Optical rectification and shift currents in GaAs and GaP response: Below and above the band gap

F. Nastos and J. E. Sipe
Department of Physics and Institute for Optical Sciences, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7

(Received 31 March 2005; revised manuscript received 16 May 2006; published 7 July 2006)

We present a full band structure scheme to calculate the electronic contribution to the second order susceptibility coefficient pertinent to optical rectification, $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ where $\omega_\Sigma \approx 0$, within the independent particle approximation for the electron dynamics, and in the dipole limit. This allows us to determine the electronic response of a bulk semiconductor to a femtosecond optical pulse over a range of central frequencies, both below and above the band gap frequency. Particularly interesting is the limit $\chi_2^{abc}(0;\omega,-\omega)$. In addition to the usual near-dc interband rectification current, shift and injection currents, associated with actual divergences in $\chi_2^{abc}(0;\omega,-\omega)$, are taken into account. Calculations for GaAs and GaP, in which injection currents are forbidden, are performed. The band energies and matrix elements are computed with the full potential linearized augmented plane wave method. For frequencies above the band gap, and for typically available pulse widths, we demonstrate that the shift current dominates the current response, being approximately two orders of magnitude larger than the rectification current. For very narrow pulse widths, on the order of a femtosecond, the rectification current becomes comparable to the shift current.

DOI: 10.1103/PhysRevB.74.035201 PACS number(s): 77.84.−s, 42.65.−k, 42.70.Mp

I. INTRODUCTION

Optical rectification (OR) is a second-order nonlinear optical effect in which an applied monochromatic optical field induces a static polarization inside a noncentrosymmetric semiconductor.[1] If the optical field comes instead from a pulsed laser source, then the induced polarization varies in time with the pulse intensity, and generates a fast current. Although OR was discovered early in the study of nonlinear optics,[2] there has been renewed interest in it since the development of femtosecond laser pulses first allowed the generation of ultrafast OR currents that emit THz radiation.[3] These sources are useful for spectroscopy in the far-infrared regime, a part of the spectrum that has not been extensively exploited.[4] Since the THz radiation is related to the charge dynamics of the electrons in the semiconductor, it is also of interest as a means of examining the microscopic processes responsible for its generation. Interactions of the laser field with the ionic lattice of a polar material can also generate THz radiation,[5] but in this work we consider exclusively the electronic response.

For applied laser fields with photon energies below the band gap, the rectification current is the only second-order nonlinear current induced in the system. For optical excitations above the band gap, there is a markedly different response.[6] Absorption processes lead to short-lived currents, whose time-variation adds to the THz generated from the interband polarization. In a clean semiconductor with filled bands and lacking inversion symmetry, there are two sources of these currents: *shift* and *injection* processes.[7] The differences between these currents and optical rectification have not been fully appreciated. See the review by von Baltz[8] for a brief survey.

Shift currents arise when there is a charge transfer associated with excitation from the valence band to the conduction band. The shift is on the order of a bondlength, and occurs on femtosecond time scales. Figure 1 approximates a "before-and-after" look at the electron density in GaAs, and helps give a qualitative feel for how shift currents arise. In panel (a) one sees the electron density, in a [110] lattice plane, for the states at the top of the valence band Γ_8 point; these are fourfold degenerate. The density is localized around the arsenic atom. The crystal can absorb photons of energy larger than the band gap to populate the states near the bottom of the conduction band. Panel (b) shows the density of the Γ_6-point wave function of the lowest conduction band, where the electron has now relocated nearer to the gallium atom. Depending on the details of the light polarization, the electron density evolves differently to reach the state depicted in (b). If the electric field is polarized along the

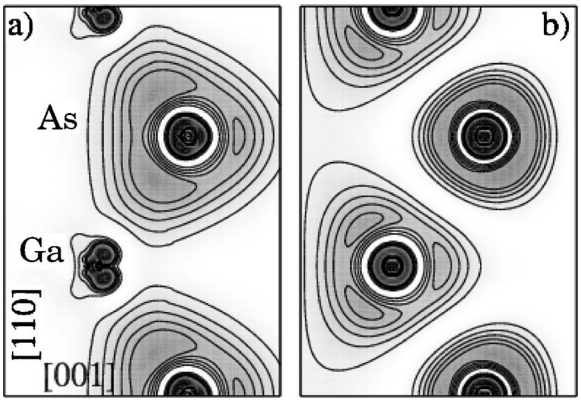

FIG. 1. A plot of the electron density in the [110] plane of GaAs. Panel (a) shows the electron density of the highest Γ valence electron, and panel (b) the electron density of the lowest Γ point conduction band. Dark regions correspond to higher densities. The densities in panel (b) have been multiplied by two since the lowest conduction band Γ point is twofold degenerate whereas the highest valence band Γ point is fourfold degenerate.

 ©2006 The American Physical Society

[100] direction, for example, an electron from the As atom can move towards any one of its four nearest neighbors with equal probability, giving no net current. But for polarizations along the [111] directions, associated with each arsenic atom is one particular gallium atom in a direction from the arsenic atom indicated by the direction of the field. It is primarily towards this gallium atom that the electron associated with the arsenic atom moves, and a net electrical current results.

Shift current is also known as the *bulk photovoltaic effect*.[9] In a perturbative density-matrix calculation, shift currents are associated with the nonzero off-diagonal components.[10] It has recently been proposed that shift currents could be observed in different systems, such as boron-nitride nanotubes.[11] A similar shift current occurs in finite systems, such as molecules, in which there is a charge transfer associated with the electronic excitation.[12]

Injection currents reflect a lack of symmetry in the crystal structure that can give rise to phase differences between transition amplitudes associated with different polarizations of light. When the crystal is photoexcited with circularly polarized light, the different excitation pathways for horizontal and vertical polarizations lead to an interference effect resulting in an asymmetric population in reciprocal space, and hence a current. Because it is generated only with circularly polarized light, the injection current is also called the *circular photovoltaic effect*.[9] The injection current, which is symmetry forbidden in zincblende crystals, has not been thoroughly studied. We do not examine injection currents in detail here, although our analytical results pertain to calculations of the effect.

In this work we address the difference between below- and above-gap excitation in the generation of the electric current in the zincblende crystals, and carefully distinguish between the rectification and shift contributions. There have been few theoretical attempts to describe the difference between below- and above-gap excitation in unbiased samples. Early descriptions of the shift current introduced the shift distance,[10] an estimate of the shift in the center of an electron charge induced by absorption, which replaced the explicit evaluation of dipole matrix elements. Analytic models for the band structure were used, which limited results to near the band edge. Later, full-band-structure expressions for the shift current, using the minimal coupling Hamiltonian, were presented and evaluated using model band structures.[13] The nonresonant polarization current has been difficult to calculate for nonzero frequencies, even within the simplest approximations. To our knowledge, Khurgin was the first to attempt to investigate the difference in below- and above-gap rectification in GaAs within a single model.[14,15] Based on permanent dipole moments, his model is limited to excitations near the band gap and uses an empirically determined value for the bond dipole moment. More generally, the validity of this model has been questioned.[16,17] Attempts to include scattering effects on the evolution of the current through kinetic theory have been reported.[18]

The difference in the response between below- and above-gap excitation has recently been experimentally time resolved at the femtosecond scale,[19] although it was noticed much earlier for longer time scales.[20] In this paper we approach these effects within a first-principles nonlinear susceptibility formalism.[21] Specific nonlinear effects are associated with different frequency components of $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, and a complete account of the response would include effects at all $\omega_\Sigma=\omega_\beta+\omega_\gamma$. Our focus is on nonlinear mixing of the frequency components of the field around ω and $-\omega$, so that the sum frequency ω_Σ is near zero.

Different approaches to a microscopic calculation of $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ appear in the literature. Typically, they are limited to the specific frequency components of the different nonlinear effects. Popular examples are second-harmonic generation[22–25] $\chi_2^{abc}(-2\omega;\omega,\omega)$, the electro-optic effect[24,26] $\chi_2^{abc}(-\omega;0,\omega)$, and below-gap rectification[6] $\chi_2^{abc}(0;\omega,-\omega)$. In approaches that directly evaluate $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ it is insufficient to know one specific frequency component if the response to a short pulse is to be evaluated, since the spectrum of the pulse cannot be approximated by a single frequency. To include the dispersion one must be able to evaluate $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ over a range of frequencies ω_β and ω_γ. In this paper we will show how this can be accomplished.

Recently, Sipe and Shkrebtii[21] (SS) presented a general formulation for calculating $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ within the independent particle approximation for the electron dynamics. The approach is a perturbative one, accounting for both interband and intraband transitions. Moreover, the formulation applies to any frequency components of $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, irrespective of whether the applied frequency is below or above the band gap. Although the formalism is within the independent particle approximation, it has been successful in studies of second-harmonic generation.[27] Here we apply the formalism of SS to the rectification current and shift current density in bulk GaAs. An outline of the paper is as follows. In Sec. II we present our main analytic result, a simplification of the general second-order susceptibility tensor for the calculation of these processes. The simplification highlights the different contributions to the current: rectification current occurring at all frequencies, and shift and injection currents when there is absorption from frequencies above the band gap. We specialize to the zincblende crystals for which the expressions reduce to a simple form. The resulting expressions involve integrations over the Brillouin zone of the crystal, and apply to all crystal classes. In Sec. III we focus on GaAs and GaP, and we use our approach to numerically evaluate the expressions. We employ the full potential linearized augmented plane wave (FLAPW) method to estimate the band energies and matrix elements, and the band gap underestimation is treated within a scissors correction. Our results are presented in Sec. IV. We first consider the continuous wave limit, including the shift distance as well as the below-gap polarizability. The temporal response to a short optical pulse is calculated, and the effect of nonlinear mixing on the current generation is discussed. We conclude in Sec. V.

II. SECOND ORDER COEFFICIENTS

A. General treatment

In this section we derive our expressions for calculating the low frequency second-order response. We use "Gaussian

centimeter-gram-second (CGS) units" throughout our derivations for convenience, but we switch to Systemé International (SI) units to present our results in the figures. While our starting point is the general expression for the second-order current density as given by Sipe and Shkrebtii, which we sketch here,[21] those authors ultimately considered only monochromatic radiation in their examples, and it is necessary to extend their formalism to deal efficiently with incident pulses. Unless otherwise stated, we use the same notation as in SS.

The current density for sum frequencies $\omega_\Sigma \approx 0$ can be written as the sum

$$\langle \mathbf{J}(t)\rangle = \langle \mathbf{J}_{\text{intra}}(t)\rangle + \langle \mathbf{J}_{\text{inter}}(t)\rangle,$$

where

$$\langle \mathbf{J}_{\text{inter}}(t)\rangle = \frac{d}{dt}\langle \mathbf{P}_{\text{inter}}(t)\rangle,$$

and

$$\langle \mathbf{J}_{\text{intra}}(t)\rangle = \langle \mathbf{J}_{\text{intra}}(t)\rangle^{\text{I}} + \langle \mathbf{J}_{\text{intra}}(t)\rangle^{\text{II}}. \tag{1}$$

This separation into different intraband terms helps identify the shift and injection contributions to the total current, which are contained in $\langle \mathbf{J}_{\text{intra}}(t)\rangle^{\text{I}}$ and $\langle \mathbf{J}_{\text{intra}}(t)\rangle^{\text{II}}$, respectively. In terms of the Fourier components of the electric field, given by

$$E^a(\omega) = \int_{-\infty}^{\infty} dt\, E^a(t) e^{i\omega t},$$

the total second-order polarization is

$$P^a(t) = \int_{-\infty}^{\infty} \frac{d\omega_\beta}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\gamma}{2\pi} \chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \times E^b(\omega_\beta) E^c(\omega_\gamma) e^{-i\omega_\Sigma t}, \tag{2}$$

where $\omega_\Sigma = \omega_\beta + \omega_\gamma$. Here, superscript Roman characters represent Cartesian coordinates, and when they are repeated they are to be summed over. The intraband electric currents discussed in the Introduction appear as singularities of the susceptibility coefficient in the limit $\omega_\Sigma \to 0$.[21]

They arise in the following way in this treatment. It is found that the current response can be described with the aid of three third-rank tensors that are finite for all values of their frequency arguments: $\chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, $\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, and $\bar{K}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$. The first of these relates the interband polarization term $\langle \mathbf{P}_{\text{inter}}(t)\rangle$ to the applied field,

$$\langle P^a_{\text{inter}}(t)\rangle = \int_{-\infty}^{\infty} \frac{d\omega_\beta}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\gamma}{2\pi} \chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \times E^b(\omega_\beta) E^c(\omega_\gamma) e^{-i\omega_\Sigma t}, \tag{3}$$

the second relates the first intraband current term of Eq. (1) to the applied field,

$$\langle J^a_{\text{intra}}(t)\rangle^{\text{I}} = \int_{-\infty}^{\infty} \frac{d\omega_\beta}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\gamma}{2\pi} \bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \times E^b(\omega_\beta) E^c(\omega_\gamma) e^{-i\omega_\Sigma t}, \tag{4}$$

and the third relates the time derivative of the second intraband current term of Eq. (1) to the applied field,

$$\frac{d}{dt}\langle J^a_{\text{intra}}(t)\rangle^{\text{II}} = \int_{-\infty}^{\infty} \frac{d\omega_\beta}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\gamma}{2\pi} \bar{K}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \times E^b(\omega_\beta) E^c(\omega_\gamma) e^{-i\omega_\Sigma t}. \tag{5}$$

Since the total expectation value of the second-order current response is

$$\langle \mathbf{J}(t)\rangle = \frac{d\mathbf{P}(t)}{dt} \tag{6}$$

we see from Eq. (2) that

$$\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = \frac{\bar{K}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)}{(-i\omega_\Sigma)^2} + \frac{\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)}{(-i\omega_\Sigma)} + \chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \tag{7}$$

explicitly showing the divergences as $\omega_\Sigma \to 0$.

