Effects of nonlocality on second-harmonic generation in bulk semiconductors

J. L. Cabellos,[1] Bernardo S. Mendoza,[1,*] M. A. Escobar,[1] F. Nastos,[2] and J. E. Sipe[2]
[1]*Department of Photonics, Centro de Investigaciones en Óptica, León, Guanajuato, Mexico*
[2]*Department of Physics and Institute for Optical Sciences, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7*

(Received 21 July 2009; revised manuscript received 5 October 2009; published 28 October 2009)

Calculations of the second-harmonic susceptibility tensor $\chi^{abc}(-2\omega;\omega,\omega)$ are presented for bulk semiconductors within both the $\mathbf{v}\cdot\mathbf{A}$ and the $\mathbf{r}\cdot\mathbf{E}$ gauges. The description of the semiconductor states incorporates the "scissors" Hamiltonian commonly used to obtain the correct band gap. The nonlocality of the scissors correction leads to terms in $\chi^{abc}(-2\omega;\omega,\omega)$ not considered before within a sum-over-states approach to the $\mathbf{v}\cdot\mathbf{A}$ gauge. Using this expression, we show that the results of the two gauges give the same result for $\chi^{abc}(-2\omega;\omega,\omega)$, within very good numerical accuracy. As part of the derivation, we clarify the well-known result for the linear optical response which states that the scissors correction rigidly shifts the spectrum along the energy axis, keeping the line-shape intact. The calculation is presented for GaAs using an *all-electron* scheme and a *pseudopotential* scheme.

DOI: 10.1103/PhysRevB.80.155205 PACS number(s): 71.15.Mb, 42.65.Ky, 42.65.An, 78.66.−w

I. INTRODUCTION

The development of nonlinear optical materials is an active area of research, with works ranging from the study and growth of nonlinear crystals to the design of metamaterials. Perhaps the simplest nonlinear process, second-harmonic generation (SHG), is one of the most important for the generation of frequencies, as a spectroscopic probe, and because its reverse process—spontaneous parametric downconversion, which is described by the same nonlinear susceptibility—can be used to generate entangled photons for application in quantum information processing.

The numerical calculation of any nonlinear optical response is a nontrivial task, and different methodologies and numerical approaches have been employed. Our interest is in strategies that can be applied to study the nonlinear optical response of a material over a wide frequency range, and in a regime where a perturbative treatment is appropriate. The first attempt along these lines is the work of Butcher and McLean,[1] where the resulting equations appeared to be plagued by divergences that appear in the dc (static) limit. Aspnes[2] showed that in the static limit these divergences are only apparent, in that the coefficients that multiply the divergent terms vanish; but his proof was limited to cubic crystals. Ghahramani *et al.*,[3] gave a more general proof of the disappearance of these apparent divergence for cold ($T=0$ K), undoped semiconductors of any crystal class. Levine,[4] presented a formula for the nonlinear second-order susceptibility tensor where the scissors approximation is properly introduced, but the expressions are difficult to compute. These studies used what is sometimes called the "velocity gauge" or "$\mathbf{v}\cdot\mathbf{A}$ gauge" for the treatment of the coupling of an electron to the electromagnetic field, where $\mathbf{v}$ is the velocity operator of the electron and $\mathbf{A}$ is the vector potential specifying the electromagnetic field. Later, Aversa and Sipe[5] showed that a divergence free expression for the nonlinear second-order susceptibility tensor $\chi^{abc}(-2\omega;\omega,\omega)$ could be more easily obtained using what is sometimes called the "length gauge" or "$\mathbf{r}\cdot\mathbf{E}$ formulation." Here $\mathbf{r}$ is the position operator and $\mathbf{E}$ is the electric field.

In the works of Rashkeev *et al.*,[6] and Hughes and Sipe,[7] the length-gauge formulation was used to evaluate $\chi^{abc}(-2\omega;\omega,\omega)$ for several zinc-blende semiconductors within an *ab initio* scheme. The more recent work of Leitsmann *et al.*,[8] extends the velocity-gauge approach to include excitonic and local field interactions in GaAs. Quasiparticle effects, at the scissors correction level, have been correctly incorporated by Nastos *et al.*[9] in the length-gauge approach, and before this Adolph and Bechstedt[10] discussed how to include these effects, even beyond the scissors approximation, within the velocity-gauge approach. Surface second-harmonic generation has also been studied within the velocity-gauge scheme with good success,[11–13] and $\chi^{abc}(-2\omega;\omega,\omega)$ spectra have been calculated for superlattices within both the length-gauge approach[14] and the velocity-gauge approach.[3]

However, a full comparison between calculations using these two different approaches has not been done. One goal of this paper is to establish the equivalence between the length-gauge and velocity-gauge schemes. Of course, it is well known that measurable quantities must be gauge invariant, and indeed we show in this paper that the expressions for $\chi^{abc}(-2\omega;\omega,\omega)$ from the two different approaches give the same result. In order to do so, we derive a expression for $\chi^{abc}(-2\omega;\omega,\omega)$ within the velocity gauge that properly takes into account the nonlocal nature of the scissors Hamiltonian. In all previous calculations of $\chi^{abc}(-2\omega;\omega,\omega)$ within the velocity gauge, the scissors implementation was carried out by following its implementation for linear optical response. We show that this naïve procedure, of shifting the conduction energies and renormalizes the velocity matrix elements, which works for the linear response, does not work at all for the nonlinear response. The expression we derive for $\chi^{abc}(-2\omega;\omega,\omega)$ contains two terms directly obtained from the scissors Hamiltonian that are clearly required to obtain gauge invariance within the scissors implementation. Earlier, Nastos *et al.*[9] showed the correct way of calculating $\chi^{abc}(-2\omega;\omega,\omega)$ using the scissors Hamiltonian within the length gauge, and in the present paper we can identify a unified approach to the calculation of $\chi^{abc}(-2\omega;\omega,\omega)$, with and without the scissors correction, that gives the same result for the velocity gauge and length gauge. While we can verify this analytically for linear response,[5] for the SHG response coefficient we can only confirm it numerically. Nonetheless, with this confidence acquired our approach can serve as a model for the gauge-invariant calculation of other nonlinear optical response coefficients.

The paper is organized as follows. In Sec. II we present the most important steps involved in the derivation of second-order susceptibility tensor $\chi^{abc}(-2\omega;\omega,\omega)$, within the length-gauge and velocity-gauge approaches. In Sec. III we show the results of the numerical evaluation of $\chi^{abc}(-2\omega;\omega,\omega)$, taking as an example the zinc-blende bulk semiconductor GaAs, and discuss them. We calculate the expressions for $\chi^{abc}(-2\omega;\omega,\omega)$ with *ab initio* programs based on density-functional theory (DFT) within the local-density approximation (LDA), using all-electron and pseudopotential schemes. Finally, in Sec. IV we present our conclusions.

II. THEORY

In this section we present the strategy used to calculate the second-order nonlinear response. Although this has already been discussed in earlier studies, we consider both the velocity-gauge and the length-gauge responses within a common formalism. Our derivation includes terms not included before in the velocity gauge, for both the linear and the nonlinear responses. For the nonlinear response, the terms are crucial for establishing numerically that both gauges give the same result, as they must.

A. Perturbation approach

We use the independent particle approximation and neglect local field and excitonic effects and treat the electromagnetic field classically, while the matter is described quantum mechanically. We can describe the system using a scaled one-electron density operator ρ, with which we can calculate the expectation value of a single-particle observable $\mathcal{O}$ as $\langle\mathcal{O}\rangle=\mathrm{Tr}(\rho\mathcal{O})$, with $\mathcal{O}$ the associated quantum-mechanical operator and Tr the trace. The density operator satisfies $i\hbar(d\rho/dt)=[H(t),\rho]$, with $H(t)$ as the total single-electron Hamiltonian, written as

$$H(t)=H_0+H_I(t),$$

where H_0 is the unperturbed time-independent Hamiltonian, and $H_I(t)$ is the time-dependent potential energy due to the interaction of the electron with the electromagnetic field; H_0 has eigenvalues $\hbar\omega_n(\mathbf{k})$ and eigenstates $|n\mathbf{k}\rangle$ (Bloch states) labeled by a band index n and crystal momentum $\mathbf{k}$. To proceed with the solution of ρ it is convenient to use the interaction picture, where a unitary operator $U=\exp(iH_0t/\hbar)$ transforms any operator $\mathcal{O}$ into $\widetilde{\mathcal{O}}=U\mathcal{O}U^{\dagger}$. Even if $\mathcal{O}$ does not depend on t, $\widetilde{\mathcal{O}}$ does through the explicit time dependence of U. The dynamical equation for $\widetilde{\rho}$ is given by

$$i\hbar\frac{d\widetilde{\rho}}{dt}=[\widetilde{H}_I(t),\widetilde{\rho}]$$

with solution

$$i\hbar\widetilde{\rho}(t)=i\hbar\widetilde{\rho}_0+\int_{-\infty}^{t}dt'[\widetilde{H}_I(t'),\widetilde{\rho}(t')], \tag{1}$$

where $\widetilde{\rho}_0=\widetilde{\rho}(t=-\infty)$ is the unperturbed density matrix. We look for the standard perturbation series solution, $\widetilde{\rho}(t)=\widetilde{\rho}^{(0)}+\widetilde{\rho}^{(1)}+\widetilde{\rho}^{(2)}+\cdots$, where the superscript denotes the order (power) with which each term depends on the perturbation $H_I(t)$. From Eq. (1) the Nth order term is

$$\widetilde{\rho}^{(N)}(t)=\frac{1}{i\hbar}\int_{-\infty}^{t}dt'[\widetilde{H}_I(t'),\widetilde{\rho}^{(N-1)}(t')]. \tag{2}$$

The series is generated by the unperturbed density operator $\widetilde{\rho}^{(0)}\equiv\widetilde{\rho}_0$, assumed to be the diagonal Fermi-Dirac distribution, $\langle n\mathbf{k}|\widetilde{\rho}_0|n\mathbf{k}\rangle=f[\hbar\omega_n(\mathbf{k})]\equiv f_n$. For a clean, cold semiconductor $f_n=1$ for n a valence (v) or occupied band and zero for n a conduction (c) or empty band. This we assume throughout. We remark that the expectation values satisfy $\langle\mathcal{O}\rangle=\mathrm{Tr}(\rho\mathcal{O})=\mathrm{Tr}(\widetilde{\rho}\widetilde{\mathcal{O}})$.

We first look for the expectation value of the macroscopic current density, $\mathbf{J}$, given by

$$\langle\mathbf{J}\rangle=\frac{e}{\Omega}\mathrm{Tr}(\rho\dot{\mathbf{r}}), \tag{3}$$

where $\dot{\mathbf{r}}$ is the time derivative of the position operator of the electron of charge e,

$$\mathbf{v}\equiv\dot{\mathbf{r}}=\frac{1}{i\hbar}[\mathbf{r},H] \tag{4}$$

with $\mathbf{v}$ the velocity operator of the electron, and Ω is the normalization volume. We calculate the macroscopic polarization density $\mathbf{P}$, related to $\langle\mathbf{J}\rangle$ by $\langle\mathbf{J}\rangle=d\mathbf{P}/dt$. For a perturbing (Maxwell macroscopic) electromagnetic field, $\mathbf{E}(t)=\mathbf{E}(\omega)e^{-i\widetilde{\omega}t}+\text{c.c.}$, where $\widetilde{\omega}=\omega+i\eta$ and $\eta>0$ is used to adiabatically turn on the interaction, we write the second-order nonlinear polarization as,

$$P^{a(2)}(2\omega)=\chi^{abc}(-2\omega;\omega,\omega)E^b(\omega)E^c(\omega), \tag{5}$$

where $\chi^{abc}(-2\omega;\omega,\omega)$ is the nonlinear susceptibility responsible of SHG. The superscripts in Eq. (5) denote Cartesian components, and if repeated are to be summed over. Without loss of generality we can always define $\chi^{abc}(-2\omega;\omega,\omega)$ to satisfy intrinsic permutation symmetry, $\chi^{abc}(-2\omega;\omega,\omega)=\chi^{acb}(-2\omega;\omega,\omega)$; if it did not, the part that did not satisfy intrinsic permutation symmetric would make no contribution to the second-order polarization. This part could be dropped since it would have no physical significance.

