PHYSICAL REVIEW B VOLUME 49, NUMBER 7 15 FEBRUARY 1994-I

Optical second-harmonic susceptibilities: Frequency-dependent formulation with results for GaP and GaAs

Zachary H. Levine
Department of Physics, The Ohio State University, Columbus, Ohio 43210-1106

(Received 3 August 1993)

A formula for frequency-dependent optical second-harmonic susceptibilities for insulating crystals, including local-field corrections, is presented. In analogy to a formula given previously for the static limit, the formula is analyzed in terms of "virtual electron" (*ccv*) and "virtual hole" (*vvc*) terms to obtain a numerically stable form. Numerical results for GaP and GaAs agree with most of the experimental values suggested in a recent review article by Roberts. Miller's Δ_M is predicted to vary by 13% from the laser lines at 10.6 to 1.32 μm for GaP.

I. INTRODUCTION

Technical difficulties delayed the accurate calculation of second-harmonic susceptibilities $\chi^{(2)}$ for more than two decades. The initial discovery[1] of second-harmonic generation lead to rapid advances in the theory (see, e.g., Ref. 2). For example, the present understanding of the classical electrodynamics of second-harmonic generation dates to this period. Bond-charge models were used to predict second-harmonic susceptibilities, most notably by Levine.[3] Although band-theoretic formulas were derived, they were too complicated and delicate for their application at the time. One problem concerned apparent divergences in the theory.[4,5] Another problem concerned uncertainties regarding local-field corrections.[2,6] Yet another difficulty stemmed from insufficiently accurate matrix elements.[6,7]

Recent research has seen considerable progress on some of these classic problems. For example, a proof that the second-harmonic susceptibility of an insulator to a vector potential is finite in the static limit has been given recently.[8] A band-theoretic formalism with local-field corrections within the vector theory has been presented recently[9] (recapitulating, in part, earlier work[10]), but with an attempt at a realistic numerical evaluation.[11] Recently, the second-harmonic susceptibility was evaluated for some 15 III-V and II-VI semiconductors using the self-consistent orthogonalized-linear-combination of atomic-orbitals method (OLCAO);[12] this reference contains an excellent summary of the theoretical situation in its introduction.

In Ref. 13, I presented a formula within band theory for optical second-harmonic response to an external scalar potential for a crystal, including local-field corrections. The use of a scalar potential simplifies the formalism compared to a vector theory; however, it limits the calculation to the Kleinman-allowed symmetries.[13,14] Fortunately, these are the largest ones.[15] Previously, we evaluated $\chi^{(2)}$ in the static limit for GaAs, GaP, AlAs, AlP,[16,17] selenium,[18,19] α-quartz,[19] urea,[20] and four polytypes of SiC.[21] Typically, we have achieved agreement within experimental uncertainties for these materials. The goal of the present work is twofold: to present a version of the frequency-dependent formula including local-field corrections that has no apparent divergences, and to evaluate the formula numerically for both GaP, a material used as a reference standard for second-harmonic susceptibilities, and GaAs.

II. FORMULA FOR $\chi^{(2)}(-2\omega;\omega,\omega)$

Considering the second-harmonic response to be the sum of the response of individual electrons in a crystal governed by a reduced Hamiltonian $H_{\mathbf{k}}$ leads to the formula

$$
\begin{aligned}
\chi^{(2)}(-2\omega;\omega,\omega)=-\frac{i}{2}\frac{1}{(2\pi)^3}\int_{\mathrm{BZ}}d\mathbf{k}\sum_{n\pm}^{v}\{&\omega^{-3}[\pm\mathrm{Im}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm2\omega)H_2|n\mathbf{k}\rangle\mp2\mathrm{Im}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm\omega)H_2|n\mathbf{k}\rangle\\
&\pm\mathrm{Im}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm2\omega)H_1G_{n\mathbf{k}}(\pm\omega)H_1|n\mathbf{k}\rangle\\
&+\tfrac{1}{2}\mathrm{Im}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(-\omega)H_1G_{n\mathbf{k}}(\omega)H_1|n\mathbf{k}\rangle]\\
&+\omega^{-2}[2\mathrm{Re}\langle n\mathbf{k}|H_2G_{n\mathbf{k}}(\pm\omega)\phi(\omega)|n\mathbf{k}\rangle\\
&\qquad+\mathrm{Re}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm2\omega)H_1G_{n\mathbf{k}}(\pm\omega)\phi(\omega)|n\mathbf{k}\rangle\\
&\qquad+\mathrm{Re}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm2\omega)\phi(\omega)G_{n\mathbf{k}}(\pm\omega)H_1|n\mathbf{k}\rangle\\
&\qquad+\mathrm{Re}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm\omega)H_1G_{n\mathbf{k}}(\mp\omega)\phi(\omega)|n\mathbf{k}\rangle]\\
&\qquad+\omega^{-1}[\mp\mathrm{Im}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}(\pm2\omega)\phi(\omega)G_{n\mathbf{k}}(\pm\omega)\phi(\omega)|n\mathbf{k}\rangle\\
&\qquad\quad-\tfrac{1}{2}\mathrm{Im}\langle n\mathbf{k}|\phi(\omega)G_{n\mathbf{k}}(-\omega)H_1G_{n\mathbf{k}}(\omega)\phi(\omega)|n\mathbf{k}\rangle]\}
\end{aligned}
\tag{1}
$$