Despite these divergences, we can of course still consider the limiting case of monochromatic irradiation where

$$\mathbf{E}(\omega_\beta) = 2\pi[\mathbf{E}_0 \delta(\omega_\beta - \omega) + \mathbf{E}_0^* \delta(\omega_\beta + \omega)], \tag{8}$$

or

$$\mathbf{E}(t) = \mathbf{E}_0 e^{-i\omega t} + \text{c.c.}$$

Returning to Eqs. (3)–(5), it is clear that $\bar{K}_2^{abc}(0;\omega,-\omega)$ describes a current appearing in the crystal at a constant rate, since using Eq. (8) in Eq. (5) yields

$$\frac{d}{dt}\langle J^a_{\text{intra}}(t)\rangle^{\text{II}} = \bar{K}_2^{abc}(0;\omega,-\omega) E_0^b (E_0^c)^* + \text{c.c.}$$

This is the injection current. Similarly, we see that $\bar{\sigma}_2^{abc}(0;\omega,-\omega)$ is associated with a constant shift current accompanying the absorption of radiation, since using Eq. (8) in Eq. (4) yields

$$\langle J^a_{\text{intra}}(t)\rangle^{\text{I}} = \bar{\sigma}_2^{abc}(0;\omega,-\omega) E_0^b (E_0^c)^* + \text{c.c.}$$

As well, a constant polarization appearing in the crystal is described by $\chi^{abc}_{2\text{inter}}(0;\omega,-\omega)$, since using Eq. (8) in Eq. (3) gives

$$\langle P^a_{\text{inter}}(t)\rangle = \chi^{abc}_{2\text{inter}}(0;\omega,-\omega) E_0^b (E_0^c)^* + \text{c.c.} \tag{9}$$

One might be tempted to identify this last term as the optical rectification. This would be incorrect, since it would exclude intraband contributions to the near-static polarization coming from Eqs. (4) and (5): Although $\bar{\sigma}_2^{abc}(0;\omega,-\omega)$ and $\bar{K}_2^{abc}(0;\omega,-\omega)$ strictly vanish for $\hbar\omega$ below the band gap, an expansion of the susceptibilities $\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,-\omega_\gamma)$ and $\bar{K}_2^{abc}(-\omega_\Sigma;\omega_\beta,-\omega_\gamma)$ about $\omega_\Sigma = 0$ leads to terms linear in ω_Σ that survive at all frequencies. When these are used in Eq. (7), contributions to a static polarization that are well be-

haved as $\omega_\Sigma \to 0$ arise. These add with $\chi^{abc}_{2\text{inter}}(0;\omega,-\omega)$ to form a tensor we call $\chi^{abc}_{2\text{rect}}(0;\omega,-\omega)$, which describes the total constant polarization,

$$P^a_{\text{rect}}(t) = \chi^{abc}_{2\text{rect}}(0;\omega,-\omega)E^b_0(E^c_0)^* + \text{c.c.}, \tag{10}$$

induced in the crystal at frequencies below the gap. It is Eq. (10), not Eq. (9), that describes the full optical rectification for continuous wave excitation. This shows that, even for continuous wave excitation, a direct identification of the terms $\bar{K}^{abc}_2(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, $\bar{\sigma}^{abc}_2(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, and $\chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ with injection, shift, and rectification current, respectively, is incorrect. For pulsed excitation, the situation is even more complicated and care is required to identify the different components of the current. The work of SS focused on excitation above the band and the continuous-wave limit, so this issue did not arise there. For extremely short pulses any such kind of division of the current into different components probably loses all significance, but for long pulses in the optical regime, one can still phenomenologically identify an injection, shift, and rectification current by the time dependence of their source terms. The identification is more easily made if the symmetry properties of the response tensors $\bar{K}^{abc}_2(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, $\bar{\sigma}^{abc}_2(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, and $\chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ are identified, which we do in the next section through a microscopic analysis. Once this is done, we return to a description of pulsed excitation.

B. Response coefficients

Microscopic expression for the second-order nonlinear response tensors $\bar{K}^{abc}_2(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, $\bar{\sigma}^{abc}_2(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, and $\chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ have been derived by SS within the independent particle model. The interband polarization coefficient is written as $\chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)=\chi^{abc}_{2\text{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)+\tilde{\chi}^{abc}_{2\text{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$. The term $\chi^{abc}_{2\text{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ survives if only interband transitions are considered, and $\tilde{\chi}^{abc}_{2\text{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ contains both interband and intraband contributions. These two coefficients are given by

$$\chi^{abc}_{2\text{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = \frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}\left[\left(\frac{r^a_{nm}r^b_{mp}r^c_{pn}}{\omega_{mp}-\rho_\beta\omega_{mn}}(\rho_\beta f_{pm}F_+(\omega_{mp},\omega_\beta)+\rho_\gamma f_{np}F_+(\omega_{pn},\omega_\gamma)-f_{nm}F_+(\omega_{mn},\omega_\Sigma))\right)+(bc\beta\gamma\leftrightarrow cb\gamma\beta)\right], \tag{11}$$

and

$$\tilde{\chi}^{abc}_{2\text{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = \frac{ie^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m}f_{nm}\left\{r^a_{nm}\left[\left(\frac{r^b_{mn;c}}{\rho_\gamma\omega_{mn}}-\frac{r^b_{mn}\Delta^c_{mn}}{\rho^2_\gamma\omega^2_{mn}}\right)+\left(\frac{r^c_{mn;b}}{\rho_\beta\omega_{mn}}-\frac{r^c_{mn}\Delta^b_{mn}}{\rho^2_\beta\omega^2_{mn}}\right)\right]F_+(\omega_{mn},\omega_\Sigma)\right.$$
$$\left.+\left[\left(r^b_{mn}\left[\frac{\rho_\beta r^a_{nm;c}}{\rho_\gamma\omega_{mn}}+\frac{\rho^2_\beta r^a_{nm}\Delta^c_{mn}}{\rho^2_\gamma\omega^2_{mn}}\right]F_+(\omega_{mn},\omega_\beta)\right)+(bc\beta\gamma\leftrightarrow cb\gamma\beta)\right]\right\}. \tag{12}$$

In these equations, and ones that follow, $\rho_i=\omega_i/\omega_\Sigma$ for $i=\alpha$ or β, $\omega_{mn}(\mathbf{k})=\omega_m(\mathbf{k})-\omega_n(\mathbf{k})$ is the frequency difference between bands m and n, and we have dropped the explicit $\mathbf{k}$ dependence of matrix elements and band frequencies. The symbol $F_\pm$ denotes

$$F_\pm(\omega_{nm},\omega_\beta)=\frac{1}{\omega_{nm}-\omega_\beta-i\eta}\pm\frac{1}{\omega_{nm}+\omega_\beta+i\eta},$$

and $(bc\beta\gamma\leftrightarrow cb\gamma\beta)$ indicates that the preceding term should be repeated switching b with c, and β with γ. The small positive real number η is taken to the 0^+ limit at the end of the derivation. For nondegenerate points in the Brillouin zone (BZ), the interband position matrix elements $r^a_{nm}(\mathbf{k})$ are related to the velocity matrix elements, $v^a_{nm}(\mathbf{k})$, by

$$r^a_{nm}(\mathbf{k}) = -iv^a_{nm}(\mathbf{k})/\omega_{nm}(\mathbf{k}), \tag{13}$$

where

$$v^a_{nm}(\mathbf{k})\delta(\mathbf{k}-\mathbf{k}') = \frac{\hbar}{mi}\int d^3r\langle n\mathbf{k}|\mathbf{r}\rangle\frac{\partial}{\partial r^a}\langle\mathbf{r}|m\mathbf{k}\rangle,$$

and $\langle\mathbf{r}|m\mathbf{k}\rangle$ is the Bloch state.[28] The quantity $\Delta^a_{nm}(\mathbf{k})=(v^a_{nn}(\mathbf{k})-v^a_{mm}(\mathbf{k}))/m$ represents the difference between the group velocities in different bands at a given k point. We also require the generalized derivative of $r^a_{nm}(\mathbf{k})$ with respect to k^b, given by

$$r^a_{nm;b}(\mathbf{k}) = -\frac{r^a_{nm}(\mathbf{k})\Delta^b_{nm}(\mathbf{k})+r^b_{nm}(\mathbf{k})\Delta^a_{nm}(\mathbf{k})}{\omega_{nm}(\mathbf{k})}-\frac{i}{\omega_{nm}(\mathbf{k})}\times\sum_p\left[\omega_{np}r^a_{np}(\mathbf{k})r^b_{pm}(\mathbf{k})-\omega_{pm}r^b_{np}(\mathbf{k})r^a_{pm}(\mathbf{k})\right], \tag{14}$$

which arises from the intraband transitions that appear when one considers the nonlinear response.[29] It is easy to verify that $[r^a_{mn;b}(-\mathbf{k})]^*=-r^a_{nm;b}(\mathbf{k})$, which helps simplify our com-

putations. In Eq. (14) the index p runs over all bands. When the energy band p is degenerate with another band s (and s is either m or n) so that $\omega_{ps}(\mathbf{k})=0$, the prescription detailed in SS shows that matrix element $\mathbf{r}_{ps}(\mathbf{k})$ vanishes.

In terms of these microscopic quantities $\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ is given by (Ref. 21),

$$\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = -\frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3} \times\sum_{n,m} f_{nm}\{r^b_{mn}r^c_{nm;a}F_-(\omega_{mn},\omega_\beta) + r^c_{mn}r^b_{nm;a}F_-(\omega_{mn},\omega_\gamma)\}, \quad (15)$$

and $\bar{K}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ is given by (Ref. 30),

$$\bar{K}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = -\frac{ie^3}{8\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m} f_{nm}\Delta^a_{mn}[r^c_{nm},r^b_{mn}] \times[F_+(\omega_{mn},\omega_\beta)-F_+(\omega_{mn},\omega_\gamma)],$$

where $[r^c_{nm},r^b_{mn}]\equiv r^c_{nm}r^b_{mn}-r^b_{nm}r^c_{mn}$.

With these expressions in hand, a direct calculation of the current response using Eqs. (6) and (7) would be possible. But it would involve much computational overhead, since at each pair of relevant frequencies the tensors $\chi^{abc}_{\mathrm{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, $\chi^{abc}_{\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, $\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$, and $\bar{K}^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ would have to be evaluated and stored, which is impractical. To simplify our calculations, we analytically rewrite these coefficients in terms of effective single-frequency response coefficients that can be more simply computed.