The unperturbed Hamiltonian is given by

$$H_0=\frac{p^2}{2m_e}+V(\mathbf{r}) \tag{6}$$

with m_e the mass of the electron, $\mathbf{p}$ its canonical momentum, and $V(\mathbf{r})$ is the local periodic crystal potential, where we neglect spin-orbit terms. This Hamiltonian is used to solve the Kohn-Sham equations[15] of DFT, for convenience usually within the LDA. As is well known, the use of these solutions as single particle states leads to an underestimation of the band gap. A standard procedure to correct for this is to use the so-called "scissors approximation," by which one rigidly shifts the conduction bands in energy so that the band gap corresponds to the accepted experimental band gap; this is often in fairly good agreement with the GW band gap based on a more sophisticated calculation.[16] Concurrently, one uses the LDA wave functions, since they produce band structures with dispersion relations similar to those predicted by the GW approximation. Mathematically, one adds the scissors (nonlocal) term $S(\mathbf{r},\mathbf{p})$, to the unperturbed or unscissored Hamiltonian H_0, i.e.,

$$H_0^S = H_0 + S(\mathbf{r},\mathbf{p}),$$

where

$$S(\mathbf{r},\mathbf{p}) = \hbar\Delta\sum_n \int d^3k(1-f_n)|n\mathbf{k}\rangle\langle n\mathbf{k}|, \tag{7}$$

with $\hbar\Delta$ the rigid ($\mathbf{k}$-independent) energy correction to be applied. Several properties of $S(\mathbf{r},\mathbf{p})$ are shown in the appendix. The unscissored and scissored Hamiltonians satisfy

$$H_0\psi_{n\mathbf{k}}(\mathbf{r}) = \hbar\omega_n(\mathbf{k})\psi_{n\mathbf{k}}(\mathbf{r}),$$

$$H_0^S\psi_{n\mathbf{k}}(\mathbf{r}) = \hbar\omega_n^S(\mathbf{k})\psi_{n\mathbf{k}}(\mathbf{r}),$$

where

$$\omega_n^S(\mathbf{k}) = \omega_n(\mathbf{k}) + (1-f_n)\Delta, \tag{8}$$

and $\psi_{n\mathbf{k}}(\mathbf{r})=\langle\mathbf{r}|n\mathbf{k}\rangle$ is the coordinate representation of the ket $|n\mathbf{k}\rangle$. We emphasize that the scissored Hamiltonian has the same eigenfunctions as the unscissored Hamiltonian.

B. Velocity gauge formalism

To calculate the optical response in the velocity gauge, we use the minimal substitution through which, in the presence of an electromagnetic field, the Hamiltonian is written as

$$H^S = \frac{1}{2m_e}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + V(\mathbf{r}) + S\left(\mathbf{r},\mathbf{p} - \frac{e}{c}\mathbf{A}\right), \tag{9}$$

where $\mathbf{A}$ is the vector potential; one obtains the magnetic field as $\mathbf{B}=\nabla\times\mathbf{A}$ and the electric field as $\mathbf{E}=-(1/c)\dot{\mathbf{A}}$, with c the speed of light in vacuum. In general these electric and magnetic fields are taken to be the macroscopic Maxwell fields. We assume the long-wavelength limit, in which $\mathbf{A}$ is uniform and only depends on time. Furthermore, we take a harmonic perturbation of the form $\mathbf{A}(t)=\mathbf{A}(\omega)e^{-i\tilde{\omega}t}+\mathbf{A}^*(\omega)e^{i\tilde{\omega}^*t}$, where only the "positive frequency" term will be kept in the following, because that term will contribute to the positive frequency part of the linear and second-harmonic responses. Expanding the scissors operator according to,[9]

$$S\left(\mathbf{r},\mathbf{p} - \frac{e}{c}\mathbf{A}\right) = S(\mathbf{r},\mathbf{p}) + \frac{e}{c}\frac{i}{\hbar}\mathbf{A}\cdot[\mathbf{r},S(\mathbf{r},\mathbf{p})] + \frac{1}{2!}\left(\frac{e}{c}\frac{i}{\hbar}\right)^2\{\mathbf{A}\cdot\mathbf{r},[\mathbf{A}\cdot\mathbf{r},S(\mathbf{r},\mathbf{p})]\} + \cdots,$$

leads to the following scissored Hamiltonian up to second order in $\mathbf{A}$:

$$H^S = H_0^S + H_{I,1} + H_{I,2},$$

where

$$H_{I,1} = -\frac{e}{c}\mathbf{A}\cdot\mathbf{v}^{\Sigma}, \tag{10}$$

$$H_{I,2} = -\frac{ie^2}{2\hbar c^2}[r^b,v^{S,c}]A^bA^c + \frac{e^2}{2m_ec^2}A^2 \tag{11}$$

are the linear and nonlinear (second-order) interaction Hamiltonians. The $e^2A^2/(2m_ec^2)$ term is only a function of time contributing to a global phase factor to the electron wave function that has no effect on expectation values, so it can be dropped. We have defined

$$\mathbf{v}^S = -\frac{i}{\hbar}[\mathbf{r},S(\mathbf{r},\mathbf{p})] \tag{12}$$

as the contribution to the velocity operator due to the nonlocal scissors term, and

$$\mathbf{v}^{\Sigma} = \frac{\mathbf{p}}{m_e} + \mathbf{v}^S, \tag{13}$$

as the scissored velocity operator. From Eq. (4) the current operator $\mathbf{j}=e\dot{\mathbf{r}}$, up to second order in $\mathbf{A}$, is

$$j^a = j_0^a + j_1^a + j_2^a,$$

with

$$j_0^a = ev^{\Sigma,a},$$

$$j_1^a = -\frac{e^2}{cm_e}A^a + \frac{ie^2}{\hbar c}[r^a,v^{S,b}]A^b,$$

$$j_2^a = -\frac{e^3}{2\hbar^2c^2}[r^a,(r^b,v^{S,c})]A^bA^c,$$

operators of zero, first and second orders in $\mathbf{A}$, respectively. From Eq. (3),

$$\langle J^{(1)a}\rangle = \frac{1}{\Omega}\mathrm{Tr}(j_0^a\rho^{(1)}) + \frac{1}{\Omega}\mathrm{Tr}(j_1^a\rho^{(0)}), \tag{14}$$

is the linear macroscopic current density and

$$\langle J^{(2)a}\rangle = \frac{1}{\Omega}\mathrm{Tr}(j_0^a\rho^{(2)}) + \frac{1}{\Omega}\mathrm{Tr}(j_1^a\rho^{(1)}) + \frac{1}{\Omega}\mathrm{Tr}(j_2^a\rho^{(0)}), \tag{15}$$

is the nonlinear (second-order) macroscopic current density.

1. Linear response

We calculate the linear response, within the velocity gauge, and show that there is a term not previously included when the scissored Hamiltonian is used. Indeed, we show that by coincidence the "usual" way of including the scissor correction leads to the correct result. That is, the scissors correction only gives a rigid shift in the energy axis of the unscissored spectrum by an amount equal to Δ; the line shape of the spectrum is the same for both the scissored and the unscissored Hamiltonians.[9,17–19] In the following, we show that if the usual procedure is used for the nonlinear response the resulting scissored susceptibility is wrong. The derivation of the linear response here is important for making sense of our later results, and also sets some of the intermediate results that will be used in the calculation of the nonlinear response. In passing, we also show that the linear response is gauge invariant, since we obtain the same analytic result for the linear susceptibility in both gauges. This agreement holds with and without the scissors correction.

We start by taking matrix elements of Eq. (2) to obtain

$$\tilde{\rho}^{(1)}_{mn}(t) = \frac{ei}{\hbar c}\int_{-\infty}^{t} dt' A^b(t') \sum_\ell [\tilde{v}^{\Sigma,b}_{m\ell}(t')\tilde{\rho}^{(0)}_{\ell n}(t') - \tilde{\rho}^{(0)}_{m\ell} \times(t')\tilde{v}^{\Sigma,b}_{\ell n}(t')],$$

where the sum over ℓ is over all states, and we have used Eq. (10). Since $U(t)=\exp(iH_0^S t/\hbar)$, we get $\tilde{v}^{\Sigma,b}_{m\ell}(t')=v^{\Sigma,b}_{m\ell}e^{i\omega^S_{m\ell}t'}$, $\tilde{\rho}^0_{\ell n}(t')=f_\ell\delta_{\ell n}$, and $\omega^S_{m\ell}=\omega^S_m(\mathbf{k})-\omega^S_\ell(\mathbf{k})$, where we have omitted the dependence on $\mathbf{k}$ from the already crowded notation. Then $\tilde{\rho}^{(1)}_{mn}(t)=\rho^{(1)}_{mn}e^{i\omega^S_{mn}t}e^{-i\tilde{\omega}t}$, with

$$\rho^{(1)}_{mn}(t) = \frac{e}{\hbar c}\frac{v^{\Sigma,b}_{mn}f_{nm}}{\omega^S_{mn}-\tilde{\omega}}A^b(\omega), \tag{16}$$

where $f_{nm}=f_n-f_m$. Using $\mathrm{Tr}(\rho^{(0)})/\Omega=n_0$, with n_0 the electronic density, and $d\mathbf{P}/dt=\langle\mathbf{J}\rangle$ to write $P^a(\omega)=(i/\tilde{\omega})\langle J^a(\omega)\rangle=\chi^{ab}(-\omega;\omega)E^b(\omega)$, we get from Eq. (14) that

$$\begin{aligned}\chi^{ab}(-\omega;\omega) &= \frac{e^2}{\hbar\tilde{\omega}^2}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{v^{\Sigma,a}_{nm}v^{\Sigma,b}_{mn}f_{nm}}{\omega^S_{mn}-\tilde{\omega}} - \frac{e^2 n_0}{m_e\tilde{\omega}^2}\delta_{ab} + \frac{ie^2}{\hbar\tilde{\omega}^2}\frac{1}{\Omega}\mathrm{Tr}(\rho^{(0)}\mathcal{F}^{ab}) \\ &= \frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n} f_{nm}v^{\Sigma,a}_{nm}v^{\Sigma,b}_{mn} \times\left(\frac{1}{(\omega^S_{mn})^2(\omega^S_{mn}-\tilde{\omega})}+\frac{1}{(\omega^S_{mn})^2\tilde{\omega}}+\frac{1}{\omega^S_{mn}\tilde{\omega}^2}\right) \\ &\quad - \frac{e^2 n_0}{m_e\tilde{\omega}^2}\delta_{ab} + \frac{ie^2}{\hbar\tilde{\omega}^2}\frac{1}{\Omega}\mathrm{Tr}(\rho^{(0)}\mathcal{F}^{ab}),\end{aligned} \tag{17}$$

is the linear susceptibility within the scissored Hamiltonian; we used a partial fraction expansion in the first term after the first equal sign. We have defined

$$\mathcal{F}^{ab} = [r^a, v^{S,b}], \tag{18}$$

used the fact that the f_{nm} factor allow us to write $m\neq n$ and that in the continuous limit of $\mathbf{k}$ $(1/\Omega)\Sigma_{\mathbf{k}}\to\int d^3k/(8\pi^3)$.

