman,[18] contains elements of both bonding orbital and band theory, and its literal interpretation in terms of energy band structure is not clear. However, the use of the *intrastate* matrix element $\langle a|\bar{x}|a\rangle$ is analogous to the *intraband* contribution of the scalar potential representation. The intermediate conduction band Γ_1 does not appear explicitly but its function is assumed in the ground-state mixing term [e.g., as in Eq. (2.21), where the state Γ_{15} appears only in matrix elements], and it appears as the ratio C/E_g. PVV's result differs from Bell's result by substituting C for V, E_g for E_0', and multiplying by

$$0.46\left(\frac{a_0}{10a_B}\right)\left(\frac{E_g}{\text{eV}}\right)\left(\frac{E_h}{E_g}\right)^4, \tag{4.2}$$

and thus has a somewhat more complicated dependence and is generally larger. Kleinman's modification,[18] which makes use of the f-sum rule, can be expressed by

$$\chi_{123}=\frac{3e\hbar}{m}\chi^{(1)}\frac{QC}{E_hE_g^2}\left(\frac{E_g}{13.6\text{ eV}}\right)^{1/2}. \tag{4.3}$$

All expressions are similar, differing by factors of order unity. The bonding orbital theories[14] are the real-space complements of the band picture and are related only in the average sense.

Spin-orbit (so) splitting has been included only to the extent that it affects various energy gaps and interband reduced masses. However, in the presence of so splitting, it is possible to get an additional contribution to $\chi^{(2)}_{123}$ from the so-split valence bands and the lowest conduction band.[9,10,18] Rustagi[10] has calculated its effect using a two-band model and found it to be substantially smaller than the measured value. Cardona and Pollak[21] have performed a rough calculation to show that it should account for about 10% of the three-band effect considered here. In further support of this result, Aspnes[43] has estimated the contribution of the so interaction to $\chi^{(2)}$ in the vicinity of the E_1 transitions for GaAs from the linear electroreflectance data of Kyser and Rehn[44]; the value obtained amounts to about 10% of the calculated band contribution at these points, in good agreement with the estimate of Cardona and Pollak.

Extension of the theory of crystals of wurtzite symmetry was not attempted here because the zinc-blende results indicated that the simple model band structures used were not adequate. Although our calculations indicate that the discrepancy between theory and experiment is reduced as the band structure is made more accurate and Coulomb effects are included, we emphasize that a detailed calculation with an accurately computed energy-band structure would be required in order to determine satisfactorily whether these corrections are adequate.

Note added in proof. Since this paper was submitted, several of the experimental values of $\chi^{(2)}_{123}$ given in Table I have been revised downward substantially [B. F. Levine and C. G. Bethea, Appl. Phys. Letters 20, 275 (1972)]. For GaAs, $10^8\chi^{(2)}_{123}(\text{esu})=43$ should replace 90; for GaP, 20 should replace 52; for ZnS, 8 should replace 17 (values for the remaining compounds are not given). Thus the predictions of 38, 24, and 10, of the average band model for GaAs, GaP, and ZnS, respectively, are now in very good agreement with experiment. However, this model is crude and this agreement should not preclude more rigorous comparisons of the theory by evaluating Eq. (2.9) with accurately calculated energy-band structures.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge many useful and informative conversations with D. A. Kleinman, who stimulated the interest in this work, and to M. I. Bell for sending a preprint of his work prior to publication. The author also wishes to thank M. I. Bell, E. I. Blount, G. D. Boyd, M. Cardona, B. F. Levine, and J. E. Rowe for discussions in regard to various aspects of this work.