We begin with the easier response coefficients. Clearly, Eq. (15) separates into single-frequency effective coefficients,

$$\bar{\sigma}_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = -i(\Lambda^{abc}(\omega_\beta)+\Lambda^{acb}(\omega_\gamma)),$$

where $\Lambda^{abc}(\omega)$ can be expressed as the sum: $\Lambda^{abc}(\omega)=\Lambda^{abc}_{\mathrm{I}}(\omega)+i\Lambda^{abc}_{\mathrm{II}}(\omega)$. The term $\Lambda^{abc}_{\mathrm{II}}(\omega)$ is real and explicitly given by

$$\Lambda^{abc}_{\mathrm{II}}(\omega) = -\frac{i\pi e^3}{2\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m} f_{nm}r^b_{mn}r^c_{nm;a}\delta(\omega_{mn}-\omega). \quad (16)$$

Because of the form of $F_-(\omega_{mn},\omega)$, appearing in Eq. (15), the Kramers-Krönig equation relates $\Lambda^{abc}_{\mathrm{I}}(\omega)$ to $\Lambda^{abc}_{\mathrm{II}}(\omega)$,

$$\Lambda^{abc}_{\mathrm{I}}(\omega) = \frac{2\omega}{\pi}\int_0^\infty d\omega'\mathcal{P}\left(\frac{1}{\omega'^2-\omega^2}\right)\Lambda^{abc}_{\mathrm{II}}(\omega'), \quad (17)$$

where $\mathcal{P}$ indicates the principal part. We also have the relations,

$$\Lambda^{abc}_{\mathrm{I}}(-\omega) = -\Lambda^{abc}_{\mathrm{I}}(\omega),$$

and

$$\Lambda^{abc}_{\mathrm{II}}(-\omega) = \Lambda^{abc}_{\mathrm{II}}(\omega).$$

The delta function appearing in the integrand of Eq. (16) restricts $\Lambda^{abc}_{\mathrm{II}}(\omega)$ to be nonzero only for frequencies such that $\hbar|\omega|$ is greater than the band gap.

The susceptibility tensor $\bar{K}^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ for $\langle\mathbf{J}_{\mathrm{intra}}(t)\rangle^{\mathrm{II}}$ can also be separated into single-frequency terms,

$$\bar{K}^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = \Gamma^{abc}(\omega_\beta)-\Gamma^{abc}(\omega_\gamma),$$

where

$$\Gamma^{abc}(\omega) = -\frac{ie^3}{8\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m} f_{nm}[r^c_{nm},r^b_{mn}]F_+(\omega_{mn},\omega).$$

Unlike $\Lambda^{abc}(\omega)$, which contains the factor $F_-(\omega_{mn},\omega)$, $\Gamma^{abc}(\omega)$ contains $F_+(\omega_{mn},\omega)$ in its integrand; thus we can write $\Gamma^{abc}(\omega)=\Gamma^{abc}_{\mathrm{I}}(\omega)+i\Gamma^{abc}_{\mathrm{II}}(\omega)$, where the two purely real terms $\Gamma^{abc}_{\mathrm{I}}(\omega)$ and $\Gamma^{abc}_{\mathrm{II}}(\omega)$ are related by another Kramers-Krönig equation,

$$\Gamma^{abc}_{\mathrm{I}}(\omega) = \frac{2}{\pi}\int_0^\infty d\omega'\mathcal{P}\left(\frac{\omega'}{\omega'^2-\omega^2}\right)\Gamma^{abc}_{\mathrm{II}}(\omega'), \quad (18)$$

and they have the symmetry relations,

$$\Gamma^{abc}_{\mathrm{I}}(-\omega) = \Gamma^{abc}_{\mathrm{I}}(\omega),$$

and

$$\Gamma^{abc}_{\mathrm{II}}(-\omega) = -\Gamma^{abc}_{\mathrm{II}}(\omega).$$

After some algebra, we obtain

$$\Gamma^{abc}_{\mathrm{II}}(\omega) = -\frac{ie^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m} f_{nm}[r^c_{nm},r^b_{mn}]\Delta^a_{mn}\delta(\omega_{mn}-\omega)$$

for the imaginary part. For frequencies ω where $\hbar|\omega|$ is below the band gap $\Gamma^{abc}_{\mathrm{II}}(\omega)=0$, and $\Gamma^{abc}_{\mathrm{I}}(\omega)$ contributes to the nonresonant polarization.

We now turn to the more complicated coefficients describing $\langle\mathbf{P}_{\mathrm{inter}}(t)\rangle$, namely $\chi^{abc}_{2\mathrm{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ and $\tilde{\chi}^{abc}_{2\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$. In Appendix A we show that $\tilde{\chi}^{abc}_{2\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ can be written as the sum,

$$\tilde{\chi}^{abc}_{2\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) = \left\{\frac{-1}{\omega_\gamma}[\Lambda^{abc}(\omega_\beta)]^* + \frac{1}{\omega_\gamma}\Lambda^{abc}(\omega_\Sigma) + \frac{\omega_\beta}{\omega_\gamma}\left[1-\frac{\omega_\Sigma}{\omega_\gamma}\right]W^{abc}(\omega_\beta) + \frac{\omega_\Sigma^2}{\omega_\gamma^2}W^{abc}(\omega_\Sigma)\right\} + (bc\beta\gamma\leftrightarrow cb\gamma\beta). \quad (19)$$

Here $W^{abc}(\omega)$ is another effective single-frequency component. It contains the factor $F_+(\omega_{mn},\omega)$ in its integrand, so like $\Gamma^{abc}(\omega)$ above, the real and imaginary parts of $W^{abc}(\omega)$ are related through a Kramers-Krönig relation equivalent to Eq. (18). To simplify the computations, we note that the imaginary parts of $W^{abc}(\omega)$ and $\Gamma^{abc}(\omega)$ are related by $W^{abc}_{\mathrm{II}}=\omega^{-2}\Gamma^{abc}_{\mathrm{II}}(\omega)$.

So far the analytical rewriting of the susceptibility tensors has been exact. In principle, the above expressions could be

used in similar calculations for any frequency component of $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$. The remaining term $\chi_{2\text{ter}}^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ is not so supple, and to express it in a computable form requires approximations. In Appendix B we show that in the limit $|\rho_\beta| \gg 1$ and $|\rho_\gamma| \gg 1$ we can write,

$$\chi_{2\text{ter}}^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \approx \left\{ \left[1 - \frac{\omega_\beta}{\omega_\gamma} - \frac{\omega_\beta\omega_\Sigma}{\omega_\gamma^2} \right] C^{abc}(\omega_\beta) + \frac{\omega_\Sigma}{\omega_\beta}\left[1 + \frac{\omega_\beta^2}{\omega_\gamma^2} \right] \tilde{C}^{abc}(\omega_\beta) + \frac{\omega_\Sigma}{\omega_\beta} E^{abc}(\omega_\Sigma) + \frac{\omega_\Sigma^2}{\omega_\beta^2} \tilde{E}^{abc}(\omega_\Sigma) \right\} + (bc\beta\gamma \leftrightarrow cb\gamma\beta). \tag{20}$$

Like $\Gamma^{abc}(\omega)$, $S^{abc}(\omega)$, and $W^{abc}(\omega)$, the terms $C^{abc}(\omega)$, $\tilde{C}^{abc}(\omega)$, $E^{abc}(\omega)$, and $\tilde{E}^{abc}(\omega)$, contain $F_+(\omega_{mn},\omega)$ in the integrands, and so their real and imaginary parts are related as they are for $\Gamma^{abc}(\omega)$ in Eq. (18). After simplifying we find

$$C_{\text{II}}^{abc}(\omega) = -\frac{\pi e^3}{4\hbar^2}\int \frac{d^3k}{8\pi^3} \sum_{n,m,p} f_{nm} r^b_{nm} \left(\frac{r^a_{mp} r^c_{pn}}{\omega_{mp}} + \frac{r^c_{mp} r^a_{pn}}{\omega_{np}} \right) \times \delta(\omega_{mn} - \omega),$$

$$\tilde{C}_{\text{II}}^{abc}(\omega) = -\frac{\pi e^3}{4\hbar^2}\int \frac{d^3k}{8\pi^3} \sum_{n,m,p} f_{nm} r^b_{nm} \omega_{mn} \left(\frac{r^a_{mp} r^c_{pn}}{\omega^2_{mp}} - \frac{r^c_{mp} r^a_{pn}}{\omega^2_{np}} \right) \times \delta(\omega_{mn} - \omega),$$

$$E_{\text{II}}^{abc}(\omega) = \frac{\pi e^3}{4\hbar^2}\int \frac{d^3k}{8\pi^3} \sum_{n,m,p} f_{nm} \frac{r^a_{nm}}{\omega_{mn}} [r^b_{mp}, r^c_{pn}] \delta(\omega_{mn} - \omega),$$

and

$$\tilde{E}_{\text{II}}^{abc}(\omega) = \frac{\pi e^3}{4\hbar^2}\int \frac{d^3k}{8\pi^3} \sum_{n,m,p} f_{nm} \frac{r^a_{nm}}{\omega^2_{mn}} (r^b_{mp} r^c_{pn} \omega_{mp} + r^c_{mp} r^b_{pn} \omega_{np}) \times \delta(\omega_{mn} - \omega),$$

for the imaginary parts of these terms.

C. The response to a pulse and simplifications for zincblende

We now return to our expressions for pulsed excitation, and phenomenologically identify the injection, shift, and rectification contributions to the response. We consider pulses in the optical regime described by a carrier wave of frequency $\omega_\circ$, modulated by an envelope function $E_{\text{env}}(t)$, so that

$$\mathbf{E}(t) = \mathbf{E}_{\text{env}}(t) e^{-i\omega_\circ t} + \mathbf{E}^*_{\text{env}}(t) e^{i\omega_\circ t}.$$

For this discussion, we consider an unchirped pulse, so that $\mathbf{E}_{\text{env}}(t)$ is real, but the resulting equations can be easily generalized to a chirped pulse. The Fourier transform of the incident field is then

$$\mathbf{E}(\omega) = \mathbf{E}_{\text{env}}(\omega + \omega_\circ) + \mathbf{E}_{\text{env}}(\omega - \omega_\circ),$$

and using this in Eq. (2) the near-static interband response becomes

$$\langle P^a_{\text{inter}}(t)\rangle = \int_{-\infty}^{\infty} \frac{d\omega_\beta}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\gamma}{2\pi} \chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma) \times [E^b_{\text{env}}(\omega_\beta + \omega_\circ) E^c_{\text{env}}(\omega_\gamma - \omega_\circ) + E^b_{\text{env}}(\omega_\beta - \omega_\circ) E^c_{\text{env}}(\omega_\gamma + \omega_\circ)],$$

where we have dropped the terms containing the factors $E^a(\omega+\omega_\circ)E^b(\omega+\omega_\circ)$ and $E^a(\omega-\omega_\circ)E^b(\omega-\omega_\circ)$ since they correspond to sum-frequency and second-harmonic generation. Translating the integration variables and using permutation symmetry simplifies this to

$$\langle P^a_{\text{inter}}(t)\rangle = 2\int_{-\infty}^{\infty} \frac{d\omega_\beta}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega_\gamma}{2\pi} \chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\circ+\omega_\beta,-\omega_\circ+\omega_\gamma) \times E^b_{\text{env}}(\omega_\beta) E^c_{\text{env}}(\omega_\gamma). \tag{21}$$

The frequency spread of the pulse is narrow enough that the susceptibility tensor can be expanded about the carrier frequency to give

$$\chi^{abc}_{2\text{inter}}(-\omega_\Sigma;\omega_\circ+\omega_\beta,-\omega_\circ+\omega_\gamma) = \chi^{abc}_{2\text{inter}}(0;\omega_\circ,-\omega_\circ) + \omega_\beta \frac{\partial}{\partial\omega} \chi^{abc}_{2\text{inter}}(-\omega;\omega_\circ+\omega,-\omega_\circ)\Big|_{\omega=0} + \omega_\gamma \frac{\partial}{\partial\omega} \chi^{abc}_{2\text{inter}}(-\omega;\omega_\circ,-\omega_\circ+\omega)\Big|_{\omega=0} + \cdots.$$

Using this expansion in Eq. (21), keeping only terms to first order in ω, and simplifying gives

$$\langle P^a_{\text{inter}}(t)\rangle = 2\chi^{abc}_{2\text{inter}}(0;\omega_\circ,-\omega_\circ) E^b_{\text{env}}(t) E^c_{\text{env}}(t) + 2\frac{\partial}{\partial\omega} \text{Im}[\chi^{abc}_{2S}(-\omega;\omega_\circ,-\omega_\circ+\omega)]\Big|_{\omega=0} \times \frac{\partial}{\partial t}[E^b_{\text{env}}(t) E^c_{\text{env}}(t)] + 2\frac{\partial}{\partial\omega} \text{Im}[\chi^{abc}_{2A}(-\omega;\omega_\circ,-\omega_\circ+\omega)]\Big|_{\omega=0} \times \left[E^b_{\text{env}}(t) \frac{\partial}{\partial t} E^c_{\text{env}}(t) - E^c_{\text{env}}(t) \frac{\partial}{\partial t} E^b_{\text{env}}(t) \right]. \tag{22}$$

where $\chi^{abc}_{2S} \equiv (\chi^{abc}_{2\text{inter}} + \chi^{acb}_{2\text{inter}})/2$ and $\chi^{abc}_{2A} \equiv (\chi^{abc}_{2\text{inter}} - \chi^{acb}_{2\text{inter}})/2$. To obtain this final form we again used permutation symmetry and that $\chi^{abc*}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta;\omega_\gamma) = \chi^{abc}_{2\text{inter}}(\omega_\Sigma;-\omega_\beta;-\omega_\gamma)$.