From time reversal symmetry we have that $v^S_{mn}(-\mathbf{k})=-v^S_{nm}(\mathbf{k})$ and $\omega^S_{mn}(-\mathbf{k})=\omega^S_{mn}(\mathbf{k})$ with which it follows that the contribution to $\chi^{ab}(-\omega;\omega)$ coming from the $1/\tilde{\omega}$ cancels out. By simple subindex manipulation, the third term, combined with the fourth term in the right-hand side of Eq. (17), gives

$$\frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n} f_n\frac{v^{\Sigma,a}_{nm}v^{\Sigma,b}_{mn}+v^{\Sigma,a}_{mn}v^{\Sigma,b}_{nm}}{\omega^S_{mn}} - \frac{e^2 n_0}{m_e}\delta_{ab} \equiv \zeta^{ab}. \tag{19}$$

The last term on the right-hand side of Eq. (17) reduces to

$$\frac{ie^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_n f_n\mathcal{F}^{ab}_{nn} \equiv \eta^{ab} \tag{20}$$

where

$$\mathcal{F}^{ab}_{nn} = i\Delta\sum_{m(\neq n)} f_{nm}(r^a_{nm}r^b_{mn}+r^b_{nm}r^a_{mn}), \tag{21}$$

summing m over all v and c states different from n (see the appendix). Finally, Eq. (17) reduces to

$$\chi^{ab}(-\omega;\omega) = \frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{v^{\Sigma,a}_{nm}v^{\Sigma,b}_{mn}f_{nm}}{(\omega^S_{mn})^2(\omega^S_{mn}-\tilde{\omega})} + \frac{\zeta^{ab}}{\tilde{\omega}^2} + \frac{\eta^{ab}}{\tilde{\omega}^2}, \tag{22}$$

which is the linear-response coefficient obtained within the velocity gauge, including the scissors correction. Using

$$\mathbf{v}^\Sigma_{nm} = \frac{\omega^S_{nm}}{\omega_{nm}}\mathbf{v}_{nm} \quad (n\neq m), \tag{23}$$

and $\omega^S_{mn}=\omega_{mn}-f_{mn}\Delta$, from the appendix we get that

$$\chi^{ab}(-\omega;\omega) = \frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}v^a_{nm}v^b_{mn}}{\omega^2_{mn}(\omega^S_{mn}-\tilde{\omega})} - \frac{e^2}{\tilde{\omega}^2}\int\frac{d^3k}{8\pi^3}\sum_n f_n\left[\frac{1}{m^*_n}\right]^{ab}, \tag{24}$$

where $[1/m^*_n]^{ab}$ is the effective mass tensor given in Eq. (A12).

For a clean, cold semiconductor $f_n=f_n(\mathbf{k})=1$ or 0, independent of $\mathbf{k}$ and the integration over the Brillouin zone of the term involving the effective-mass tensor vanishes identically,[3] which implies that

$$\chi^{ab}(-\omega;\omega) = \frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}v^a_{nm}v^b_{mn}}{\omega^2_{mn}(\omega^S_{mn}-\omega-i\eta)}, \tag{25}$$

where the energy denominator leads to resonances when $\omega^S_{mn}=\omega$.

A similar calculation neglecting the scissors term in the Hamiltonian leads to

$$\chi^{ab}_{\text{unscissored}}(-\omega;\omega) = \frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}v^a_{nm}v^b_{mn}}{\omega^2_{mn}(\omega_{mn}-\omega-i\eta)},$$

where now the resonances are at $\omega_{mn}=\omega$. A naïve procedure to "scissors" above result would be to take

$$\chi^{ab}_{\text{naïve}}(-\omega;\omega) = \frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}v^{\Sigma,a}_{nm}v^{\Sigma,b}_{mn}}{(\omega^S_{mn})^2(\omega^S_{mn}-\omega-i\eta)}, \tag{26}$$

an incorrect strategy, since it misses the second and third important terms on the right-hand side of Eq. (22). However, using Eq. (23) in Eq. (26) leads by coincidence to the correct result of Eq. (25). It appears that this point has not been appreciated in the literature; Eq. (22) shows the correct way to include the scissors Hamiltonian within the velocity gauge, which is not simply the usual strategy illustrated by Eq. (26).[20]

Using Eq. (A1), we rewrite Eq. (25) as

$$\chi^{ab}(-\omega;\omega)=\frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}r^a_{nm}r^b_{mn}}{\omega^S_{mn}-\omega-i\eta},$$

which is identical to the length gauge result for the scissored Hamiltonian.[9] Again, for the unscissored Hamiltonian one gets,[21]

$$\chi^{ab}_{\text{unscissored}}(-\omega;\omega)=\frac{e^2}{\hbar}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}r^a_{nm}r^b_{mn}}{\omega_{mn}-\omega-i\eta},$$

and as discussed by Nastos *et al.*,[9] in the length gauge the unscissored linear susceptibility can be "scissored" by simply shifting $\omega_{nm}\rightarrow\omega^S_{nm}$ and keeping the same matrix elements $\mathbf{r}_{nm}$. Thus, as in the velocity gauge, the scissored linear response is simply rigidly shifted in energy form its LDA result, keeping the same line shape. We remark that this constitutes a direct analytical proof of gauge invariance for the linear response. For the nonlinear response, we have not been able to construct any such analytical proof. However, we can at least provide a check on the gauge invariance through a numerical calculation. To that end we need expressions for the second-order response in the velocity and length gauges, and we now turn to the first of these.

2. Nonlinear response

Using the results of Sec. II and the previous subsection, we find that to second order in **A** the density matrix is given by

$$\begin{aligned}\tilde{\rho}^{(2)}_{mn}(t)&=-\frac{i}{\hbar}\int_{-\infty}^{t}dt'[\tilde{H}_{I,1}(t'),\tilde{\rho}^{(1)}(t')]_{mn}\\&\quad-\frac{i}{\hbar}\int_{-\infty}^{t}dt'[\tilde{H}_{I,2}(t'),\tilde{\rho}^{(0)}(t')]_{mn}\\&=\frac{e^2}{\hbar^2c^2}\Bigg[\sum_{\ell(\neq n)}\frac{f_{n\ell}v^{\Sigma,b}_{m\ell}v^{\Sigma,c}_{\ell n}}{\omega^S_{\ell n}-\tilde{\omega}}-\sum_{\ell(\neq m)}\frac{f_{\ell m}v^{\Sigma,c}_{m\ell}v^{\Sigma,b}_{\ell n}}{\omega^S_{m\ell}-\tilde{\omega}}\\&\quad+\frac{i}{2}f_{nm}\mathcal{F}^{bc}_{mn}\Bigg]\frac{A^b(\omega)A^c(\omega)}{\omega^S_{mn}-2\tilde{\omega}}e^{-i2\tilde{\omega}t}e^{i\omega^S_{mn}t}\\&=\rho^{(2)}_{mn}e^{i\omega^S_{mn}t},\end{aligned}$$

where, as in the linear response, only the positive frequency terms are used. The $\mathcal{F}^{ab}_{mn}$ term is obtained in the appendix in Eq. (A7). The macroscopic current density can be calculated through Eq. (15), where we take each term separately

$$\begin{aligned}\frac{1}{\Omega}\text{Tr}(j^a_0\rho^{(2)})&=e\int\frac{d^3k}{8\pi^3}\sum_{mn}v^{\Sigma,a}_{nm}\rho^{(2)}_{nm}\\&=\frac{e^3}{\hbar^2c^2}\Bigg[\int\frac{d^3k}{8\pi^3}\sum_{mn}\frac{v^{\Sigma,a}_{nm}}{\omega^S_{mn}-2\tilde{\omega}}\\&\quad\times\left(\sum_{\ell(\neq n)}\frac{f_{n\ell}v^{\Sigma,b}_{m\ell}v^{\Sigma,c}_{\ell n}}{\omega^S_{\ell n}-\tilde{\omega}}-\sum_{\ell(\neq m)}\frac{f_{\ell m}v^{\Sigma,c}_{m\ell}v^{\Sigma,b}_{\ell n}}{\omega^S_{m\ell}-\tilde{\omega}}\right)\\&\quad+\frac{i}{2}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}\frac{f_{nm}v^{\Sigma,a}_{nm}\mathcal{F}^{bc}_{mn}}{\omega^S_{mn}-2\tilde{\omega}}\Bigg]\\&\quad\times A^b(\omega)A^c(\omega)e^{-i2\tilde{\omega}t},\end{aligned}\tag{27}$$

$$\begin{aligned}\frac{1}{\Omega}\text{Tr}(j^a_1\rho^{(1)})&=\frac{e^2}{cm_e}\frac{1}{\Omega}\text{Tr}(\rho^{(1)})A^a(\omega)e^{-i\tilde{\omega}t}\\&\quad+\frac{ie^2}{\hbar c}\frac{1}{\Omega}\text{Tr}(\mathcal{F}^{ab}\rho^{(1)})A^b(\omega)e^{-i\tilde{\omega}t}\\&=\frac{ie^3}{\hbar^2c^2}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}f_{nm}\frac{\mathcal{F}^{ab}_{nm}v^{\Sigma,c}_{mn}}{\omega^S_{mn}-\tilde{\omega}}A^b(\omega)A^c(\omega)e^{-i2\tilde{\omega}t},\end{aligned}\tag{28}$$

since $\text{Tr}(\rho^{(1)})=0$ [see Eq. (16)], and finally

$$\begin{aligned}\frac{1}{\Omega}\text{Tr}(j^a_2\rho^{(0)})&=-\frac{e^3}{2\hbar^2c^2}\int\frac{d^3k}{8\pi^3}\sum_{mn}\rho^{(0)}_{mn}[r^a,(r^b,v^{S,c})]_{nm}\\&\quad\times A^b(\omega)A^c(\omega)e^{-i2\tilde{\omega}t}\\&=-\frac{e^3}{2\hbar^2c^2}\int\frac{d^3k}{8\pi^3}\sum_{n}f_n[r^a,\mathcal{F}^{bc}]_{nn}\\&\quad\times A^b(\omega)A^c(\omega)e^{-i2\tilde{\omega}t}\\&=-\frac{e^3}{2\hbar^2c^2}\int\frac{d^3k}{8\pi^3}\\&\quad\times\sum_n f_n\left(r^a_{nm}\mathcal{F}^{bc}_{mn}-\mathcal{F}^{bc}_{nm}r^a_{mn}+i\frac{\partial}{\partial k^a}\mathcal{F}^{bc}_{e,nn}\right)\\&\quad\times A^b(\omega)A^c(\omega)e^{-i2\tilde{\omega}t},\end{aligned}\tag{29}$$

where we used the expression for $[r^a,\mathcal{F}^{bc}]_{nn}$ derived in the appendix. Again employing time-reversal symmetry, we can take $r^a_{nm}(-\mathbf{k})=r^a_{mn}(\mathbf{k})$, $\mathbf{r}_{mn;\mathbf{k}}(\mathbf{k})=-\mathbf{r}_{nm;\mathbf{k}}(-\mathbf{k})$, $\mathcal{F}^{bc}_{nm}(-\mathbf{k})=\mathcal{F}^{bc}_{mn}(\mathbf{k})$, and $\mathcal{F}^{bc^*}_{nm}(\mathbf{k})=-\mathcal{F}^{bc}_{mn}(\mathbf{k})$. If we add the **k** and the **−k** contributions in Eq. (29), we get a perfect cancellation of the terms within the parenthesis, and so the contribution from $\text{Tr}(j^a_2\rho^{(0)})$ vanishes.