0163-1829/94/49(7)/4532(7)/$06.00

which is equivalent to Eq. (5.13) of Ref. 13 with the notational modification given in Ref. 17. In Eq. (1), $H_1=\mathbf{q}\cdot\nabla_{\mathbf{k}}H_{\mathbf{k}}$ (which is a generalization of the momentum operator $\mathbf{p}$), $H_2=\frac{1}{2}\mathbf{q}\cdot\nabla_{\mathbf{k}}\mathbf{q}\cdot\nabla_{\mathbf{k}}H_{\mathbf{k}}$, where $\mathbf{q}$ is the wave vector of the external perturbation, and ω its frequency. The local field in linear response is given by $\phi(\omega)$, where the external perturbation is taken to have unit strength. Atomic units are used. The band index n runs over the valence (i.e., occupied) states. Also, $G_{n\mathbf{k}}(\omega)=\sum_m |m\mathbf{k}\rangle\langle m\mathbf{k}|/(\epsilon_{n\mathbf{k}}-\epsilon_{m\mathbf{k}}+\omega)$ is a Green operator; the sum on m is over all states. Of course, $H_{\mathbf{k}}|n\mathbf{k}\rangle=\epsilon_{n\mathbf{k}}|n\mathbf{k}\rangle$. (Note that $\sum_{\pm}\frac{1}{2}=1$.)

Equation (1) is not the best form numerically because of the apparent divergences in the formula: (i) for the static limit for an insulator, (ii) due to degenerate states, and (iii) at the frequencies of valence-to-valence transitions. Equation (5.19) of Ref. 13 addresses the divergences of type (i). However, certain terms in Eq. (5.19) remain manifestly divergent, but they can be removed using algebra, and the difference of the sum rule of Sec. III of Ref. 17 and Sec. VI of Ref. 13. Taking this sum rule into account leads to a formula without type (i) divergences