The expansion demonstrates how the dispersion of the susceptibility leads to a more complicated temporal response. The first term follows the pulse intensity profile, and as in the continuous wave case, the tensor $\chi^{abc}_{2\text{inter}}(0;\omega_\circ,-\omega_\circ)$ describes a polarization that is proportional to the instantaneous electric field squared. The second and third terms follow the time derivative of the intensity profile, and arise entirely from the off-diagonal components in $\chi^{xyz}_{2\text{inter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$.

One can apply the same procedure to simplify Eqs. (4) and (5) and express $\langle J^a_{\text{intra}}(t)\rangle^{\text{I}}$ and $\langle J^a_{\text{intra}}(t)\rangle^{\text{II}}$ in terms of the

product $E^b_{\rm env}(t)E^c_{\rm env}(t)$ and its derivatives, but the zincblende symmetry of GaAs and GaP, which we focus on for the rest of our presentation, allows some simplifications. For this crystal class, third rank tensors must be symmetric under the exchange of any pair of indices, so that $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)=\chi_2^{acb}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$. Thus the third term in the expansion for $\langle P^a_{\rm inter}(t)\rangle$ vanishes as does the similar term in an analogous expansion for $\langle J^a_{\rm intra}(t)\rangle^{\rm I}$ and $\langle J^a_{\rm intra}(t)\rangle^{\rm II}$. In fact, since the tensor $\bar{K}^{abc}(\omega)$ governing $\langle J^a_{\rm intra}(t)\rangle^{\rm II}$ is antisymmetric under the exchange of its last two indices, the intraband current $\langle J^a_{\rm intra}(t)\rangle^{\rm II}$, which includes the injection current, vanishes for GaAs and GaP and so we will not consider it any further in this paper; it will be the subject of a future communication. For $\langle J^a_{\rm intra}(t)\rangle^{\rm I}$ we have

$$\langle J^a_{\rm intra}(t)\rangle^{\rm I}=2\bar{\sigma}_2^{abc}(0;\omega_\circ,-\omega_\circ)E^b_{\rm env}(t)E^c_{\rm env}(t)+2\frac{\partial}{\partial\omega}{\rm Im}[\bar{\sigma}_2^{abc}(-\omega;\omega_\circ,-\omega_\circ+\omega)]\Big|_{\omega=0}\times\frac{\partial}{\partial t}[E^b_{\rm env}(t)E^c_{\rm env}(t)].$$

When we add $d\langle P^a_{\rm inter}(t)\rangle/dt$ to $\langle J^a_{\rm intra}(t)\rangle^{\rm I}$ to construct the total current $\langle J^a(t)\rangle$ we see that there will be terms that vary as $E^a_{\rm env}(t)E^b_{\rm env}(t)$, as $\partial[E^a_{\rm env}(t)E^b_{\rm env}(t)]/\partial t$, and such higher order derivatives. We identify the term in $\langle J^a(t)\rangle$ that varies as $E^a_{\rm env}(t)E^b_{\rm env}(t)$ as the shift current,

$$J^a_{\rm shift}(t)=2\bar{\sigma}_2^{abc}(0;\omega_\circ,-\omega_\circ)E^b_{\rm env}(t)E^c_{\rm env}(t)\equiv 2\sigma^{abc}_{\rm shift}(\omega_\circ)E^b_{\rm env}(t)E^c_{\rm env}(t),\tag{23}$$

defining $\sigma^{abc}_{\rm shift}(\omega_\circ)$. Removing this contribution, the remainder would then relate $P^a(t)$ to $E^a_{\rm env}(t)E^b_{\rm env}(t)$, $\partial[E^a_{\rm env}(t)E^b_{\rm env}(t)]/\partial t$, and higher order derivatives. This expression we take as identifying the optical rectification $P^a_{\rm rect}(t)$, the generalization of Eq. (10) to treat pulses.

The zincblende symmetry of GaAs and GaP leads to some final simplifications of our expressions. Symmetry requires that there be only one independent nonzero component in which the indices are $a=x$, $b=y$, and $c=z$, for any permutation of $\{x,y,z\}$. For the shift current, the tensor component $\sigma^{xyz}_{\rm shift}(\omega_\circ)$ can be written in terms of our effective single-frequency tensors as follows:

$$\sigma^{xyz}_{\rm shift}(\omega)=2\Lambda^{xyz}_{\rm II}(\omega).\tag{24}$$

For optical rectification, keeping the expansion of $P^a_{\rm rect}(t)$ to the instantaneous intensity, we have

$$P^a_{\rm rect}(t)=2\chi^{abc}_{\rm 2rect}(\omega_\circ)E^b_{\rm env}(t)E^c_{\rm env}(t)$$

and, in terms of the single-frequency tensors, from Eqs. (19) and (20) we find that

$$\chi^{xyz}_{\rm 2rect}(\omega)=\chi^{xyz}_{\rm 2ter}(0;\omega,-\omega)+\tilde{\chi}^{xyz}_{\rm 2tra}(0;\omega,-\omega)+\frac{\partial\Lambda^{xyz}_{\rm I}(\omega)}{\partial\omega},$$

$$=2\frac{\Lambda^{xyz}_{\rm I}(\omega)}{\omega}+4C^{xyz}_{\rm I}(\omega)+\frac{\partial\Lambda^{xyz}_{\rm I}(\omega)}{\partial\omega}.\tag{25}$$

Here $\chi^{xyz}_{\rm 2rect}$ contains the entire interband polarization response of $\chi^{xyz}_{\rm 2inter}$, and also contains the virtual polarization contributions in $\bar{\sigma}^{xyz}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$. The imaginary parts have vanished since the effective response tensors are symmetric under any permutation of the Cartesian indices. This leaves the rectification tensor purely real, as is expected from considering intrinsic permutation symmetry and zincblende symmetry. At first sight this might seem surprising, since one generally expects response tensors to contain both real and imaginary parts, which are related by Kramers-Krönig type relations. Even in nonlinear response this holds, for example, for the second-harmonic response tensor.[31] There the real part of the tensor is associated with conversion between fundamental and the second-harmonic frequencies, while the imaginary part is associated with the removal of energy from the total electromagnetic field.[21] In linear response, of course, the real part of the susceptibility is similarly associated with a modification of the propagation of the signal through the medium, while the imaginary part is associated with its absorption. In our case there is a similar "pairing" of two effects, but they are associated with the rectification [described by Eq. (25)] and the shift current [described by Eq. (24)]; it is the presence of both of them in χ_2 that allows it to satisfy causality.

This concludes the main analytic result of the paper. Assuming that only small deviations from $\omega_\Sigma=0$ in the sum frequency are required, we have expressed the second-order nonlinear response in terms of effective single-frequency components.

III. COMPUTATIONAL DETAILS

In this section we describe our procedure to evaluate the coefficients, $\Lambda^{abc}(\omega)$, $W^{abc}(\omega)$, $C^{abc}(\omega)$, $\tilde{C}^{abc}(\omega)$, $E^{abc}(\omega)$, and $\tilde{E}^{abc}(\omega)$ for the zincblende semiconductors GaAs and GaP. For each of these single-frequency coefficients, only the imaginary part of the Kramers-Krönig pair is evaluated directly; the other part is found from the corresponding Kramers-Krönig relation. We use the linear analytic tetrahedron method (LATM) to perform the required BZ integrations.[32] For the LATM, the energy eigenvalues and momentum matrix elements are required on a dense tetrahedral grid in the Brillouin zone. We estimate these with Kohn-Sham density functional theory[33,34] using the popular FLAPW code WIEN2K.[35,36]

In the FLAPW method, real space is partitioned into nonoverlapping atomic spheres, collectively known as the "muffin-tin" region, and an interstitial region. In the muffin-tin region a linear combination of Schrödinger equation solutions and their energy derivatives are used to expand the wave functions. In the interstitial region a plane-wave expansion is used, and each plane-wave is augmented by an atomi-

clike function inside the sphere. Spin-orbit effects are included through a second variational step. We refer the reader to Singh[37] for further details and references. Following usual procedures, a self-consistent calculation is first made to generate a converged Kohn-Sham potential.

For the results presented below we used the room temperature lattice constants of 10.683 a_0 for GaAs and 10.300 a_0 for GaP. We chose a muffin-tin radius of $R_{\mathrm{mt}}=2.31\ a_\circ$ around each atom for GaAs, and 2.23 $a_\circ$ for GaP. We found the parameters $R_{\mathrm{mt}}\times K_{\mathrm{max}}=9.0$, and $l_{\mathrm{max}}=12$ to be sufficient for convergence; here K_{max} is the magnitude of the largest vector in the plane-wave expansion, and l_{max} is the maximum partial wave inside the atomic spheres. For the exchange-correlation potential we use the local density approximation (LDA) as parametrized by Perdew and Wang.[38] In our tests using generated gradient approximation (GGA) functionals[39] we found little qualitative difference in the response functions. For the self-consistent calculation 1240 k points were used in the irreducible Brillouin zone (IBZ), as determined by Blöchl *et al.*[40] and the total energy was converged to within 10^{-5} Ry. The resulting self-consistent potential was then used to solve the Kohn-Sham equations on a tetrahedral grid of k points tailored for our BZ integrations.

Many different schemes have been presented for dividing the IBZ into tetrahedra. Some common methods are given by Harrison[41] and Blöchl *et al.*[40] To generate our grid we use an implementation due to MacDonald *et al.*,[42] in which the zincblende IBZ is first divided into three tetrahedra described by the sets of vertices: $\{\Gamma,W,U,X\}$, $\{\Gamma,W,X,L\}$ and $\{\Gamma,W,L,K\}$. Each of these tetrahedra is divided into eight smaller tetrahedra by introducing a new k point halfway along each tetrahedron edge. In turn, each new tetrahedron can be further divided as necessary to produce a mesh dense enough for a converged spectrum. The integrals over each resulting tetrahedron are done analytically by separately linearizing the integrand and the energies over the tetrahedron.[43] Using this approach it is easy to keep track of contributions to the spectrum arising from different sectors of the IBZ, and then to refine the mesh as necessary to help with convergence, but the initial specification and division of the IBZ into tetrahedra must be done by hand for each crystal class.