Using $\langle\mathbf{J}\rangle^{(2)}=d\mathbf{P}^{(2)}/dt$ for the second-harmonic response, we get $\mathbf{P}^{(2)}(2\omega)=(i/2\tilde{\omega})\langle\mathbf{J}^{(2)}(2\omega)\rangle$, and from Eq. (5), Eqs. (27) and (28) we find

$$\chi^{abc}(-2\omega;\omega,\omega)=\frac{e^3}{2\hbar^2\tilde{\omega}^3}\Bigg[-i\int\frac{d^3k}{8\pi^3}\sum_{mn}\frac{v_{nm}^{\Sigma,a}\{v_{m\ell}^{\Sigma,b}v_{\ell n}^{\Sigma,c}\}}{\omega_{mn}^S-2\tilde{\omega}}\Bigg(\sum_{\ell(\neq n)}\frac{f_{n\ell}}{\omega_{\ell n}^S-\tilde{\omega}}-\sum_{\ell(\neq m)}\frac{f_{\ell m}}{\omega_{m\ell}^S-\tilde{\omega}}\Bigg)+\frac{1}{2}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}f_{nm}\Bigg(\frac{v_{nm}^{\Sigma,a}\{\mathcal{F}_{mn}^{bc}\}}{\omega_{mn}^S-2\tilde{\omega}}+2\frac{\{\mathcal{F}_{nm}^{ab}v_{mn}^{\Sigma,c}\}}{\omega_{mn}^S-\tilde{\omega}}\Bigg)\Bigg],$$

where $\{\}$ implies the symmetrization of the Cartesian indices bc, i.e., $\{u^b s^c\}=(u^b s^c+u^c s^b)/2$. We take half of this expression and add to it the corresponding expression with $\mathbf{k}$ replaced by $-\mathbf{k}$ in the integrand; this of course give a result equivalent to our first expression, and using time-reversal symmetry we simplify it to yield

$$\chi^{abc}(-2\omega;\omega,\omega)=\frac{e^3}{2\hbar^2\tilde{\omega}^3}\Bigg[\int\frac{d^3k}{8\pi^3}\sum_{(mn)\neq\ell}\frac{\mathrm{Im}[v_{nm}^{\Sigma,a}\{v_{m\ell}^{\Sigma,b}v_{\ell n}^{\Sigma,c}\}]}{\omega_{mn}^S-2\tilde{\omega}}\Bigg(\frac{f_{n\ell}}{\omega_{\ell n}^S-\tilde{\omega}}-\frac{f_{\ell m}}{\omega_{m\ell}^S-\tilde{\omega}}\Bigg)+\frac{1}{2}\int\frac{d^3k}{8\pi^3}\sum_{m\neq n}f_{nm}\Bigg(\frac{\mathrm{Re}[v_{nm}^{\Sigma,a}\{\mathcal{F}_{mn}^{bc}\}]}{\omega_{mn}^S-2\tilde{\omega}}+2\frac{\mathrm{Re}[\{\mathcal{F}_{nm}^{ab}v_{mn}^{\Sigma,c}\}]}{\omega_{mn}^S-\tilde{\omega}}\Bigg)\Bigg]. \tag{30}$$

Following Ghahramani *et al.*,[3] we use partial fractions to write the energy denominator of the first term on the right-hand side of Eq. (30) as

$$\frac{A}{\tilde{\omega}^3}+\frac{B}{\tilde{\omega}^2}+\frac{C}{\tilde{\omega}}+F, \tag{31}$$

where the odd terms in ω, A, and C, can be shown to give zero contribution.[3] For the second term on the right-hand side of Eq. (30) we expand the denominators in partial fractions to obtain

$$\frac{1}{\tilde{\omega}^3(\omega_{mn}^S-\tilde{\omega})}=\frac{1}{\tilde{\omega}(\omega_{mn}^S)^3}+\frac{1}{\tilde{\omega}^3\omega_{mn}^S}+\frac{1}{\tilde{\omega}^2(\omega_{mn}^S)^2}+\frac{1}{(\omega_{mn}^S)^3(\omega_{mn}^S-\tilde{\omega})}, \tag{32}$$

and

$$\frac{1}{\tilde{\omega}^3(\omega_{mn}^S-2\tilde{\omega})}=\frac{4}{\tilde{\omega}(\omega_{mn}^S)^3}+\frac{1}{\tilde{\omega}^3\omega_{mn}^S}+\frac{2}{\tilde{\omega}^2(\omega_{mn}^S)^2}+\frac{8}{(\omega_{mn}^S)^3(\omega_{mn}^S-2\tilde{\omega})}. \tag{33}$$

Using time-reversal symmetry and simple manipulation of the band indices, we can show that all the odd terms in $\tilde{\omega}$ coming from Eqs. (32) and (33) give zero contribution. Collecting the B and F terms of Eq. (31), and the nonzero terms of Eqs. (32) and (33) we obtain

$$\chi^{abc}(-2\omega;\omega,\omega)=\frac{e^3}{2\hbar^2}\int\frac{d^3k}{8\pi^3}\Bigg[\sum_{n(m\neq\ell)}\Bigg(\frac{\mathrm{Im}[v_{mn}^{\Sigma,a}\{v_{n\ell}^{\Sigma,b}v_{\ell m}^{\Sigma,c}\}]}{\omega_{nm}^S-2\omega_{\ell m}^S}-\frac{\mathrm{Im}[v_{n\ell}^{\Sigma,a}\{v_{\ell m}^{\Sigma,b}v_{mn}^{\Sigma,c}\}]}{\omega_{\ell n}^S-2\omega_{\ell m}^S}\Bigg)\frac{f_{m\ell}}{(\omega_{\ell m}^S)^3(\omega_{\ell m}^S-\tilde{\omega})}-16\sum_{\ell(m\neq n)}\frac{f_{mn}}{\omega_{\ell m}^S-2\omega_{nm}^S}\frac{\mathrm{Im}[v_{m\ell}^{\Sigma,a}\{v_{\ell n}^{\Sigma,b}v_{nm}^{\Sigma,c}\}]}{(\omega_{\ell m}^S)^3(\omega_{\ell m}^S-2\tilde{\omega})}-16\sum_{m(\ell\neq n)}\frac{f_{\ell n}}{\omega_{\ell m}^S-2\omega_{\ell n}^S}\frac{\mathrm{Im}[v_{m\ell}^{\Sigma,a}\{v_{\ell n}^{\Sigma,b}v_{nm}^{\Sigma,c}\}]}{(\omega_{\ell m}^S)^3(\omega_{\ell m}^S-2\tilde{\omega})}+\sum_{m\neq n}\frac{f_{nm}}{(\omega_{mn}^S)^3}\Bigg(4\frac{\mathrm{Re}[v_{nm}^{\Sigma,a}\{\mathcal{F}_{mn}^{bc}\}]}{\omega_{mn}^S-2\tilde{\omega}}+\frac{\mathrm{Re}[\{\mathcal{F}_{nm}^{ab}v_{mn}^{\Sigma,c}\}]}{\omega_{mn}^S-\tilde{\omega}}\Bigg)\Bigg] \tag{34}$$

as the nondivergent contribution, to which the divergent term

$$\chi_D^{abc}(-2\omega;\omega,\omega)=\frac{e^3}{2\hbar^2\tilde{\omega}^2}\int\frac{d^3k}{8\pi^3}\Bigg[\sum_{(mn)\neq\ell}b_{\ell mn}\,\mathrm{Im}[v_{nm}^{\Sigma,a}\{v_{m\ell}^{\Sigma,b}v_{\ell n}^{\Sigma,c}\}]+\sum_{m\neq n}\frac{f_{nm}}{(\omega_{mn}^S)^2}(\mathrm{Re}[v_{nm}^{\Sigma,a}\{\mathcal{F}_{mn}^{bc}\}]+\mathrm{Re}[\{\mathcal{F}_{nm}^{ab}v_{mn}^{\Sigma,c}\}])\Bigg]$$

must be added, where

$$b_{\ell mn}=\frac{f_{m\ell}}{\omega_{nm}^S\omega_{\ell m}^S}\Bigg(\frac{2}{\omega_{nm}^S}+\frac{1}{\omega_{\ell m}^S}\Bigg)+\frac{f_{n\ell}}{\omega_{nm}^S\omega_{n\ell}^S}\Bigg(\frac{2}{\omega_{nm}^S}+\frac{1}{\omega_{n\ell}^S}\Bigg),$$

comes from the B term of Eq. (31). Following the steps of Ghahramani *et al.*,[3] we can show that for a clean, cold semiconductor $\chi_D^{abc}(-2\omega;\omega,\omega)=0$.[22]

Finally, we insert the explicit values for the f_n factors and take the limit of $\eta\to 0$ in Eq. (34) to find

$$\mathrm{Im}[\chi_{\mathrm{v}}^{abc}(-2\omega;\omega,\omega)] = \frac{\pi|e|^3}{2\hbar^2}\int\frac{d^3k}{8\pi^3}\Bigg[\sum_{vc}\frac{16}{(\omega_{cv}^S)^3}\left(\sum_{c'}\frac{\mathrm{Im}[v_{vc}^{\Sigma,a}\{v_{cc'}^{\Sigma,b}v_{c'v}^{\Sigma,c}\}]}{\omega_{cv}^S-2\omega_{c'v}^S}-\sum_{v'}\frac{\mathrm{Im}[v_{vc}^{\Sigma,a}\{v_{cv'}^{\Sigma,b}v_{v'v}^{\Sigma,c}\}]}{\omega_{cv}^S-2\omega_{cv'}^S}\right)\delta(\omega_{cv}^S-2\omega)$$
$$+\sum_{(vc)\neq\ell}\frac{1}{(\omega_{cv}^S)^3}\left(\frac{\mathrm{Im}[v_{\ell c}^{\Sigma,a}\{v_{cv}^{\Sigma,b}v_{v\ell}^{\Sigma,c}\}]}{\omega_{c\ell}^S-2\omega_{cv}^S}-\frac{\mathrm{Im}[v_{v\ell}^{\Sigma,a}\{v_{\ell c}^{\Sigma,b}v_{cv}^{\Sigma,c}\}]}{\omega_{\ell v}^S-2\omega_{cv}^S}\right)\delta(\omega_{cv}^S-\omega)$$
$$-\sum_{vc}\frac{1}{(\omega_{cv}^S)^3}(4\,\mathrm{Re}[v_{vc}^{\Sigma,a}\{\mathcal{F}_{cv}^{bc}\}]\delta(\omega_{cv}^S-2\omega)+\mathrm{Re}[\{\mathcal{F}_{vc}^{ab}v_{cv}^{\Sigma,c}\}]\delta(\omega_{cv}^S-\omega))\Bigg], \tag{35}$$