$$\begin{aligned}
\chi^{(2)}(-2\omega;\omega,\omega)=-\frac{i}{2}\frac{1}{(2\pi)^3}\int_{\mathrm{BZ}}d\mathbf{k}\sum_n^v\{&12\mathrm{Im}\langle n\mathbf{k}|H_2\mathcal{G}_{n\mathbf{k}}(2\omega)\mathcal{G}_{n\mathbf{k}}(\omega)H_1|n\mathbf{k}\rangle\\
&+12\mathrm{Im}\langle n\mathbf{k}|H_1F_{n\mathbf{k}}\mathcal{G}_{n\mathbf{k}}(\omega)H_1\mathcal{G}_{n\mathbf{k}}(2\omega)\mathcal{G}_{n\mathbf{k}}(\omega)H_1|n\mathbf{k}\rangle\\
&+6\mathrm{Im}\langle n\mathbf{k}|H_1\mathcal{G}_{n\mathbf{k}}(\omega)H_1\mathcal{G}_{n\mathbf{k}}(2\omega)\mathcal{G}_{n\mathbf{k}}(\omega)F_{n\mathbf{k}}H_1|n\mathbf{k}\rangle\\
&+4\mathrm{Re}\langle n\mathbf{k}|H_2\mathcal{G}_{n\mathbf{k}}(\omega)G_{n\mathbf{k}}\phi(\omega)|n\mathbf{k}\rangle\\
&+2\mathrm{Re}\langle n\mathbf{k}|H_1F_{n\mathbf{k}}[\mathcal{G}_{n\mathbf{k}}(2\omega)+\mathcal{G}_{n\mathbf{k}}(\omega)]H_1\mathcal{G}_{n\mathbf{k}}(\omega)G_{n\mathbf{k}}\phi(\omega)|n\mathbf{k}\rangle\\
&+2\mathrm{Re}\langle n\mathbf{k}|H_1F_{n\mathbf{k}}\mathcal{G}_{n\mathbf{k}}(2\omega)\phi(\omega)\mathcal{G}_{n\mathbf{k}}(\omega)G_{n\mathbf{k}}H_1|n\mathbf{k}\rangle\\
&+8\mathrm{Re}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}\mathcal{G}_{n\mathbf{k}}(2\omega)\phi(\omega)G_{n\mathbf{k}}H_1|n\mathbf{k}\rangle\\
&+\mathrm{Re}\langle n\mathbf{k}|H_1[4\mathcal{G}_{n\mathbf{k}}(2\omega)-2\mathcal{G}_{n\mathbf{k}}(\omega)]H_1\mathcal{G}_{n\mathbf{k}}(\omega)\phi(\omega)|n\mathbf{k}\rangle\\
&+4\mathrm{Re}\langle n\mathbf{k}|H_1\mathcal{G}_{n\mathbf{k}}(2\omega)\phi(\omega)\mathcal{G}_{n\mathbf{k}}(\omega)H_1|n\mathbf{k}\rangle\\
&+\mathrm{Re}\langle n\mathbf{k}|H_1G_{n\mathbf{k}}[8\mathcal{G}_{n\mathbf{k}}(2\omega)+2\mathcal{G}_{n\mathbf{k}}(\omega)]H_1G_{n\mathbf{k}}\phi(\omega)|n\mathbf{k}\rangle\\
&-\mathrm{Re}[\langle n\mathbf{k}|\phi(\omega)|n\mathbf{k}\rangle\langle n\mathbf{k}|H_1[16\mathcal{G}_{n\mathbf{k}}(2\omega)-2\mathcal{G}_{n\mathbf{k}}(\omega)]G_{n\mathbf{k}}^2H_1|n\mathbf{k}\rangle]\\
&-16\mathrm{Re}[\langle n\mathbf{k}|H_1|n\mathbf{k}\rangle\langle n\mathbf{k}|H_1\mathcal{G}_{n\mathbf{k}}(2\omega)G_{n\mathbf{k}}^2\phi(\omega)|n\mathbf{k}\rangle]\\
&+2\mathrm{Im}\langle n\mathbf{k}|H_1F_{n\mathbf{k}}\mathcal{G}_{n\mathbf{k}}(2\omega)\phi(\omega)\mathcal{G}_{n\mathbf{k}}(\omega)\phi(\omega)|n\mathbf{k}\rangle\\
&+4\mathrm{Im}\langle n\mathbf{k}|H_1\mathcal{G}_{n\mathbf{k}}(2\omega)\phi(\omega)\mathcal{G}_{n\mathbf{k}}(\omega)F_{n\mathbf{k}}\phi(\omega)|n\mathbf{k}\rangle\\
&+2\mathrm{Im}\langle n\mathbf{k}|\phi(\omega)F_{n\mathbf{k}}\mathcal{G}_{n\mathbf{k}}(\omega)H_1\mathcal{G}_{n\mathbf{k}}(\omega)\phi(\omega)|n\mathbf{k}\rangle\\
&-6\mathrm{Im}[\langle n\mathbf{k}|\phi(\omega)|n\mathbf{k}\rangle\langle n\mathbf{k}|H_1\mathcal{G}_{n\mathbf{k}}(2\omega)\mathcal{G}_{n\mathbf{k}}(\omega)\phi(\omega)F_{n\mathbf{k}}\phi(\omega)|n\mathbf{k}\rangle]\}\ ,
\end{aligned}\qquad (2)$$

where

$$\mathcal{G}_{n\mathbf{k}}(\omega)=G_{n\mathbf{k}}(\omega)G_{n\mathbf{k}}(-\omega)(1-|n\mathbf{k}\rangle\langle n\mathbf{k}|)$$

and $F_{n\mathbf{k}}=\epsilon_{n\mathbf{k}}-H_{\mathbf{k}}$ is the pseudoinverse of

$$G_{n\mathbf{k}}=\textstyle\sum_{m\neq n}|m\mathbf{k}\rangle\langle m\mathbf{k}|/(\epsilon_{n\mathbf{k}}-\epsilon_{m\mathbf{k}})\ ,$$

i.e.,

$$G_{n\mathbf{k}}F_{n\mathbf{k}}=1-|n\mathbf{k}\rangle\langle n\mathbf{k}|\ .$$

[In Refs. 18 and 19, the symbol $\mathcal{G}_{n\mathbf{k}}(\omega)$ has a different definition. Note that $G_{n\mathbf{k}}$ and $G_{n\mathbf{k}}(\omega)$ are distinct.] In addition to the use of the sum rule, the relation

$$\mathcal{G}_{n\mathbf{k}}(\omega)(F_{n\mathbf{k}}-\omega^2G_{n\mathbf{k}})=G_{n\mathbf{k}}$$

was applied five times to simplify Eq. (5.19) of Ref. 13. It is simple to verify that the static limit of Eq. (2) is given by Eq. (2) of Ref. 17, using the relation $\lim_{\omega\to 0}\mathcal{G}_{n\mathbf{k}}(\omega)=G_{n\mathbf{k}}^2$.