As the number of k points is increased, we find that convergence of the GaAs spectra is slow around the band edge, which has been noted by others calculating different optical responses for this crystal.[22,24] Even when we calculate the response functions with 18 513 k points in the irreducible zone we find that near the band edge the spectra are still not completely converged. We attribute this slow convergence to the small effective mass of the LDA conduction bands, and to overcome it we perform a two-step calculation. First, we calculate the spectra using our 18 513 k point mesh. We then repeat the calculation but use 18 513 points in a region localized around the Γ point. At a photon energy of 2.5 eV, we then match the spectrum accurate for high energies to the valid one at low energies. To match the values requires scaling down the high energy spectrum by about 1%. We note that although 18 513 points may sound like an excessive number to use, all the calculations are performed on a single modern workstation. We anticipate that for many crystals, where the bands are more linear than is typical for zincblende crystals, the response calculations will require fewer k points for convergence. We emphasize accurate integration near the band edge both because we make extensive use of Kramers-Krönig relations, and because we want to establish accurate FLAPW-LDA results for comparisons to other bandstructure methods.

In our calculations for the optical response functions we use 32 conduction bands for both GaAs and GaP. Using 20 conduction bands instead produces very similar response functions. It is well known that the LDA underestimates the conduction band energies. A popular fix for this, which we follow in this work, is simply to shift the conduction band energies such that the band gap agrees with the experimental value. The argument for this is that in many cases, including GaAs, the experimental band structure is close to the LDA one, the major difference being a rigid shift in energy. Also, for some systems, the GW wave functions have a significant overlap with the LDA wave functions,[44] indicating that the matrix elements computed *via* the LDA are a reasonable approximation.

The details of our scissors implementation is different from previous work concerned with the same underlying formalism for the response.[21,24,25] We will briefly describe the scheme, but refer to the reader to the literature[45] for more details. First, the LDA velocity matrix elements for each required k point on our mesh were computed from within the WIEN2K package. Using these and the unscissored LDA-KS eigenvalues, we construct the components $r^a_{mn}(\mathbf{k})$ using Eq. (13). If we were calculating the linear response, this would correspond to the accepted scissors implementation, in which there is cancellation of the correction to the momentum matrix elements with the correction to the energies.[24,46] Here though, unlike an earlier approach,[24] when we calculate the generalized derivative $r^a_{mn;b}(\mathbf{k})$ from Eq. (14) we use the unscissored velocity matrix elements and energy eigenvalues. Since the position matrix element should be computed from the unscissored quantities, so should its generalized derivative. The scissors correction is then applied to all the explicitly remaining energies in our expressions. For GaAs and GaP we used scissor shifts of 1.25 eV and 1.7 eV, respectively.

The calculations for the nonlinear response tensors require a sum over intermediate states, labeled by the index p, but for $C^{abc}_{\mathrm{II}}(\omega)$ and $\tilde{C}^{abc}_{\mathrm{II}}(\omega)$ there is an energy denominator in the integrand associated with this index that potentially introduces numerical difficulties. If the band p is degenerate with the band m, then the basis can be chosen so that the relevant component of position matrix element $r^a_{mp}(\mathbf{k})$ is zero.[21] But if the bands are not degenerate and have a small energy difference, then this denominator can introduce error into the integrand, since the wave functions are not as converged as the energies. This type of issue appears for other nonlinear response calculations, such as the second-harmonic response.[47]

To determine that no such error is occurring, we simply set up a tolerance for what the band separation must be to qualify for a degeneracy, and set the integrand to zero at k points and band combinations which qualify as a degeneracy.

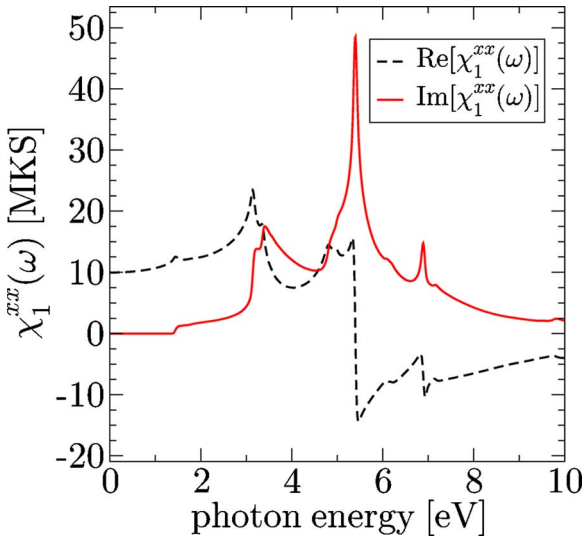

FIG. 2. (Color online) The real and imaginary parts of $\chi_1^{xx}(\omega)$ for GaAs.

We found the spectra in question to be relatively unvarying for tolerances between 10 and 50 meV. For smaller tolerances, different choices of tolerance lead to qualitatively different spectra, and for larger tolerances the overall magnitude of the spectra begin to significantly diminish, indicating that we are excluding too much of the BZ in the integration. For the figures, and results, we settled on a tolerance of 30 meV. An alternate procedure is to lift the degeneracies by hand. Doing this gives nearly identical results.

The imaginary parts of the spectra all go to zero sufficiently fast enough that evaluating the imaginary parts to 20 eV with a spacing of 1 meV was sufficient to evaluate the Kramers-Krönig relations. Before evaluating the Kramers-Krönig pairs, the imaginary parts of our spectra are smeared by convolving them with a 25 meV full width at half maximum (FWHM) Gaussian, as is usually done to account phenomenologically for population decay and dephasing. We chose 25 meV since it roughly corresponds to room temperature broadening of the energy bands.

To establish a reference to other optical response calculations, we first present the real and imaginary parts of the linear response function $\chi_1^{xx}(\omega)$ for GaAs in Fig. 2. The usual independent particle microscopic expression for $\chi_1^{xx}(\omega)$ is used; see, for example, Eq. (34) in SS.

Our results for the real and imaginary parts of the effective single-frequency response coefficients $\Lambda^{xyz}(\omega)$, $C^{xyz}(\omega)$, $\tilde{C}^{xyz}(\omega)$, and $\tilde{E}^{xyz}(\omega)$ are presented in Fig. 3 for GaAs. We do not show the effective single-frequency tensors for GaP since those spectra are very similar to the ones for GaAs. This is expected given the similar electronic structures of the two materials. Since we use such a fine mesh for integration over the Brillouin zone, resonances associated with critical points are quite pronounced, especially in the nonlinear response.

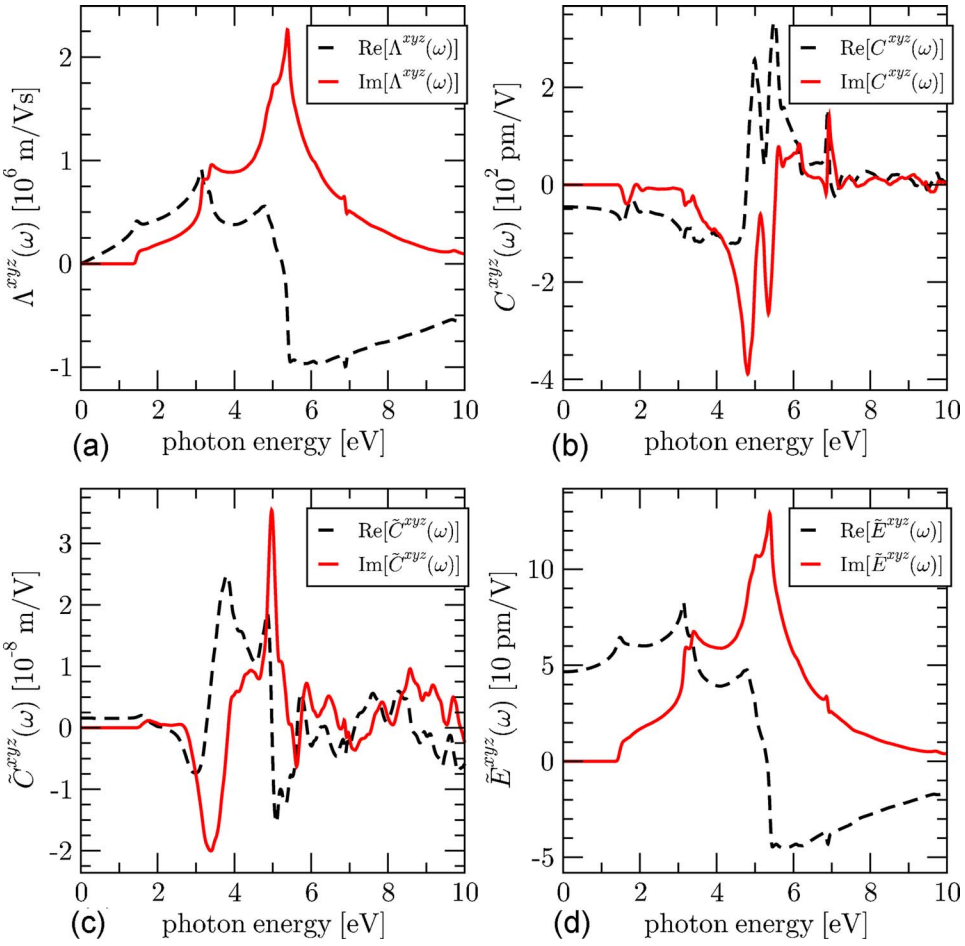

FIG. 3. (Color online) The effective single-frequency tensors required to form the optical rectification tensors in GaAs. Panels (a), (b), (c), and (d) show the real and imaginary parts of $\Lambda^{xyz}(\omega)$, $C^{xyz}(\omega)$, $\tilde{C}^{xyz}(\omega)$, and $\tilde{E}^{xyz}(\omega)$, respectively.

Without the Gaussian smearing we find that the nonlinear response spectra have a noticeable resonance at approximately 7 eV. These resonances are narrow enough that the smearing significantly reduces them.

The $\Lambda^{xyz}(\omega)$ spectrum is linear with ω below the band gap, as expected from Eq. (17). The response functions $C^{xyz}(\omega)$ and $\tilde{C}^{xyz}(\omega)$ show much resonance structure. This is not surprising, given the denominator in their integrands. Similarly strong resonance structures are found for other calculations for the second-order response.[47,48] The spike in the $C^{xyz}(\omega)$ spectrum for GaP is due to the small heavy-hole and light-hole splitting in the valence bands. With a coarser k mesh the resonances are not as pronounced, a usual consequence of finite-element methods such as the LATM. The effect of these resonances on the spectra is minimal below the band gap. We performed calculations without the spin-orbit interaction, and found that the general structure of the spectra were similar.

The shape of the imaginary parts of the $\tilde{E}^{xyz}(\omega)$ and $\Lambda^{xyz}(\omega)$ spectra appears to be very similar. This is because any terms in the second-order optical response functions explicitly containing the velocity matrix element differences $\mathbf{\Delta}_{mn}(\mathbf{k})$ are zero by crystal symmetry for the zincblende crystals.[25] The generalized derivative $r^a_{nm;b}(\mathbf{k})$ in $\Lambda^{abc}_{\mathrm{II}}(\omega)$ contains such a term, and so only the second term of Eq. (14) contributes. One then sees then that the resulting integrands are similar, differing by a factor of $1/(2\omega_{mn})$. But the evaluation of these two components under the scissors approximation is still different, since all the frequencies appearing in the integrand of $E^{abc}(\omega)$ should be scissors corrected, whereas the frequencies in the generalized derivative, appearing in the $\Lambda^{abc}(\omega)$ integrand, are the unmodified Kohn-Sham ones.

Care must be taken in the calculation of $\partial\Lambda_{\mathrm{I}}(\omega)/\partial\omega$ to evaluate Eq. (25). Artificial kinks in the frequency dependence of $\Lambda_{\mathrm{I}}(\omega)$, which are too small to be noticed in the plots, arise at energies corresponding to the tetrahedral microzone edges because of the piece-wise linearization of the band structure over the BZ. Hence a straightforward numerical differentiation of the calculated $\Lambda_{\mathrm{I}}(\omega)$ to yield $\partial\Lambda_{\mathrm{I}}(\omega)/\partial\omega$ would produce unphysical oscillations. To avoid these, we piece-wise spline fit the spectrum $\Lambda_{\mathrm{I}}(\omega)$ on a less dense energy grid before we take the derivative.