as the imaginary part of the nonlinear SHG susceptibility for the scissored Hamiltonian within the velocity-gauge formalism, where we have used the subscript v to denote it. The $\mathrm{Re}[\chi_{\mathrm{v}}^{abc}(-2\omega;\omega,\omega)]$ is obtained through the Kramers-Kroning transformation. Taking $\Delta=0$ we get

$$\mathrm{Im}[\chi_{\mathrm{v},\Delta=0}^{abc}(-2\omega;\omega,\omega)] = \frac{\pi|e|^3}{2\hbar^2}\int\frac{d^3k}{8\pi^3}\Bigg[\sum_{vc}\frac{16}{(\omega_{cv})^3}\left(\sum_{c'}\frac{\mathrm{Im}[v_{vc}^{a}\{v_{cc'}^{b}v_{c'v}^{c}\}]}{\omega_{cv}-2\omega_{c'v}}-\sum_{v'}\frac{\mathrm{Im}[v_{vc}^{a}\{v_{cv'}^{b}v_{v'v}^{c}\}]}{\omega_{cv}-2\omega_{cv'}}\right)\delta(\omega_{cv}-2\omega)$$
$$+\sum_{(vc)\neq\ell}\frac{1}{(\omega_{cv})^3}\left(\frac{\mathrm{Im}[v_{\ell c}^{a}\{v_{cv}^{b}v_{v\ell}^{c}\}]}{\omega_{c\ell}-2\omega_{cv}}-\frac{\mathrm{Im}[v_{v\ell}^{a}\{v_{\ell c}^{b}v_{cv}^{c}\}]}{\omega_{\ell v}-2\omega_{cv}}\right)\delta(\omega_{cv}-\omega)\Bigg], \tag{36}$$

since $\mathcal{F}_{nm}^{ab}|_{\Delta=0}=0$ (see the appendix). This equation is identical to one obtained earlier for the unscissored Hamiltonian.[3] However, as far as we know the expression for $\mathrm{Im}[\chi^{abc}(-2\omega;\omega,\omega)]$ given in Eq. (35), and the last two terms proportional to Δ through $\mathcal{F}_{nm}^{ab}$ have been neglected in the literature until now. As we show below (Sec. III), these terms are crucial for the gauge invariance of the second-order response within the scissored Hamiltonian.

Indeed, in the past the scissors implementation within the velocity gauge has been performed by taking Eq. (36) and simply replacing ω_{mn} by ω_{mn}^S and $\mathbf{v}_{mn}$ with $\mathbf{v}_{mn}^{\Sigma}$, as the usual scissoring of the linear response would (wrongly) suggest. This strategy leads to

$$\mathrm{Im}[\chi_{\mathrm{v,wrong}}^{abc}(-2\omega;\omega,\omega)] = \frac{\pi|e|^3}{2\hbar^2}\int\frac{d^3k}{8\pi^3}\Bigg[\sum_{vc}\frac{16}{(\omega_{cv}^S)^3}\left(\sum_{c'}\frac{\mathrm{Im}[v_{vc}^{\Sigma,a}\{v_{cc'}^{\Sigma,b}v_{c'v}^{\Sigma,c}\}]}{\omega_{cv}^S-2\omega_{c'v}^S}-\sum_{v'}\frac{\mathrm{Im}[v_{vc}^{\Sigma,a}\{v_{cv'}^{\Sigma,b}v_{v'v}^{\Sigma,c}\}]}{\omega_{cv}^S-2\omega_{cv'}^S}\right)\delta(\omega_{cv}^S-2\omega)$$
$$+\sum_{(vc)\neq\ell}\frac{1}{(\omega_{cv}^S)^3}\left(\frac{\mathrm{Im}[v_{\ell c}^{\Sigma,a}\{v_{cv}^{\Sigma,b}v_{v\ell}^{\Sigma,c}\}]}{\omega_{c\ell}^S-2\omega_{cv}^S}-\frac{\mathrm{Im}[v_{v\ell}^{\Sigma,a}\{v_{\ell c}^{\Sigma,b}v_{cv}^{\Sigma,c}\}]}{\omega_{\ell v}^S-2\omega_{cv}^S}\right)\delta(\omega_{cv}^S-\omega)\Bigg], \tag{37}$$

a wrong result, since we are missing the important contribution from $\mathcal{F}_{mn}^{ab}$ given in Eq. (35). It is obvious that the coincidence that takes place in the linear response does not arise here at all, since if we substitute $\mathbf{v}_{nm}^{\Sigma}=(\omega_{nm}^S/\omega_{nm})\mathbf{v}_{nm}$ in Eq. (37) we do not get the last two terms on the right-hand side of Eq. (35)!

C. Length gauge formalism

Within this gauge, the interaction Hamiltonian is given by

$$H_I(t) = -e\mathbf{r}\cdot\mathbf{E}(t). \tag{38}$$

As discussed in Nastos *et al.*,[9] the length-gauge formalism for the scissored Hamiltonian can be easily worked out by simply using the unscissored Hamiltonian for the unperturbed system with $-e\mathbf{r}\cdot\mathbf{E}(t)$ as the interaction, and then at the end of the calculation only replacing ω_{nm} by ω_{nm}^S to obtain the scissored results for any susceptibility expression, whether linear or nonlinear. Indeed, $\mathbf{r}_{nm}$ and $\mathbf{r}_{nm;\mathbf{k}}$, as stated before, are calculated within the unscissored (LDA) Hamiltonian. We use $H(t)=H_0-e\mathbf{r}\cdot\mathbf{E}(t)$ as the time-dependent Hamiltonian, that from Eq. (4) gives $\dot{\mathbf{r}}=\mathbf{v}=\mathbf{p}/m_e$.

Taking the matrix elements of Eq. (2) but now with the $H_I(t)$ of Eq. (38), we obtain $[\tilde{\rho}_L^{(1)}(t)]_{nm}=B_{nm}^bE^b(\omega)e^{i(\omega_{nm}-\tilde{\omega})t}$, with

$$B_{nm}^b = \frac{e}{\hbar}\frac{f_{mn}r_{nm}^b}{\omega_{nm}-\tilde{\omega}},$$

and

$$[\tilde{\rho}_L^{(2)}(t)]_{nm} = \frac{e}{i\hbar}\frac{1}{\omega_{nm}-2\tilde{\omega}}\left[i\sum_{\ell}(r_{n\ell}^bB_{\ell m}^c-B_{n\ell}^cr_{\ell m}^b) - (B_{nm}^c)_{;k^b}\right]E^b(\omega)E^c(\omega)e^{i(\omega_{nm}-2\tilde{\omega})t}.$$

We have used the fact that for a cold semiconductor $\partial f_n/\partial\mathbf{k}=0$ and thus the intraband contribution to the linear term vanishes identically. From Eq. (3) we can obtain[5]

$$\chi^{abc}_{L,e}(-2\omega;\omega,\omega) = \frac{e^3}{\hbar^2}\int \frac{d^3k}{8\pi^3}\left[\sum_{\ell(n\neq m)} \frac{2f_{nm}}{\omega_{mn}-2\tilde{\omega}} + \sum_{m(\ell\neq n)} \frac{f_{\ell n}}{\omega_{\ell n}-\tilde{\omega}} + \sum_{n(\ell\neq m)} \frac{f_{m\ell}}{\omega_{m\ell}-\tilde{\omega}}\right] \frac{r^a_{nm}\{r^b_{m\ell}r^c_{\ell n}\}}{\omega_{\ell n}-\omega_{m\ell}},$$

as the contribution from interband processes only, and

$$\chi^{abc}_{L,i}(-2\omega;\omega,\omega) = \frac{ie^3}{\hbar^2}\int \frac{d^3k}{8\pi^3}\sum_{m\neq n} f_{nm}\left[\frac{2r^a_{nm}\{r^b_{mn;k^c}\}}{\omega_{mn}(\omega_{mn}-2\tilde{\omega})} + \frac{\{r^a_{nm;k^c}r^b_{mn}\}}{\omega_{mn}(\omega_{mn}-\tilde{\omega})} + \frac{1}{\omega^2_{mn}}\left(\frac{1}{\omega_{mn}-\tilde{\omega}} - \frac{4}{\omega_{mn}-2\tilde{\omega}}\right) r^a_{nm}\{r^b_{mn}\mathcal{V}^c_{mn}\} - \frac{\{r^b_{nm;k^a}r^c_{mn}\}}{2\omega_{mn}(\omega_{mn}-\tilde{\omega})}\right],$$

as the contribution from intraband processes only, where $r^b_{nm;k^a}$ is the generalized derivative of $\mathbf{r}$, and is explicitly given by,[5]

$$r^b_{nm;k^a} = \frac{r^a_{nm}\mathcal{V}^b_{mn} + r^b_{nm}\mathcal{V}^a_{mn}}{\omega_{nm}} + \frac{i}{\omega_{nm}}\sum_{\ell}(\omega_{\ell m}r^a_{n\ell}r^b_{\ell m} - \omega_{n\ell}r^b_{n\ell}r^a_{\ell m}) \quad (n\neq m), \tag{39}$$

where $\mathcal{V}^a_{nm}=(p^a_{nn}-p^a_{mm})/m_e$ is the difference between the electron velocity at bands n and m, and the sum over ℓ is over all the valence and conduction states.

We notice that the above expressions for $\chi^{abc}_{L,e,i}(-2\omega;\omega,\omega)$ are divergence free at $\omega=0$, that both satisfy the intrinsic permutation symmetry $\chi^{abc}_{L,e,i}(-2\omega;\omega,\omega)=\chi^{acb}_{L,e,i}(-2\omega;\omega,\omega)$, and finally that the full susceptibility $\chi^{abc}_{L}(-2\omega;\omega,\omega)=\chi^{abc}_{L,e}(-2\omega;\omega,\omega)+\chi^{abc}_{L,i}(-2\omega;\omega,\omega)$, where the subscript L denotes the length gauge. Again using time-reversal symmetry, we can take $\mathbf{r}_{mn}(\mathbf{k})=\mathbf{r}_{nm}(-\mathbf{k})$ and $\mathbf{r}_{mn;\mathbf{k}}(\mathbf{k})=-\mathbf{r}_{nm;\mathbf{k}}(-\mathbf{k})$, along with the hermiticity condition $\mathbf{r}_{mn}=\mathbf{r}^*_{nm}$, which implies that $\mathbf{r}_{mn;\mathbf{k}}=\mathbf{r}^*_{nm;\mathbf{k}}$, and arrive at the following results for the imaginary parts of $\chi^{abc}_{i,e}$:

$$\text{Im}[\chi^{abc}_{L,e}] = \frac{\pi|e|^3}{\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{(vc)\neq\ell}\left[\frac{2\,\text{Re}[r^a_{vc}\{r^b_{c\ell}r^c_{\ell v}\}]}{\omega^S_{c\ell}-\omega^S_{\ell v}}\delta(\omega^S_{cv}-2\omega) + \left(\frac{\text{Re}[r^a_{v\ell}\{r^b_{\ell c}r^c_{cv}\}]}{\omega^S_{cv}-\omega^S_{\ell c}} + \frac{\text{Re}[r^a_{\ell c}\{r^b_{cv}r^c_{v\ell}\}]}{\omega^S_{v\ell}-\omega^S_{cv}}\right)\delta(\omega^S_{cv}-\omega)\right] \tag{40}$$