Most discussions of second-harmonic generation consider terms equivalent to the third and fourth terms of Eq. (1) which correspond to the second and third terms of Eq. (2). Additional terms arise in the present formalism due to terms in H_2, which vanish for a local potential but can contribute if a nonlocal potential is chosen, and the local-field corrections (i.e., any term containing ϕ).

In his discussion of the formalism for second-harmonic generation, Aspnes[5] suggested a classification of terms as *vvv*, *vvc*, or *ccv*, depending upon whether 0, 1, or 2 of the intermediate states are in the conduction bands. (The extra *v* arises from the starting bra $\langle n\mathbf{k}|$ and terminating ket $|n\mathbf{k}\rangle$.) We have previously extended this notation to speak of *cv* and *vv* terms as well.[17] This division is fruitful, because the *vvv* (and *vv*) terms vanish in Eq. (2), as they did in Aspnes's formulas. The *ccv* terms are given conveniently by restricting the Green's operators which appear in Eq. (2) to the conduction bands. [This restriction causes terms 11, 12, and 16 to be eliminated from Eq. (2) altogether; these terms are the diagonal elements terms of the *vvc* contribution.] The *vvc* terms may be reworked to avoid apparent divergences associated with occupied orbital degeneracies and orbital energies separated by ω or 2ω. The result is

$$\begin{aligned}\chi^{(2)}|_{vvc}(-2\omega;\omega,\omega)=-\frac{i}{2}\sum_{nm}^{v}\Big[&(\epsilon_{n\mathbf{k}}-\epsilon_{m\mathbf{k}})\mathrm{Im}\{\langle n\mathbf{k}|H_1|m\mathbf{k}\rangle\\
&\times\langle m\mathbf{k}|H_1\mathcal{G}_{m\mathbf{k}}(2\omega)\mathcal{G}_{m\mathbf{k}}(\omega)(6F_{m\mathbf{k}}^2+9F_{m\mathbf{k}}F_{n\mathbf{k}}+6F_{n\mathbf{k}}^2-4\omega^2)\\
&\times\mathcal{G}_{n\mathbf{k}}(\omega)\mathcal{G}_{n\mathbf{k}}(2\omega)H_1|n\mathbf{k}\rangle\}\\
&+\mathrm{Re}\{\langle n\mathbf{k}|H_1|m\mathbf{k}\rangle\langle m\mathbf{k}|H_1(-4G_{m\mathbf{k}}G_{n\mathbf{k}}+\mathcal{G}_{m\mathbf{k}}(2\omega)\mathcal{G}_{m\mathbf{k}}(\omega)\\
&\times[-2F_{m\mathbf{k}}^2-10F_{m\mathbf{k}}F_{n\mathbf{k}}+\omega^2(-4+16G_{m\mathbf{k}}F_{n\mathbf{k}})])\\
&\times\mathcal{G}_{n\mathbf{k}}(\omega)\phi(\omega)|n\mathbf{k}\rangle\}\\
&+\mathrm{Re}\{\langle n\mathbf{k}|\phi(\omega)|m\mathbf{k}\rangle\langle m\mathbf{k}|H_1\mathcal{G}_{m\mathbf{k}}(2\omega)\mathcal{G}_{m\mathbf{k}}(\omega)\\
&\times[-5F_{m\mathbf{k}}^3F_{n\mathbf{k}}-4F_{m\mathbf{k}}^2F_{n\mathbf{k}}^2-5F_{m\mathbf{k}}F_{n\mathbf{k}}^3\\
&+\omega^2(8F_{m\mathbf{k}}^3G_{n\mathbf{k}}+10F_{m\mathbf{k}}^2+42F_{m\mathbf{k}}F_{n\mathbf{k}}+10F_{n\mathbf{k}}^2+8G_{m\mathbf{k}}F_{n\mathbf{k}}^3)\\
&+\omega^4(-40F_{m\mathbf{k}}G_{n\mathbf{k}}-16-40G_{m\mathbf{k}}F_{n\mathbf{k}})+32\omega^6G_{m\mathbf{k}}G_{n\mathbf{k}}]\\
&\times\mathcal{G}_{n\mathbf{k}}(\omega)\mathcal{G}_{n\mathbf{k}}(2\omega)H_1|n\mathbf{k}\rangle\}\\
&+(\epsilon_{n\mathbf{k}}-\epsilon_{m\mathbf{k}})\mathrm{Im}\{\langle n\mathbf{k}|H_1|m\mathbf{k}\rangle\langle m\mathbf{k}|\phi(\omega)\mathcal{G}_{m\mathbf{k}}(\omega)\mathcal{G}_{n\mathbf{k}}(\omega)\phi(\omega)|n\mathbf{k}\rangle\}\\
&+\mathrm{Im}\{\langle n\mathbf{k}|\phi(\omega)|m\mathbf{k}\rangle\langle m\mathbf{k}|\phi(\omega)\mathcal{G}_{m\mathbf{k}}(\omega)(4F_{m\mathbf{k}}+2F_{n\mathbf{k}})\mathcal{G}_{n\mathbf{k}}(2\omega)H_1|n\mathbf{k}\rangle\}\Big]\ ,\end{aligned}\tag{3}$$