IV. RESULTS

A. Continuous wave limit shift current and shift distance

We first investigate the continuous wave limit of the shift current, where $\omega_\beta = -\omega_\gamma$. We use this opportunity to indicate that in Fig. 2 of SS, the spectrum is incorrect due to a computational error. The correct spectrum is given by $2\epsilon_\circ\Lambda^{xyz}_{\mathrm{II}}(\omega)$, with $\Lambda^{xyz}_{\mathrm{II}}(\omega)$ as presented here in SI units.

From J^a_{shift} [see Eq. (23)], we can predict the average distance d_{shift} that an electron center of mass shifts upon absorption of a photon. Since in this limit the shift current is proportional to the absorption rate $\dot{n}$, we have $d_{\mathrm{shift}} = |J^a_{\mathrm{shift}}/(e\dot{n})|$, where $\dot{n}$ is the carrier absorption rate. For zincblende systems and linearly polarized light along the [111] direction, this "shift distance" is given by $d_{\mathrm{shift}}(\omega) = 8\pi\Lambda^{xyz}_{\mathrm{II}}(\omega)/\epsilon_2^{xx}(\omega)$,[21] where $\epsilon_2^{xx}(\omega)$ is the imaginary part of the dielectric function. As mentioned in the Introduction, the Γ-point wave functions at the top of the valence band are localized around the arsenic atoms, and at the bottom of the conduction band they are localized around the gallium atoms. Absorption near the band edge then is expected to lead to a shift distance of about the bond length. In these calculations the GaAs and GaP bond lengths are 2.45 and 2.36 Å, respectively. In Figs. 4 and 5 we plot the frequency dependence of the shift distance for GaAs and GaP, respectively. We see that in GaAs just above the band edge the shift distance is slightly less, but still very close to the bond length. As the frequency increases, it quickly approaches and exceeds the bond length. There is a small dip just below 2 eV associated with the split-off band. Moving higher in energy,

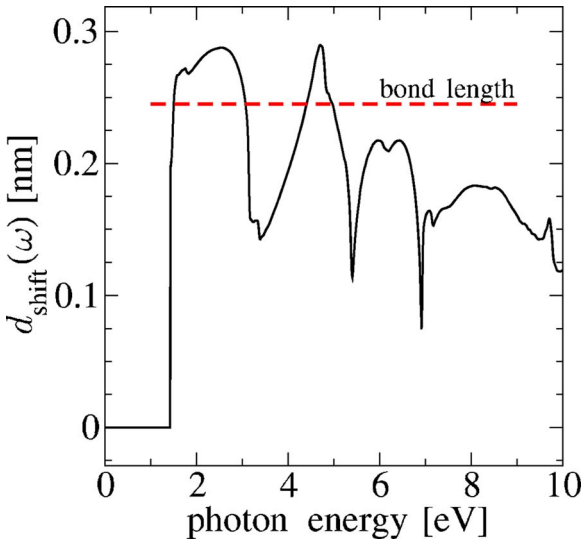

FIG. 4. (Color online) Frequency dependence of the GaAs shift distance.

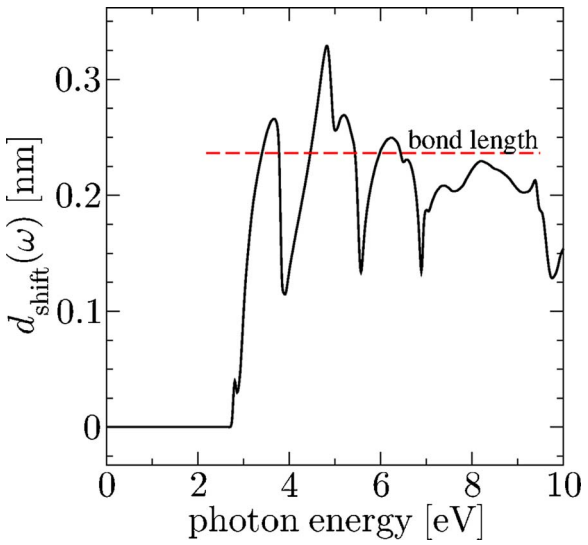

FIG. 5. (Color online) Frequency dependence of the GaP shift distance.

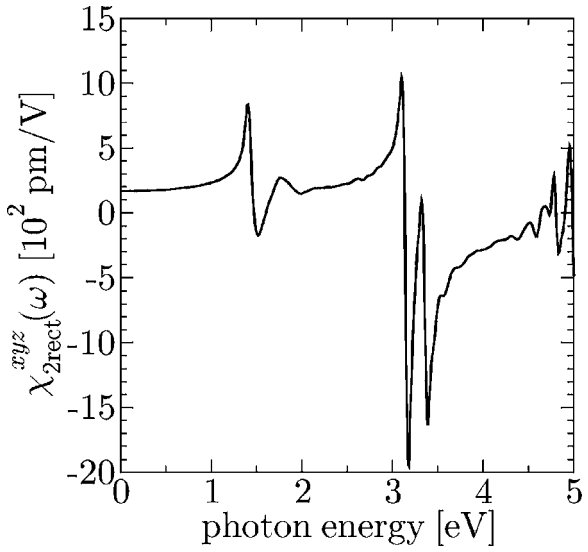

FIG. 6. Frequency dependence of the rectification tensor for GaAs.

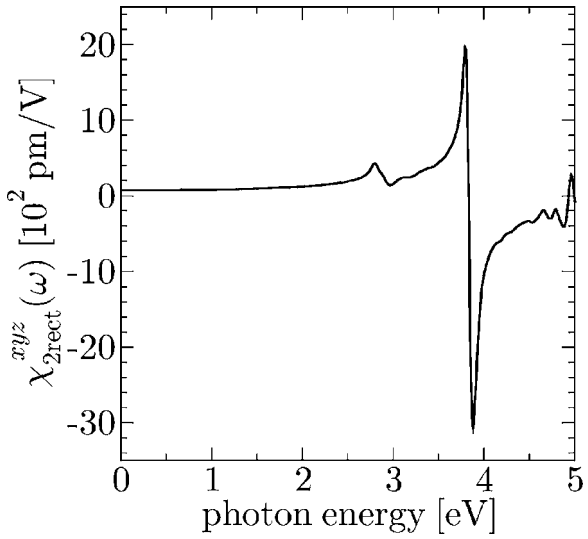

FIG. 7. Frequency dependence of the rectification tensor for GaP.

the band structure and wave function symmetries are no longer as simple as near the band edge. We see that there are pronounced dips at certain photon energies, where the shift distance falls to about half its peak value. Upon comparison to the linear response spectra these dips correspond to energies of high absorption, where valence and conduction band energies are nearly parallel. This happens along high symmetry lines in the BZ, and the wave functions of k points in these regions of the BZ are not as simple to model as those of the high-symmetry points.

An immediately apparent difference in the GaP shift distance, when compared to the GaAs shift distance, is that at the band edge the spectrum has a value around 1 Å, significantly less than the bond length. Just above the band edge there is a significant drop in the shift distance before it rapidly rises to the bond length. If the calculations are repeated without including the spin-orbit interaction, then this dip does not occur, and the spectrum immediately rises to the bond length.

Earlier work by von Baltz and Kraut[10] refers to the shift distance as the *anisotropy distance*. The shift distance has been experimentally estimated in the case of ferroelectrics and pyroelectrics.[10] Within a pseudopotential scheme, Hornung *et al.* calculated the anisotropy distance for doped GaP to be 0.9 nm for photon energies of 0.4 eV, a factor of four larger than the bond length[49] but in agreement to experiment.[50] To our knowledge, no similar experiment on insulating GaP or GaAs, measuring the anisotropy distance from photon energies crossing the direct band gap, has been reported. It is surprising that an effect as simple and as intrinsic as this has not been experimentally investigated more thoroughly.

Besides the excitation process, the effect of scattering on the electron density after excitation is interesting in its own right. The real space motion of the electrons as they relax back to their ground states is not well understood. If [111] linearly polarized light is used to excite the electrons, then it is known that the excitation occurs directed along the bond from the cation atom to the anion. While the electrons are in the conduction band, if enough scattering events occur and the electron momentum distribution is randomized, then the electrons will lose "memory" of which cation they originally came from, and will relax equally likely to any of the neighboring cations. So, while the shift current was induced preferentially along a direction, the relaxation current density could very well be isotropic if the scattering is significant enough. This would lead to a net current even if the applied field were monochromatic. This should be contrasted with shift currents generated in finite systems: in these systems, the electron typically only has one pathway back to its donor state. As the electron relaxes it produces a back current, and in the continuous wave limit there is no net current. It would be interesting to see if an experimental investigation of the shift current, perhaps through time-domain THz spectroscopy, could reveal information about the scattering processes in semiconductor crystals.

B. Rectification tensor

We now turn to rectification current in the continuous wave limit. In Figs. 6 and 7 we plot the rectification tensor spectrum for GaAs and GaP. The tensors display an expected resonance at critical points. At the band gap, the resonance in the GaAs spectrum is large enough to make the spectrum negative just above the band gap. This change in sign of the bulk rectification tensor has been experimentally seen in the work of Zhang *et al.*,[51] in which they investigated the off-normal incidence rectification in a GaAs (111) crystal. Unfortunately, a more detailed quantitative comparison cannot be presented, since they did not extract a value for the $\chi^{xyz}_{2\mathrm{rect}}$ tensor component at photon energies above the band gap. Moreover, although present, the photovoltaic effect (shift current) was not considered in their analysis. For GaP we find that the spectrum remains positive until higher photon energies associated with the E_1 transition at around 3.9 eV. Although a resonance at the band gap is present, it is not

strong enough to cause a sign change in the rectification spectrum.

For the static limit we find that $\chi_2^{xyz}(0;0,0) = 176.7$ pm/V, and 72.8 pm/V for GaAs and GaP, respectively. These results for the static limit are in agreement with other calculations done for the static limit $\chi_2^{xyz}(0;0,0)$ of the zincblende crystals.[52] Our static limits are also in close agreement with the experimental low frequency values of 166 pm/V for GaAs and 74 pm/V for GaP, as presented by Roberts,[53] but determined from second-harmonic measurements. There is a wide variation in the low-frequency values of χ_2^{xyz} for GaAs from electro-optic measurements. These values range from 100 to 200 pm/V[54,55] and the best we can say is that our value for GaAs falls within this range.

C. Temporal response

We now investigate the importance of the frequency components in $\chi_2^{abc}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ for which $\omega_\beta \neq -\omega_\gamma$ on the nonlinear response of GaAs to a short laser pulse. We assume a Gaussian pulse profile given by

$$\mathbf{E}_{\text{env}}(t) = \mathbf{E}_\circ e^{-(2t^2/\sigma^2)\ln 2}.$$

The carrier frequency is $\hbar\omega_\circ = 1.55$ eV, $E_\circ$ is the peak field strength, and σ is the full temporal width at half maximum of the intensity. For convenience we assume the laser is incident on a (110) GaAs crystal, and the field is polarized in the $[\bar{1}10]$ direction so that $\mathbf{E}_\circ = E_\circ(-\hat{\mathbf{x}}+\hat{\mathbf{y}})/\sqrt{2}$.