[1]J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Phys. Rev. 127, 1918 (1962).
[2]N. Bloembergen and Y. R. Shen, Phys. Rev. 133, A37 (1964).
[3]J. F. Ward, Rev. Mod. Phys. 37, 1 (1965).
[4]F. N. H. Robinson, Bell System Tech. J. 46, 913 (1967).
[5]S. S. Jha and N. Bloembergen, Phys. Rev. 171, 891 (1968); W. K. Burns and N. Bloembergen, Phys. Rev. B 4, 3437 (1971).
[6]R. Loudon, Proc. Phys. Soc. (London) 80, 952 (1962).
[7]P. N. Butcher and T. P. McLean, Proc. Phys. Soc. (London) 81, 219 (1963); 83, 579 (1964).
[8]P. L. Kelley, J. Phys. Chem. Solids 24, 607 (1963); 24, 1113 (1963).
[9]A. G. Aronov and G. E. Pikus, Fiz. Tverd. Tela 10, 825 (1968) [Sov. Phys. Solid State 10, 648 (1968)].
[10]K. Rustagi, J. Phys. Chem. Solids 30, 2547 (1969).
[11]M. I. Bell, in *Electronic Density of States*, edited by L. H. Bennett, Natl. Bur. Std. (U. S.) Spec. Publ. No. 323 (U. S. GPO, Washington, D. C., 1971), p. 757.
[12]H. Pursey, P. A. Page, and M. J. P. Musgrove, J. Phys. C 2, 1085 (1969).
[13]V. S. Bagaev, T. Ya. Belousova, Yu. N. Berozashvili, and D. S. Lordkipanidze, Fiz. i Tekh. Poluprovodnikov. 3, 1687 (1969) [Sov. Phys. Semicond. 3, 1418 (1970)].
[14]Chr. Flytzanis and J. Ducuing, Phys. Letters 26A, 315 (1968); Phys. Rev. 178, 1218 (1969); C. L. Tang and Chr. Flytzanis, Phys. Rev. B 4, 2520 (1971).
[15]B. F. Levine, Phys. Rev. Letters 22, 787 (1969); 24, 440 (1970).
[16]C. R. Jeggo and G. D. Boyd, J. Appl. Phys. 41, 2741 (1970).
[17]J. C. Phillips and J. A. Van Vechten, Phys. Rev. 183, 709 (1969).
[18]D. A. Kleinman, Phys. Rev. B 3, 3139 (1970).
[19]R. C. Miller and W. A. Nordland, Opt. Commun. 1, 400 (1970); Phys. Rev. B 3, 4896 (1970).
[20]C. W. Higginbotham, M. Cardona, and F. H. Pollak, Phys. Rev. 184, 821 (1969).
[21]M. Cardona and F. H. Pollak, in *The Physics of Opto-Electronic Materials*, edited by W. A. Alberg, Jr. (Plenum, New York, 1971), p. 81.
[22]J. A. Van Vechten, M. Cardona, D. E. Aspnes, and R. M. Martin, in *Proceedings of the Tenth International Conference on Semiconductors*, AEC Division of Technical Information Publication No. CONF-700801 (U. S. AEC, Oak Ridge, Tenn., 1970), p. 82.
[23]C. Kittel, *Quantum Theory of Solids* (Wiley, New

York, 1966), p. 183.

[24]D. E. Aspnes, Bull. Am. Phys. Soc. 16, 46 (1971).

[25]An expression illustrating the dispersion characteristics of $\chi^{(2)}_{123}(\omega_1, \omega_2)$ for small ω_1, ω_2 has been given previously by R. K. Chang, J. Ducuing, and N. Bloembergen [Phys. Rev. Letters 15, 415 (1965)], but the coefficients are represented as parameters only and cannot be evaluated for a specific energy-band model.

[26]See, for instance, N. Bloembergen, *Nonlinear Optics* (Benjamin, New York, 1965); P. N. Butcher, Engineering Experiment Station, Bulletin No. 200 (Ohio State University, 1965) (unpublished), where gauge invariance is discussed as a general property of the Hamiltonian.

[27]E. I. Blount, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1962), Vol. 13, p. 305.

[28]F. Stern, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1963), Vol. 15, p. 299.

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

[30]A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1962), Vol. II. p. 712ff, and see particularly Eq. (XVI. 74), p. 717.

[31]F. H. Pollak, C. W. Higginbotham, and M. Cardona, J. Phys. Soc. Japan Suppl. 21, 20 (1966).

[32]F. Herman, J. Electron. 1, 103 (1955).

[33]M. Cardona, J. Phys. Chem. Solids 24, 1543 (1963).

[34]H. -G. Junginger and W. van Haeringen, Phys. Status Solidi 37, 709 (1970).

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

[36]J. C. Phillips, Phys. Rev. Letters 20, 550 (1968); *Covalent Bonding in Molecules and Solids* (University of Chicago Press, Chicago, 1969).

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

[38]P. Y. Yu and M. Cardona, Phys. Rev. B 2, 3193 (1970).

[39]D. E. Aspnes and J. E. Rowe, in Ref. 22, p. 422.

[40]M. D. Sturge, Phys. Rev. 127, 768 (1962).

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

[42]D. D. Sell and P. V. Lawaetz, Phys. Rev. Letters 26, 311 (1971).

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

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