and

$$\text{Im}[\chi^{abc}_{L,i}] = \frac{\pi|e|^3}{\hbar^2}\int\frac{d^3k}{8\pi^3}\sum_{vc}\left[\left(\frac{2\,\text{Im}[r^a_{vc}\{r^b_{cv;k^c}\}]}{\omega^S_{cv}} - \frac{4\,\text{Im}[r^a_{vc}\{r^b_{cv}\mathcal{V}^c_{cv}\}]}{(\omega^S)^2_{cv}}\right)\delta(\omega^S_{cv}-2\omega) + \left(\frac{\text{Im}[\{r^a_{vc;k^c}r^b_{cv}\}]}{\omega^S_{cv}} + \frac{\text{Im}[r^a_{vc}\{r^b_{cv}\mathcal{V}^c_{cv}\}]}{(\omega^S)^2_{cv}} - \frac{\text{Im}[\{r^b_{vc;k^a}r^c_{cv}\}]}{2\omega^S_{cv}}\right)\delta(\omega^S_{cv}-\omega)\right], \tag{41}$$

where we have taken $\omega_{nm}\to\omega^S_{nm}$ so the above expressions are valid for the scissors Hamiltonian, $H^S=H_0+S(\mathbf{r},\mathbf{p})-e\mathbf{r}\cdot\mathbf{E}$. Recall that both $\mathbf{r}_{nm}$ and $\mathbf{r}_{nm;\mathbf{k}}$ are calculated with the unscissored (Kohn-Sham) Hamiltonian.[9] The last two equations give the nonlinear SHG susceptibility within the length gauge for the scissored Hamiltonian. Comparing Eqs. (40) and (41) with the velocity gauge result of Eq. (35) it is clear that, unlike for linear response, there is no obvious analytical scheme to prove that both gauges give the same result. In the following section we present numerical results to prove the expected gauge invariance.

III. RESULTS

In this section we evaluate the velocity- and length-gauge expressions $\chi^{abc}_{v,L}(-2\omega;\omega,\omega)$ for GaAs. In order to calculate the energies, wave functions and matrix elements we employ the "augmented plane wave plus local orbital method" using the WIEN2K code.[23] This all-electron code uses the full local-crystal potential, i.e., $V(\mathbf{r})$, just as required by our assumptions of Eq. (6), and thus the commutator $[\mathbf{r},H]=i\hbar\dot{\mathbf{r}}$ is correctly calculated for the local $V(\mathbf{r})$. We also show results calculated through the use of pseudopotentials with the ABINIT plane-wave code.[24] The pseudopotentials are nonlocal functions expressible as $V^{\text{NL}}(\mathbf{r},\mathbf{p})$,[25] just as is the scissor Hamiltonian $S(\mathbf{r},\mathbf{p})$. To really complete the calculation one would have to do the corresponding manipulations including $V^{\text{NL}}(\mathbf{r},\mathbf{p})$ in the Hamiltonian H^S [Eq. (9)], and terms would arise in the linear and nonlinear susceptibility expressions. For instance, the term $i\hbar\mathbf{v}^{\text{NL}}=[\mathbf{r},V^{\text{NL}}(\mathbf{r},\mathbf{p})]$ should be added to the velocity operator $\mathbf{v}^{\Sigma}$ given by Eq. (13). This is a research project for the future. Here we use the comparison of the all-electron and the pseudopotential calculation to get a sense of the size of error involved by neglecting the nonlocal contributions coming from $V^{\text{NL}}(\mathbf{r},\mathbf{p})$. The linear-response counterpart of these calculations are discussed in Pulci *et al.*[26] and Mendoza *et al.*[27]

Spin-orbit effects, local field effects, and the consequences of the electron-hole attraction[8] on the SHG process are neglected. Although all these effects are important for the optical response of a semiconductor, their calculation is still an open question and a numerical challenge that ought to be pursued. However this endeavor is beyond the scope of this paper. The band gap of GaAs is taken to be its experimental value of 1.52 eV. We find converged spectra for all the quantities of interest in this work, and the most important param-

TABLE I. The most important parameters used in the all-electron and pseudopotential schemes for GaAs. The empty entries are not relevant for the corresponding code. The **k** points are for the irreducible part of the first Brillouin zone, and $R_{\mathrm{MT}}K_{\mathrm{MAX}}$ is a product of the "muffin-tin" radius R and the maximum value for the plane-wave vectors **K**.[23]

Parameter (GaAs)	All-electron	Pseudopotential
Lattice parameter	$10.684a_0$	$10.684a_0$
k points	27720	27720
Unscissored band gap	0.277 eV	0.469 eV
Scissors	1.243 eV	1.051 eV
Valence bands	14 (includes semicore)	4
Conduction bands	7	7
Exchange correlation energy	LDA	LDA
Energy convergence limit	0.001 Ry	
Cut-off energy		20 Ha
$R_{\mathrm{MT}}K_{\mathrm{MAX}}$	7.0	

eters are shown in the Table I. All the spectra are calculated with an energy smearing of 0.15 eV. The linear analytic tetrahedron method is used to evaluate the Brillouin zone integrals for the imaginary part of the spectra, where special care was taken to examine the double resonances.[9] Double resonances occur if for a given frequency ω there can be resonant transitions at both frequencies ω and 2ω, that is, if there is a region in the Brillouin zone such that $\omega_{cv}(\mathbf{k})=2\omega_{c'v}(\mathbf{k})=2\omega$. For these **k** points the perturbation theory used to calculate the spectrum breaks down, since there is real population excited, which in a correct calculation must be taken into account. These points introduce sharp spikes in the spectrum that can, in principle, affect the low-frequency results, since the response at frequencies below the band gap is computed from the Kramers-Kronig relation. However, in agreement with Nastos *et al.*,[9] we find here that the double resonances affect the low-frequency results by less than 2%.

In Fig. 1 we show the imaginary part of $\chi^{xyz}_{\mathrm{v},L}(-2\omega;\omega,\omega)$ with no scissors correction ($\Delta=0$), calculated with the all-electron scheme.[28] As expected, $\mathrm{Im}[\chi^{xyz}_{\mathrm{v},L}](-2\omega;\omega,\omega)$ is zero below the gap, and above it we see a series of positive and negative peaks that can be related to electronic transitions. What is more relevant for this paper is that in the top panel we have plotted both $\chi^{xyz}_{\mathrm{v}}(-2\omega;\omega,\omega)$ and $\chi^{xyz}_{L}(-2\omega;\omega,\omega)$; they seem identical, as they must be, since gauge invariance must be fulfilled. In the bottom panel of Fig. 1, we show $\mathrm{Im}[\chi^{xyz}_{L}(-2\omega;\omega,\omega)-\chi^{xyz}_{\mathrm{v}}(-2\omega;\omega,\omega)]$, which confirms that the results for the unscissored $\mathrm{Im}[\chi^{abc}_{\mathrm{v},L}(-2\omega;\omega,\omega)]$ agree to within numerical accuracy (about 1 part in approximately 10^5), as would be expected from gauge invariance.

In Fig. 2 we show the imaginary part of $\chi^{xyz}_{\mathrm{v},L}(-2\omega;\omega,\omega)$ with a scissors shift of $\Delta=1.243$ eV, calculated with the all-electron scheme. In the top panel we compare the velocity-gauge calculation [i.e., $\mathrm{Im}[\chi^{xyz}_{\mathrm{v}}(-2\omega;\omega,\omega)]$ of Eq. (35)] with a calculation where we neglect the contributions coming from the scissors term, [i.e., $\mathrm{Im}[\chi^{xyz}_{\mathrm{v,wrong}}(-2\omega;\omega,\omega)]$ of Eq. (37)]; we see that the results disagree. In the middle panel we show $\mathrm{Im}[\chi^{xyz}_{L}(-2\omega;\omega,\omega)]$

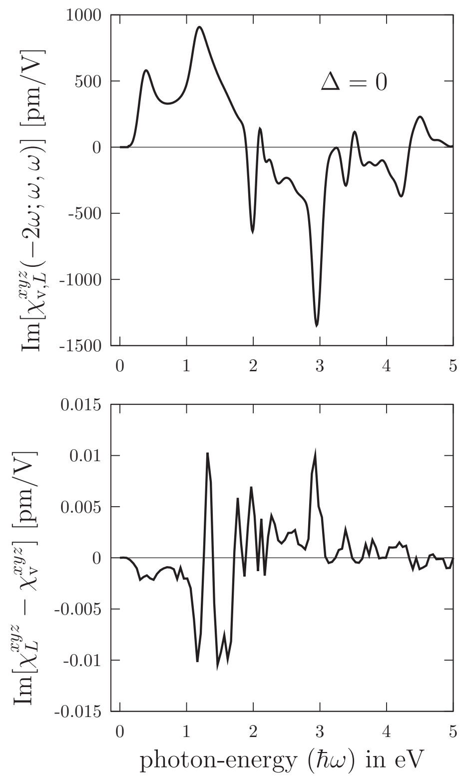

FIG. 1. (a) $\mathrm{Im}[\chi^{xyz}_{\mathrm{v},L}(-2\omega;\omega,\omega)]$ for the length-gauge and the velocity-gauge schemes, using the all-electron approach and for zero scissors correction, $\Delta=0$. (b) $\mathrm{Im}[\chi^{xyz}_{L}(-2\omega;\omega,\omega)-\chi^{xyz}_{\mathrm{v}}(-2\omega;\omega,\omega)]$ where very tiny differences between the two schemes are seen.

and in the bottom panel we show $\mathrm{Im}[\chi^{xyz}_{L}(-2\omega;\omega,\omega)-\chi^{xyz}_{\mathrm{v}}(-2\omega;\omega,\omega)]$, where it is clear that, as in the unscissored case, gauge invariance with the scissored Hamiltonian is confirmed within numerical accuracy. We stress that this fulfillment of gauge invariance is due to the terms of Eq. (35) proportional to $\mathcal{F}^{ab}_{mn}$ which in turn depends on the commutator $[\mathbf{r},\mathbf{v}^S]$ with $\mathbf{v}^S=-(i/\hbar)[\mathbf{r},S(\mathbf{r},\mathbf{p})]$. Thus, neglecting the effect of the scissors operator $S(\mathbf{r},\mathbf{p})$ in the usual perturbation procedure would lead, in general, to the wrong result for nonlinear susceptibility tensors within the velocity-gauge approach.