Expressing the interband polarization in terms of our single-frequency effective coefficients we have from Eq. (22)

$$\begin{aligned}\langle P^z_{\text{inter}}(t)\rangle = &-2\left(\frac{2}{\omega_\circ}\Lambda^{xyz}_{\text{I}}(\omega_\circ) + 4C^{xyz}_{\text{I}}(\omega_\circ)\right)E^2_{\text{env}}(t)\\ &-2\left(\frac{8}{\omega_\circ}\tilde{C}^{xyz}_{\text{II}}(\omega_\circ) + \frac{2}{\omega^2_\circ}\Lambda^{xyz}_{\text{II}}(\omega_\circ) + 4C'^{xyz}_{\text{II}}(\omega_\circ)\right.\\ &\left.+2\Lambda'^{xyz}_{\text{II}}(\omega_\circ)\right)\frac{\partial}{\partial t}E^2_{\text{env}}(t),\end{aligned} \tag{26}$$

where the prime denotes differentiation with respect to ω. Note that the bracketed quantity in the first term is only part of the optical rectification tensor in Eq. (25). The remaining contributions to the rectification appear in the intraband current $\langle J^z_{\text{intra}}(t)\rangle^{\text{I}}$ which, for the same optical pulse, is

$$\begin{aligned}\langle J^z_{\text{intra}}(t)\rangle^{\text{I}} = &\,4\Lambda^{xyz}_{\text{II}}(\omega_\circ)E^2_{\text{env}}(t) + 2\Lambda'^{xyz}_{\text{I}}(\omega_\circ)\frac{\partial}{\partial t}E^2_{\text{env}}(t)\\ &-4\Lambda''^{xyz}_{\text{II}}(\omega_\circ)E_{\text{env}}(t)\frac{\partial^2}{\partial t^2}E_{\text{env}}(t).\end{aligned} \tag{27}$$

The other single-frequency effective tensors such as $S^{xyz}_{\text{I}}(\omega)$, and $C^{xyz}_{\text{I}}(\omega)$ would appear in these expressions if we had considered more complicated pulses, such as those involving chirp. In Eq. (27) we have expanded $\langle J^z_{\text{intra}}(t)\rangle^{\text{I}}$ to include second-order derivatives of the envelope $E_{\text{env}}(t)$. This is to keep to orders consistent with the current from the polarization in Eq. (26), which would also be to second order.

The total current is the sum $\langle J^z_{\text{intra}}(t)\rangle^{\text{I}} + d\langle P^z_{\text{inter}}(t)\rangle/dt$. We can then proceed to identify the shift current $\mathbf{J}_{\text{shift}}(t)$ Eq. (23)

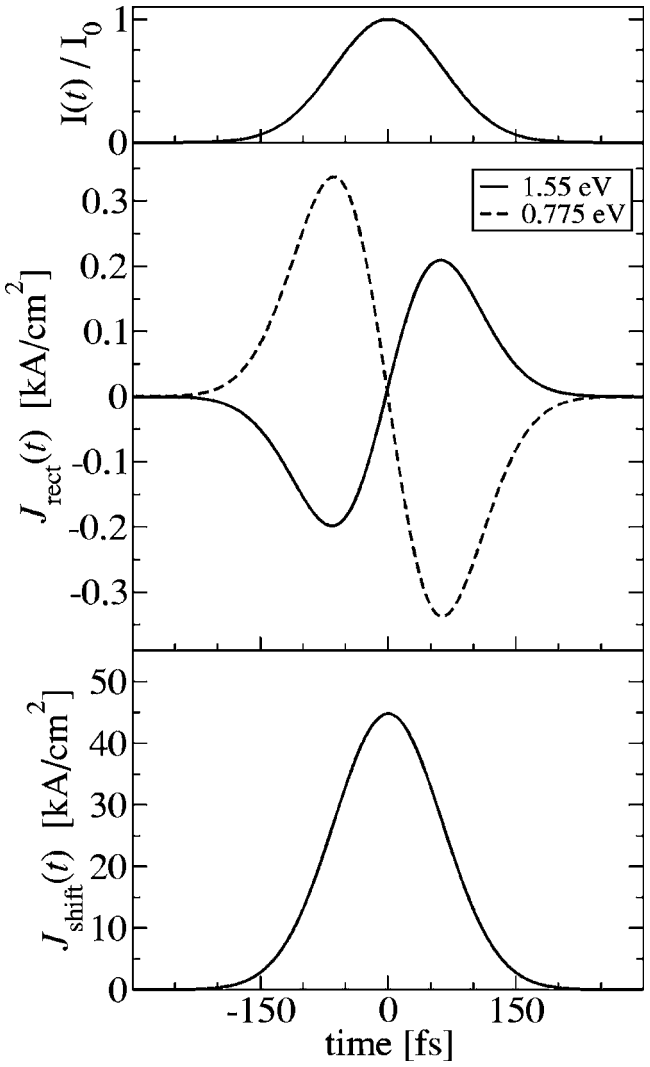

FIG. 8. The current response in GaAs to a 150 fs pulse showing the difference between above and below band gap excitation. The top panel shows the intensity profile $I(t)$ normalized to the peak intensity I_0. The middle panel shows the rectification current response to this pulse for photon energies below and just above the band gap. The bottom panel shows the shift current, which follows the pulse profile. The net current response is the sum of the bottom two panels.

which, as explained in Sec. II, varies with the pulse intensity; it is the first term on the right-hand side of Eq. (27). The remainder of the total response is the rectification current. In Fig. 8 we show the magnitude of the currents from our calculations for a 150 fs pulse. We use a peak field strength of $E_\circ = 10^7$ V/m. The top panel shows the pulse intensity profile, normalized to the peak intensity. In the middle and bottom panels, we plot the rectification and shift currents, respectively. The total current response is the sum of the two bottom panels. We have divided out the peak intensity, so that the curves can be compared to the magnitudes of the susceptibilities. For these pulse parameters the rectification current follows the time derivative of pulse, and the shift current follows the pulse profile. Clearly, the shift current dominates the response above the band gap. The sign change

in the rectification part of current reflects the change in sign of $\chi^{abc}_{2\mathrm{rect}}(0;\omega,-\omega)$ just above the band gap.

Although much experimental work on THz radiation from GaAs has been reported, most researchers do not present quantitative details of the measured THz field, and focus mainly on qualitative features such as the line shape. An experiment by Coté *et al.*[19,56] has attempted to distinguish between the temporal response of the rectification current and shift current. Since the former is proportional to the time derivative of the latter, one signature of the difference is a phase shift in the signal. This phase shift is a natural consequence of our formalism, appearing in the divergence $(-i\omega_\Sigma)^{-1}$ of Eq. (10). In their experiments, Coté *et al.* used a (110) GaAs crystal, and $[\bar{1}10]$ polarized electric fields as we have used in our calculations above. They found that the shift current densities exceeded the rectification current density by a factor of 570. For the experimental parameters we chose here, we find that the factor is roughly 180, off by a factor of three. However, a complete quantitative comparison of our results to this experiment is prohibitive; to do this would require solving for the full electron dynamics including scattering and space-charge effects, and solving the Maxwell problem for the emission of the THz radiation through the interface and the effects of the optical elements used in its detection. Steps towards dealing with the latter of these problems have been reported,[56,57] and the tensors calculated here will be an important input to full studies of current generation and transport in semiconductors.[58]

With shorter pulses, one expects the off-diagonal coefficients to play a more important role, since there will be a larger range of frequencies mixing to produce the slowly-varying response. One also expects the shorter pulse to result in a larger rectification current, since this current is proportional to the time derivative of the pulse intensity. As an example of what one should expect with shorter pulses, in Fig. 9 we plot our results assuming a Gaussian pulse, but with $\sigma=10$ fs. The spectral width of such a pulse is approaching the limit of validity for our linear expansions. Here, one notices a different rectification response. The rectification response above the band gap is now asymmetric, because the dispersion terms are now important. At frequencies closer to the band gap, where the tensors vary more rapidly, one can find even more asymmetric profiles for the rectification current. By definition, the shift current again follows the pulse profile. The most striking difference arising due to the shorter pulse is that, because the pulse profile now varies more rapidly, the rectification current is roughly two orders of magnitude larger than with the 150 fs pulse, and is comparable in magnitude to the shift current.

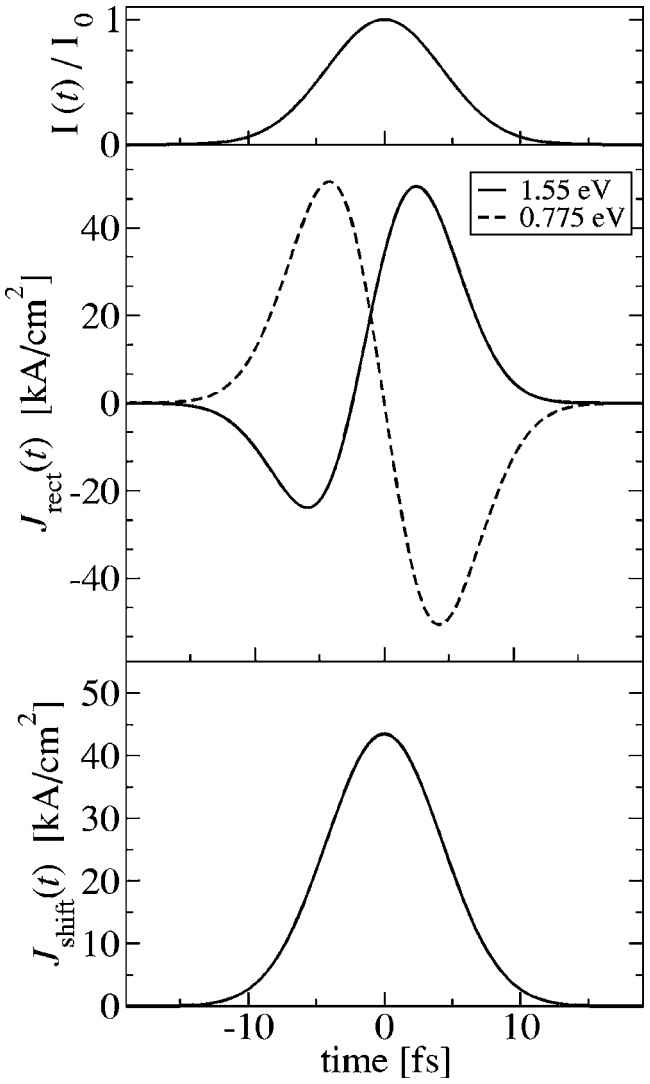

FIG. 9. Same as Fig. 8 but with a FWHM intensity of $\sigma=10$ fs.

V. CONCLUSIONS

We have derived expressions for the second-order response tensor $\chi_2(\omega_\Sigma;\omega_\beta,\omega_\gamma)$ in the optical rectification limit, where $\omega_\Sigma\approx 0$, and applied them to the zincblende structures GaAs and GaP using a FLAPW scheme. This has allowed us to compute the optical rectification response tensor $\chi_{2\mathrm{rect}}(0;\omega,-\omega)$ both below and above the band gap from a full band structure scheme. Above the band gap, the photovoltaic shift current is associated with this response. We have computed the frequency dependence of the average shift of the electron density responsible for this current. In agreement with simple models relevant to the band edge of GaAs, we find that the shift distance is approximately the bond length, but at certain laser frequencies the shift falls well below that distance. In GaP, we find that the shift distance is significantly smaller than the bond length very near the band edge.

We have also included the effects of off-diagonal dispersion in the susceptibility tensor, and shown that they become appreciable for pulses shorter than approximately 10 fs. Using the same approach we have presented here, the off-diagonal dispersion in other nonlinear effects, such as sum frequency generation, can be addressed.

Future work involves investigating the optical rectification in lower symmetry crystals, such as the *wurtzite* structures. These crystals allow the circular photovoltaic effect, which can be comparable in magnitude to the shift current

response, but are forbidden in the zincblende crystals which we have studied here.

ACKNOWLEDGMENTS

This work was supported by the National Sciences and Engineering Research Council of Canada and Photonics Research Ontario. F.N. acknowledges support by the Ontario Graduate Scholarship program, the Walter C. Sumner Foundation, and a Nortel/JDS/Uniphase photonics scholarship. We acknowledge helpful discussions with B. Adolph in the early stages of the work, B. Olejnik on the LATM method, and with D. Coté and N. Laman with regards to experimental details.