As explained above, we have also used a pseudopotential method to calculate the SHG susceptibility tensor. In this way, we can estimate the error that one makes when calculating the matrix elements of the electron's momentum operator through the use of pseudopotentials, the error arising from the nonlocal part of the pseudopotential in the commutators.[26,27] In Fig. 3 we show the absolute value of $|\chi^{xyz}_{L}(-2\omega;\omega,\omega)|=|\chi^{xyz}_{\mathrm{v}}(-2\omega;\omega,\omega)|\equiv|\chi^{xyz}(-2\omega;\omega,\omega)|$ with the scissors correction, where $\Delta=1.051$ eV for the pseudopotential code and $\Delta=1.243$ eV for the all-electron code. We notice that there is a difference between the results of the value of the static limit of $|\chi^{xyz}(-2\omega;\omega,\omega)|$ by approximately 36.8 pm/V; we obtain a static value of $|\chi^{xyz}(0;0,0)|=135.6$ pm/V for the pseudopotential calculation, and $|\chi^{xyz}(0;0,0)|=172.4$ pm/V, for the all-electron calculation. These quantities are close to the theoretical values of other studies,[9] and to the most recent experimental value close to

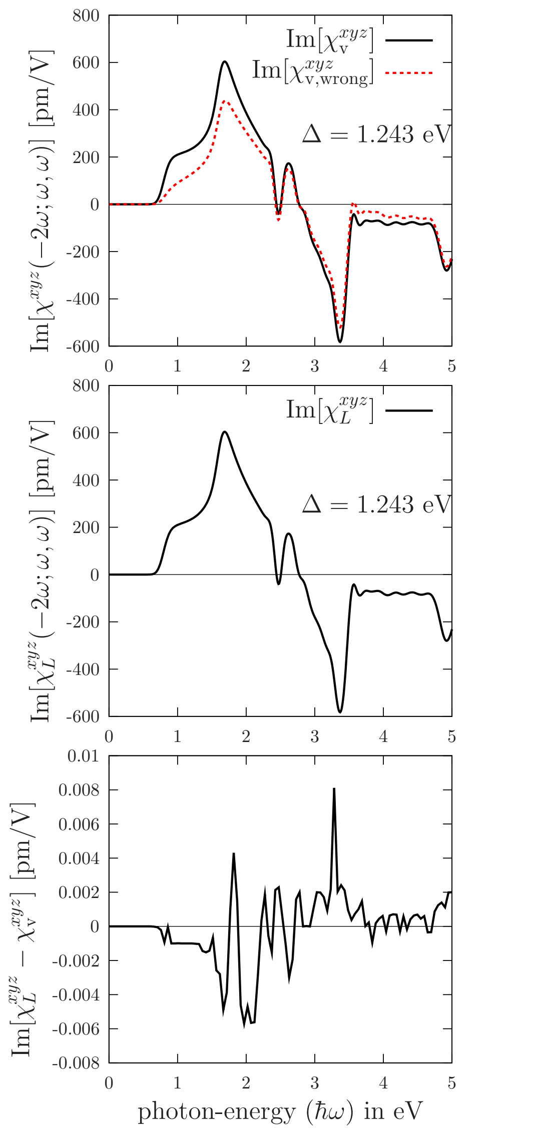

FIG. 2. (Color online) (a) Im[$\chi_v^{xyz}(-2\omega;\omega,\omega)$] and Im[$\chi_{v,\mathrm{wrong}}^{xyz}(-2\omega;\omega,\omega)$] for the velocity gauge. (b) Im[$\chi_L^{xyz}(-2\omega;\omega,\omega)$] for the longitudinal gauge. (c) Im[$\chi_L^{xyz}(-2\omega;\omega,\omega)-\chi_v^{xyz}(-2\omega;\omega,\omega)$] where very tiny differences are seen. The spectra is evaluated within the all-electron approach with $\Delta = 1.243$ eV.

the static limit of 172 pm/V at 0.118 eV.[29] We see that the corrections due to the nonlocal nature of the pseudopotentials affect not only the strength of the spectrum but also its line shape, as some resonances are energy shifted from one calculation to the other. The overall intensity correction is smaller than ~25%, and we may conclude that the pseudopotential calculation does a reasonable job for the nonlinear response. Indeed, this seems to be the case for the linear optical response as well.[27]

Although the main objective of this paper is to show how the nonlocal scissors correction must be included in the lineal and nonlineal optical responses, and how including it fulfills gauge invariance, as shown in Fig. 2, we present for reference the comparison of the theoretical results with the experimental results. In Fig. 4 we show the experimental spectrum measured by Bergfeld and Daum,[30] where in order to have a better comparison of theory and experiment, the energy scale of the theoretical results has been linearly rescaled as proposed by them.[31] Our results for the all-electron calculation show good agreement with the experimental values up to 4.3 eV. Above 4.3 eV the theoretical signal disagrees although it shows a similar line shape that is blueshifted in energy with respect to the experimental signal. We have checked that the results obtained in Refs. 6–10 qualitatively show a similar comparison with the experiment.

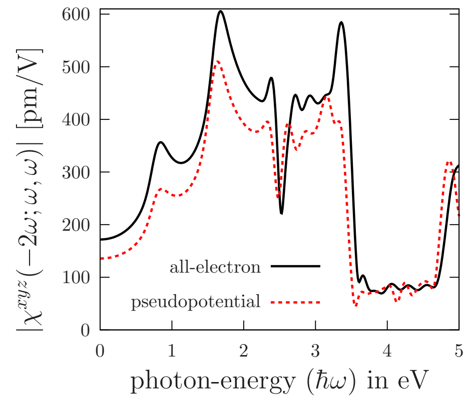

FIG. 3. (Color online) $|\chi_L^{xyz}(-2\omega;\omega,\omega)| = |\chi_v^{xyz}(-2\omega;\omega,\omega)| \equiv |\chi^{xyz}(-2\omega;\omega,\omega)|$ with a scissors correction of $\Delta = 1.243$ eV for the all-electron calculation and $\Delta = 1.051$ eV for the pseudopotential calculation.

IV. CONCLUSIONS

We have presented a comparison for the calculation of the second-harmonic susceptibility tensor using two well-known approaches, often colloquially referred to as using the velocity gauge and the "length gauge." We have done this for two Hamiltonians, the usual LDA Hamiltonian and the scissored Hamiltonian, where a rigid energy shift in the conduction

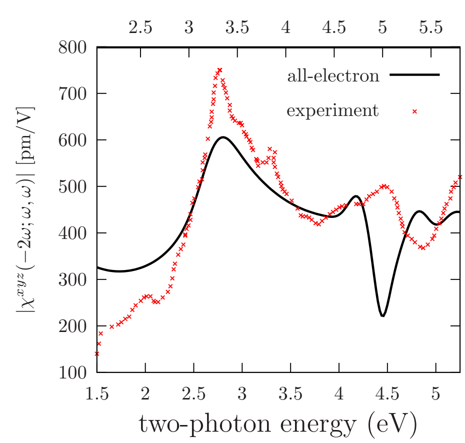

FIG. 4. (Color online) $|\chi_L^{xyz}| = |\chi_v^{xyz}| = \chi^{xyz}$ with a scissors correction of $\Delta = 1.243$ eV for the all-electron calculation, along with the experimental results of Ref. 30. The top energy scale is the original energy E_{orig} of the all-electron calculation. The bottom scale is the scaled energy, $2E_{\mathrm{mod}}$ (Ref. 31), for the theoretical spectra and the two-photon energy of the experimental results.

bands is introduced so the experimental (or GW) energy gap is obtained. We derived a expression for the velocity-gauge susceptibility $\chi_{\rm v}^{abc}(-2\omega;\omega,\omega)$, where correction terms related to the nonlocal nature of the scissors operator were obtained. These terms, not considered before in the literature, are crucial in order to obtain gauge invariance for a calculation made with the scissored Hamiltonian. For the unscissored Hamiltonian, gauge invariance is obtained with the usual $\chi^{abc}(-2\omega;\omega,\omega)$ expression for the velocity and length gauges.

We have presented our numerical results for GaAs using a DFT-LDA *ab initio* calculation, with the augmented plane wave plus local orbital all-electron method as given by WIEN2K,[23] and a plane-wave pseudopotential scheme given by the ABINIT code.[24] Besides providing a numerical demonstration of gauge invariance for the unscissored and the scissored Hamiltonian calculations, this indicates the kind and size of error that the neglect of the nonlocal nature of the pseudopotentials can be expected to produce in the calculation of $\chi_{{\rm v},L}^{abc}(-2\omega;\omega,\omega)$; it affects not only the strength of the spectrum but also its line shape. Our results compare qualitatively well with the previous work of other authors, and, in particular, with the experimental results. However, the details of each approach show that the calculation of the nonlinear response in a nontrivial matter, and better calculations of $\chi^{abc}(-2\omega;\omega,\omega)$ using more sophisticated means are still to be sought.

ACKNOWLEDGMENTS

J.L.C. and M.A.E. acknowledge partial support through CONACyT. B.S.M. acknowledges support by CONACYT under Grant No. 48915-F. F.N. acknowledges support by the Ontario Graduate Scholarship program. J.E.S. acknowledges the National Science and Engineering Research Council of Canada and Photonics Research Ontario.

APPENDIX

In this appendix we derive several results related to the scissors operator $S(\mathbf{r},\mathbf{p})$ of Eq. (7). First we sketch some well-known results, for which we follow Aversa and Sipe,[5] and Blount.[32] We write the position operator of the electron, $\mathbf{r}$, as the sum of its *interband* part $\mathbf{r}_e$ and *intraband* part $\mathbf{r}_i$, $\mathbf{r}=\mathbf{r}_e+\mathbf{r}_i$. The matrix elements of $\mathbf{r}_e$ are simply given by[5]

$$\langle n\mathbf{k}|\mathbf{r}_e|m\mathbf{k}'\rangle = \delta(\mathbf{k}-\mathbf{k}')(\mathbf{r}_e)_{nm} \to \mathbf{r}_{nm} = \frac{\mathbf{p}_{nm}}{im_e\omega_{nm}} = \frac{\mathbf{v}_{nm}}{i\omega_{nm}} \quad n\neq m, \tag{A1}$$

where the canonical momentum matrix elements are calculated according to

$$\langle n\mathbf{k}|\mathbf{p}|m\mathbf{k}'\rangle = \delta(\mathbf{k}-\mathbf{k}')\mathbf{p}_{nm} = \delta(\mathbf{k}-\mathbf{k}')\int d^3r\,\psi^*_{n\mathbf{k}}(\mathbf{r})(-i\hbar\boldsymbol{\nabla})\psi_{m\mathbf{k}}(\mathbf{r}),$$

and $\mathbf{v}_{nm}=\mathbf{p}_{nm}/m_e$. Instead of needing the matrix elements of $\mathbf{r}_i$ one actually uses its following property:[5]

$$\langle n\mathbf{k}|[\mathbf{r}_i,\mathcal{O}]|m\mathbf{k}'\rangle = i\delta(\mathbf{k}-\mathbf{k}')(\mathcal{O}_{nm})_{;\mathbf{k}}, \tag{A2}$$

where $\mathcal{O}$ is an operator and $(\mathcal{O}_{nm})_{;\mathbf{k}}$ is the generalized derivative of its matrix elements, i.e., Eq. (39) for $r^b_{nm;k^a}$. As discussed by Nastos *et al.*,[9] both $\mathbf{r}_{nm}$ [Eq. (A1)], and its generalized derivative $\mathbf{r}_{nm;\mathbf{k}}$ [Eq. (39)], are evaluated using the unscissored energies.