APPENDIX A

In this appendix, we outline how $\tilde{\chi}^{abc}_{2\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ can be decomposed into single-frequency effective tensors. From Eq. (12) we can write

$$\tilde{\chi}^{abc}_{2\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)=\left\{\frac{\rho_\beta}{\rho_\gamma}\bar{T}^{abc}(\omega_\beta;\rho_\beta,\rho_\gamma)+\frac{1}{\rho_\gamma}S^{abc}(\omega_\beta+\omega_\gamma)+\frac{1}{\rho_\gamma^2}W^{abc}(\omega_\beta+\omega_\gamma)\right\}+(bc\beta\gamma\leftrightarrow cb\gamma\beta),$$

where

$$\bar{T}^{abc}(\omega;\rho_1,\rho_2)=\frac{ie^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m}f_{nm}r^b_{mn}\left[\frac{r^a_{nm;c}}{\omega_{mn}}+\frac{\rho_1 r^a_{nm}\Delta^c_{mn}}{\rho_2\omega^2_{mn}}\right]F_+(\omega_{mn},\omega),$$

$$S^{abc}(\omega)=\frac{ie^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m}f_{nm}\frac{r^a_{nm}r^b_{mn;c}}{\omega_{mn}}F_+(\omega_{mn},\omega),$$

and

$$W^{abc}(\omega)=-\frac{ie^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m}f_{nm}\frac{r^a_{nm}r^b_{mn}\Delta^c_{mn}}{\omega^2_{mn}}F_+(\omega_{mn},\omega).$$

Note that $\rho_\beta/\rho_\gamma=\omega_\beta/\omega_\gamma$, and since we are interested in frequency components where $\omega_\beta\approx-\omega_\gamma$ but neither ω_β nor ω_γ are close to zero, we will not have any divergence problems in evaluating the above expression for $\tilde{\chi}^{abc}_{2\mathrm{tra}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ once the effective coefficients $\bar{T}^{abc}(\omega;\rho_1,\rho_2)$, $S^{abc}(\omega)$ and $W^{abc}(\omega)$ are known. Our expression for $\bar{T}^{abc}(\omega;\rho_1,\rho_2)$ still contains two frequency components. Upon further simplification to decouple the frequencies we find that

$$\bar{T}^{abc}(\omega;\rho_1,\rho_2)=[S^{bac}(\omega)]^*+\frac{\omega_1}{\omega_2}W^{abc}(\omega).$$

Finally, one can show that $S^{abc}(\omega)=\omega^{-1}\Lambda^{abc}(\omega)$, which when used in the expressions above gives Eq. (19).

APPENDIX B

Here we outline how $\chi^{abc}_{2\mathrm{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)$ can be decomposed into single-frequency tensors. From Eq. (11) we can write

$$\chi^{abc}_{2\mathrm{ter}}(-\omega_\Sigma;\omega_\beta,\omega_\gamma)=\bar{C}^{abc}(\omega_\beta,\rho_\beta)+\frac{\rho_\gamma}{\rho_\beta}\bar{D}^{abc}(\omega_\gamma,\rho_\beta)+\frac{1}{\rho_\beta}\bar{E}^{abc}(\omega_\Sigma,\rho_\beta)+(bc\beta\gamma\leftrightarrow cb\gamma\beta),$$

where

$$\bar{C}^{abc}(\omega,\rho)=\frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^a_{pm}r^b_{mn}r^c_{np}}{\rho^{-1}\omega_{mn}-\omega_{mp}}F_+(\omega_{mn},\omega),$$

$$\bar{D}^{abc}(\omega,\rho)=\frac{-e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^c_{nm}r^a_{mp}r^b_{pn}}{\rho^{-1}\omega_{np}-\omega_{mp}}F_+(\omega_{mn},\omega),$$

and

$$\bar{E}^{abc}(\omega,\rho)=\frac{-e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^a_{nm}r^b_{mp}r^c_{pn}}{\rho^{-1}\omega_{mp}-\omega_{mn}}F_+(\omega_{mn},\omega).$$

Unfortunately, these expressions still simultaneously involve two frequency components. It does not seem possible to further separate these expressions into separate frequency components, but note that for the frequency ranges relevant in OR, we have $|\rho_\beta|\gg1,|\rho_\gamma|\gg1$. Hence, we are led to try the expansions

$$\bar{C}^{abc}(\omega,\rho)=C^{abc}(\omega)+\rho^{-1}\tilde{C}^{abc}(\omega)+\cdots,$$

$$\bar{D}^{abc}(\omega,\rho)=D^{abc}(\omega)+\rho^{-1}\tilde{D}^{abc}(\omega)+\cdots,$$

and

$$\bar{E}^{abc}(\omega,\rho)\approx E^{abc}(\omega)+\rho^{-1}\tilde{E}^{abc}(\omega)+\cdots,$$

and keep only the first term. After some manipulations, following in spirit those detailed in Appendix B of SS, we find that $D^{abc}(\omega)=-C^{abc}(\omega)$, and $\tilde{D}^{abc}(\omega)=-C^{acb}(\omega)+\tilde{C}^{acb}(\omega)$, where

$$C^{abc}(\omega)=-\frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^b_{nm}r^a_{mp}r^c_{pn}}{\omega_{mp}}F_+(\omega_{mn},\omega),$$

$$\tilde{C}^{abc}(\omega)=-\frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^b_{nm}r^a_{mp}r^c_{pn}}{\omega^2_{mp}}\omega_{mn}F_+(\omega_{mn},\omega),$$

$$E^{abc}(\omega)=\frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^c_{nm}r^b_{mp}r^c_{pn}}{\omega_{mn}}F_+(\omega_{mn},\omega),$$

and

$$\tilde{E}^{abc}(\omega)=\frac{e^3}{4\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{n,m,p}f_{nm}\frac{r^a_{nm}r^b_{mp}r^c_{pn}}{\omega^2_{mn}}\omega_{mp}F_+(\omega_{mn},\omega).$$

Combining all this and simplifying gives Eq. (20).

[1] P. W. Milonni and J. H. Eberly, *Lasers* (John Wiley & Sons, New York, 1998).
[2] M. Bass, P. A. Franken, J. F. Ward, and G. Weinrich, Phys. Rev. Lett. **9**, 446 (1962).
[3] X. C. Zhang, B. B. Hu, J. T. Darrow, and D. H. Auston, Appl. Phys. Lett. **56**, 1011 (1990).
[4] A. Nahata, A. S. Weiling, and T. F. Heinz, Appl. Phys. Lett. **69**, 2321 (1996).
[5] C. Flytzanis, Phys. Rev. B **6**, 1264 (1972).
[6] S. L. Chuang, S. Schmitt-Rink, B. I. Greene, P. N. Saeta, and A. F. J. Levi, Phys. Rev. Lett. **68**, 102 (1992).
[7] Some authors use the term optical rectification to mean any second-order dc response, including the shift and injection currents. We shall keep these distinct, using the term optical rectification to imply the response associated with an induced finite polarization of the electron density which exists for excitation even below the band gap.
[8] R. von Baltz, in *Ultrafast Dynamics of Quantum Systems: Physical Processes and Spectroscopic Techniques*, edited by B. Di Bartolo, Nato ASI Series B, Vol. 372 (Plenum, 1998).
[9] B. I. Sturman and V. M. Fridkin, *The Photovoltaic and Photorefractive Effects in Noncentrosymmetric Materials* (Gordon and Breach Science Publishers, Philadelphia, 1992).
[10] R. von Baltz and W. Kraut, Phys. Rev. B **23**, 5590 (1981).
[11] P. Kral, E. J. Mele, and D. Tomanek, Phys. Rev. Lett. **85**, 1512 (2000).
[12] M. C. Beard, G. M. Turner, and C. A. Schmuttenmaer, J. Phys. Chem. A **106**, 878 (2002).
[13] H. Presting and R. von Baltz, Phys. Status Solidi B **112**, 559 (1982).
[14] J. B. Khurgin, J. Opt. Soc. Am. B **11**, 2492 (1994).
[15] J. B. Khurgin and P. Voisin, Phys. Rev. Lett. **81**, 3777 (1998).
[16] B. A. Foreman, Phys. Rev. Lett. **84**, 4513 (2000).
[17] J. B. Khurgin and P. Voisin, Phys. Rev. Lett. **84**, 4514 (2000).
[18] P. Kral, J. Phys.: Condens. Matter **12**, 4851 (2000).
[19] D. Côté, N. Laman, and H. M. van Driel, Appl. Phys. Lett. **80**, 905 (2002).
[20] D. H. Auston, A. M. Glass, and A. A. Ballman, Phys. Rev. Lett. **28**, 897 (1972).
[21] J. E. Sipe and A. I. Shkrebtii, Phys. Rev. B **61**, 5337 (2000).
[22] Z. H. Levine, Phys. Rev. B **49**, 4532 (1994).
[23] A. Dal Corso, F. Mauri, and A. Rubio, Phys. Rev. B **53**, 15638 (1996).
[24] J. L. P. Hughes and J. E. Sipe, Phys. Rev. B **53**, 10751 (1996).
[25] S. N. Rashkeev, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **57**, 3905 (1998).
[26] J. L. P. Hughes, Y. Wang, and J. E. Sipe, Phys. Rev. B **55**, 13630 (1997).
[27] S. Bergfeld and W. Daum, Phys. Rev. Lett. **90**, 036801 (2003).
[28] When using the Pauli Hamiltonian the velocity operator in coordinate representation has the form $\hat{\mathbf{v}}=(\hbar/mi)(\partial/\partial\mathbf{r})+(\hbar/4m^2c^2)\boldsymbol{\sigma}\times\boldsymbol{\nabla}V$. As is commonly done, we neglect the spin-orbit effect on the velocity, and drop the second term of this expression.
[29] C. Aversa and J. E. Sipe, Phys. Rev. B **52**, 14636 (1995).
[30] The quantity $\bar{K}_2^{abc}$ is related to K^{abc} in Ref. 21 by $\bar{K}_2^{abc}=-i\omega_\Sigma K^{abc}$.
[31] S. Scandolo and F. Bassani, Phys. Rev. B **51**, 6925 (1995).
[32] O. Jepsen and O. K. Andersen, Solid State Commun. **9**, 1763 (1971).
[33] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[34] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[35] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, and J. Luitz, WIEN2K, *An Augmented Plane Wave+Local Orbitals Program for Calculating Crystal Properties* (Karlheinz Schwarz, Techn. Universität Wien, Austria, 2001).
[36] C. Ambrosch-Draxl and J. O. Sofo, cond-mat/0402523 (unpublished).
[37] D. J. Singh, *Plane Waves, Pseudopotentials and the LAPW Method* (Kluwer Academic, Boston, 1994).
[38] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
[39] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[40] P. E. Blöchl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16223 (1994).
[41] W. A. Harrison, *Electronic Structure and the Properties of Solids* (W. H. Freeman, San Francisco, 1980).
[42] A. H. MacDonald, S. H. Vosko, and P. T. Coleridge, J. Phys. C **12**, 2991 (1979).
[43] S.-Y. Ren and W. A. Harrison, Phys. Rev. B **23**, 762 (1981).
[44] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[45] F. Nastos, B. Olejnik, K. Schwarz, and J. E. Sipe, Phys. Rev. B **72**, 045223 (2005).
[46] R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11789 (1993).
[47] B. Adolph and F. Bechstedt, Phys. Rev. B **57**, 6519 (1998).
[48] V. I. Gavrilenko and R. Q. Wu, Phys. Rev. B **61**, 2632 (2000).
[49] D. Hornung, R. von Baltz, and U. Rössler, Solid State Commun. **48**, 225 (1983).
[50] A. F. Gibson, C. B. Hatch, M. F. Kimmitt, S. Kothari, and A. Serafetinides, J. Phys. C **10**, 905 (1977).
[51] X. C. Zhang, Y. Jin, K. Yang, and L. J. Schowalter, Phys. Rev. Lett. **69**, 2303 (1992).
[52] Compare with the table of $\chi_2(0;0,0)$ values given in Ref. 25.
[53] D. A. Roberts, IEEE J. Quantum Electron. **28**, 2057 (1992).
[54] T. Y. Chang, N. van Tran, and C. K. N. Patel, Appl. Phys. Lett. **13**, 357 (1968).
[55] S. Adachi, *GaAs and Related Materials: Bulk Semiconducting and Superlattice Properties* (World Scientific, Teaneck, NJ, 1994).
[56] D. Côté, J. E. Sipe, and H. M. van Driel, J. Opt. Soc. Am. B **20**, 1374 (2003).
[57] N. Laman, M. Bieler, and H. M. van Driel, J. Appl. Phys. **98**, 103507 (2005).
[58] M. B. Johnston, D. M. Whittaker, A. Corchia, A. G. Davies, and E. H. Linfield, Phys. Rev. B **65**, 165301 (2002).