Now we establish Eq. (23). We take matrix elements of Eq. (12) and use Eq. (7) to write

$$\begin{aligned}\mathbf{v}^S_{nm} &= -\frac{i}{\hbar}\langle n\mathbf{k}|[\mathbf{r}S(\mathbf{r},\mathbf{p})-S(\mathbf{r},\mathbf{p})\mathbf{r}]|m\mathbf{k}\rangle\\ &= -i\Delta[(1-f_m)-(1-f_n)]\langle n\mathbf{k}|\mathbf{r}|m\mathbf{k}\rangle\\ &= i\Delta f_{mn}\mathbf{r}_{nm} = \frac{\Delta f_{mn}}{m_e\omega_{nm}}\mathbf{p}_{nm},\end{aligned} \tag{A3}$$

where we used Eq. (A1) since the factor, f_{mn}, yields $n\neq m$. Then the matrix elements of Eq. (13) reduce to

$$\begin{aligned}\mathbf{v}^\Sigma_{nm} &= \left(1+\frac{\Delta f_{mn}}{\omega_{nm}}\right)\langle n\mathbf{k}|\frac{\mathbf{p}}{m_e}|m\mathbf{k}\rangle = \left(\frac{\omega_{nm}+\Delta f_{mn}}{\omega_{nm}}\right)\mathbf{v}_{nm}\\ &= \left(\frac{\omega^S_n-\omega^S_m}{\omega_{nm}}\right)\mathbf{v}_{nm} = \frac{\omega^S_{nm}}{\omega_{nm}}\mathbf{v}_{nm} \quad (n\neq m),\end{aligned} \tag{A4}$$

where we used Eq. (8); thus Eq. (A4) is Eq. (23).

In order to prove Eq. (21), we start with the matrix elements of Eq. (18), which we write as

$$\mathcal{F}^{ab}_{nm} = \langle n\mathbf{k}|([r^a_i,v^{S,b}]+[r^a_e,v^{S,b}])|m\mathbf{k}\rangle.$$

The interband part is

$$\begin{aligned}\langle n\mathbf{k}|[r^a_e,v^{S,b}]|m\mathbf{k}\rangle \equiv \mathcal{F}^{ab}_{e,nm} &= \sum_\ell (r^a_{e,n\ell}v^{S,b}_{\ell m}-v^{S,b}_{n\ell}r^a_{e,\ell m})\\ &= i\Delta\sum_{\ell\neq(mn)}(f_{m\ell}r^a_{n\ell}r^b_{\ell m}-f_{\ell n}r^b_{n\ell}r^a_{\ell m}),\end{aligned} \tag{A5}$$

where we used Eqs. (A1) and (A3). For the intraband part we use the result of Eqs. (A2) and (A3) to simply write

$$\langle n\mathbf{k}|[r^a_i,v^{S,b}]|m\mathbf{k}\rangle \equiv \mathcal{F}^{ab}_{i,nm} = iv^{S,b}_{nm;k^a} = \Delta f_{nm}r^b_{nm;k^a}. \tag{A6}$$

From Eqs. (A5) and (A6) we find

$$\mathcal{F}^{ab}_{nm} = i\Delta\sum_{\ell\neq(mn)}(f_{m\ell}r^a_{n\ell}r^b_{\ell m}-f_{\ell n}r^b_{n\ell}r^a_{\ell m})+\Delta f_{nm}r^b_{nm;k^a}. \tag{A7}$$

We see that for $n=m$ the intraband contribution $\mathcal{F}^{ab}_{i,nn}=0$, whereas the interband part reduces to

$$\mathcal{F}^{ab}_{nn} = \mathcal{F}^{ab}_{e,nn} = i\Delta\sum_{m\neq n}f_{nm}(r^a_{nm}r^b_{mn}+r^b_{nm}r^a_{mn}), \tag{A8}$$

giving Eq. (21).

Now we take matrix elements of $[r^a, \mathcal{F}^{bc}]$, separating $r^a = r_i^a + r_e^a$. Then the interband part gives

$$[r_e^a, \mathcal{F}^{bc}]_{nn} = \sum_{m\neq n} (r_{nm}^a \mathcal{F}_{mn}^{bc} - \mathcal{F}_{nm}^{bc} r_{mn}^a), \tag{A9}$$

while the intraband part gives

$$[r_i^a, \mathcal{F}^{bc}]_{nn} = i\mathcal{F}_{nn;a}^{bc} = i\frac{\partial}{\partial k^a}\mathcal{F}_{nn}^{bc} = i\frac{\partial}{\partial k^a}\mathcal{F}_{e,nn}^{bc}, \tag{A10}$$

where we used Eq. (A8). Then Eqs. (A9) and (A10) are used to obtain Eq. (29).

We derive Eq. (24). Using Eq. (A4), $\omega_{mn}^S = \omega_{mn} - f_{mn}\Delta$, Eqs. (21) and (A1), Eq. (19) reduces to

$$\begin{aligned}
\zeta^{ab} &= \frac{e^2}{\hbar}\int \frac{d^3k}{8\pi^3}\sum_{m\neq n} f_n \omega_{mn}^S \frac{v_{nm}^a v_{mn}^b + v_{mn}^a v_{nm}^b}{\omega_{mn}^2} - \frac{e^2 n}{m}\delta_{ab} \\
&= \frac{e^2}{\hbar}\int \frac{d^3k}{8\pi^3}\sum_{m\neq n} f_n \omega_{mn} \frac{v_{nm}^a v_{mn}^b + v_{mn}^a v_{nm}^b}{\omega_{mn}^2} - \frac{e^2 n}{m}\delta_{ab} - \frac{e^2\Delta}{\hbar}\int \frac{d^3k}{8\pi^3}\sum_{m\neq n} f_n f_{mn} \frac{v_{nm}^a v_{mn}^b + v_{mn}^a v_{nm}^b}{\omega_{mn}^2} \\
&= -e^2 \int \frac{d^3k}{8\pi^3}\sum_n f_n \left(\frac{\delta_{ab}}{m} - \sum_{m\neq n}\frac{v_{nm}^a v_{mn}^b + v_{mn}^a v_{nm}^b}{\hbar\omega_{mn}}\right) + \frac{e^2\Delta}{\hbar}\int \frac{d^3k}{8\pi^3}\sum_n f_n \sum_{m\neq n} f_{nm}(r_{nm}^a r_{mn}^b + r_{mn}^a r_{nm}^b) \\
&= -e^2 \int \frac{d^3k}{8\pi^3}\sum_n f_n \left[\frac{1}{m_n^*}\right]^{ab} - \frac{ie^2}{\hbar}\int \frac{d^3k}{8\pi^3}\sum_n f_n \mathcal{F}_{nn}^{ab},
\end{aligned} \tag{A11}$$

where

$$\left[\frac{1}{m_n^*}\right]^{ab} = \frac{\delta_{ab}}{m_e} - \sum_{m\neq n}\frac{v_{nm}^a v_{mn}^b + v_{mn}^a v_{nm}^b}{\hbar\omega_{mn}} \tag{A12}$$

is the effective mass tensor. Identifying the second term on the right-hand side of Eq. (A11) as $-\eta^{ab}$ [see Eq. (20)], leads to

$$\zeta^{ab} + \eta^{ab} = -e^2\int \frac{d^3k}{8\pi^3}\sum_n f_n \left[\frac{1}{m_n^*}\right]^{ab}.$$

Using above in Eq. (22) gives Eq. (24).

*bms@cio.mx

[1] P. N. Butcher and T. P. McLean, Proc. Phys. Soc. London **81**, 219 (1963).
[2] D. E. Aspnes, Phys. Rev. B **6**, 4648 (1972).
[3] E. Ghahramani, D. J. Moss, and J. E. Sipe, Phys. Rev. B **43**, 8990 (1991).
[4] Z. H. Levine, Phys. Rev. B **49**, 4532 (1994).
[5] C. Aversa and J. E. Sipe, Phys. Rev. B **52**, 14636 (1995).
[6] S. N. Rashkeev, W. R. L. Lambrecht, and B. Segall, Phys. Rev. B **57**, 3905 (1998).
[7] J. L. P. Hughes and J. E. Sipe, Phys. Rev. B **53**, 10751 (1996).
[8] R. Leitsmann, W. G. Schmidt, P. H. Hahn, and F. Bechstedt, Phys. Rev. B **71**, 195209 (2005).
[9] F. Nastos, B. Olejnik, K. Schwarz, and J. E. Sipe, Phys. Rev. B **72**, 045223 (2005).
[10] B. Adolph and F. Bechstedt, Phys. Rev. B **57**, 6519 (1998).
[11] B. S. Mendoza, M. Palummo, G. Onida, and R. Del Sole, Phys. Rev. B **63**, 205406 (2001).
[12] B. S. Mendoza, A. Gaggiotti, and R. Del Sole, Phys. Rev. Lett. **81**, 3781 (1998).
[13] L. Reining, R. Del Sole, M. Cini, and J. G. Ping, Phys. Rev. B **50**, 8411 (1994).
[14] S. Sharma, J. K. Dewhurst, and C. Ambrosch-Draxl, Phys. Rev. B **67**, 165332 (2003).
[15] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[16] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[17] R. Del Sole and R. Girlanda, Phys. Rev. B **48**, 11789 (1993).
[18] Z. H. Levine and D. C. Allan, Phys. Rev. B **44**, 12781 (1991).
[19] Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).
[20] For instance, in M. Alouani and J. M. Wills, Phys. Rev. B **54**, 2480 (1996) it is found that the nonsimultaneous fulfilment of both the *f*-sum rule and the sum rule satisfied by the imaginary part of the linear dielectric function is an indication of the limitation for the scissors Hamiltonian. Indeed they concluded that the scissors correction well describes the low-lying excited states that include the E_1 and E_2 bulk critical points, and that the higher excited states are better described by the LDA Hamiltonian alone. However, they used the usual procedure to scissors the linear response. In view of the correct way of including the scissored Hamiltonian, the sum rules for both the linear and the nonlinear response ought to be reconsidered to understand their possible implications in the optical response.

[21] J. E. Sipe and A. I. Shkrebtii, Phys. Rev. B **61**, 5337 (2000).

[22] J. L. Cabellos, Ph.D. thesis, Centro de Investigaciones en Óptica, 2009.

[23] 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 Schwartz, Techn, Universität Wien, Austria, 2001).

[24] X. Gonze, J. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J. -Y. Raty, and D. C. Allan, Comput. Mater. Sci. **25**, 478 (2002).

[25] A. F. Starace, Phys. Rev. A **3**, 1242 (1971).

[26] O. Pulci, G. Onida, R. Del Sole, and A. J. Shkrebtii, Phys. Rev. B **58**, 1922 (1998).

[27] B. S. Mendoza, F. Nastos, N. Arzate, and J. E. Sipe, Phys. Rev. B **74**, 075318 (2006).

[28] The crystal class of GaAs, F-43m, gives only one independent component for the susceptibility tensor $\chi^{abc}(-2\omega;\omega,\omega)$, i.e., $\chi^{xyz}(-2\omega;\omega,\omega)=\chi^{zxy}(-2\omega;\omega,\omega)=\chi^{yzx}(-2\omega;\omega,\omega)$ plus the intrinsic permutation symmetry $\chi^{abc}(-2\omega;\omega,\omega)=\chi^{acb}(-2\omega;\omega,\omega)$.

[29] L. A. Eyres, P. J. Tourreau, T. J. Pinguet, C. B. Ebert, J. S. Harris, M. M. Fejer, L. Becouarn, B. Gerard, and E. Lallier, Appl. Phys. Lett. **79**, 904 (2001).

[30] S. Bergfeld and W. Daum, Phys. Rev. Lett. **90**, 036801 (2003).

[31] We have taken $E_{\mathrm{mod}}=a+bE_{\mathrm{orig}}$, where a and b are fitted such that the theoretical values of the bulk critical points E_1 and E_2 of the linear response (not shown in this paper) coincide with the experimental values $E_1=2.91$ eV and $E_2=4.96$ eV (Ref. 30).

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