where the sums over states in $\mathcal{G}_{n\mathbf{k}}(\omega)$, $G_{n\mathbf{k}}$, and $F_{n\mathbf{k}}$ (and their $n \rightarrow m$ counterparts) are restricted to the conduction bands. The static limit of this expression reproduces Eq. (5) of Ref. 17 as corrected in the erratum. [Specifically, the erratum reports that in term 2 of Eq. (5) of Ref. 17, $-7G_{n\mathbf{k}}^3G_{m\mathbf{k}}$ should be $-4G_{n\mathbf{k}}^3G_{m\mathbf{k}}$ and $-7G_{n\mathbf{k}}G_{m\mathbf{k}}^3$ should be $-10G_{n\mathbf{k}}G_{m\mathbf{k}}^3$.]

Surprisingly, the computational difficulty of Eqs. (2) and (3) is only somewhat worse than its static limit, which has been implemented previously.[16,17] To implement the formulas given in this section, the "virtual machine" approach[22] which we have used for other optical response properties (i.e., the dielectric function, optical activity, and static second-harmonic response[16-19]) proved to be adequate for the frequency-dependent second-harmonic susceptibility with nearly no recoding in Fortran. Instead, Eqs. (2) and (3) were encoded as about 500 statements in a special-purpose optical response language, compared to 250 statements of the same language needed for the static limit. Experience with GaP indicates that about 20% additional computational time is required per frequency for the frequency-dependent formula compared to the evaluation of its static limit alone. The computational time is dominated by changing from the plane-wave representation to the eigenstate representation and back.

The Hamiltonian used in this study is given by the Kohn-Sham local-density approximation (LDA) with a self-energy correction in the form of a "scissors" operator, i.e.,

$$H_{\mathbf{k}}=H_{\mathbf{k}}^{\mathrm{LDA}}+\Delta_{\mathbf{k}}P_{c\mathbf{k}}\ ,\tag{4}$$

where $\Delta_{\mathbf{k}}$ is a scalar and $P_{c\mathbf{k}}$ is the projection operator on to the conduction bands at a point $\mathbf{k}$ in the Brillouin zone. In practice, we choose a $\mathbf{k}$-independent Δ equal to 0.9 eV for GaP and 0.8 eV for GaAs, as suggested by *GW* calculations.[17,23] (The *GW* approximation is a particular screened Hartree-Fock approximation.) This Hamiltonian has been used successfully to predict linear and static second-harmonic optical response in a variety of systems.[16-19,24,25] To obtain the "scissors" correction, Eqs. (A5), (A8), and (A22) of Ref. 17 are applied to Eq. (2) and Eqs. (A5) and (A9) of this reference are applied to Eq. (3). The application is straightforward, with no simplification, so the resulting form is not given here.

Huang and Ching also adapt a shifted band structure in their calculation of the second-harmonic susceptibility of semiconductors.[12] However, their formula is quite different from the present one. Specifically, we have argued previously that renormalizing the matrix elements is an integral part of the "scissors" correction, and that shifting energy denominators alone is inconsistent.[24] Nevertheless, Huang and Ching use formulas with shifted energy denominators, but with unrenormalized momentum matrix elements.

Let me summarize what has been achieved. Equations (2) and (3) sum to a formula for optical second-harmonic generation for insulating crystals including local-field corrections in the presence of a frequency-dependent scalar external field. The resulting formula has no spurious apparent divergences as the optical frequency approaches zero or a valence-valence eigenvalue difference, or for degenerate eigenstates. The all-valence "*vvv*" terms have been removed analytically from the formulas. The static limit does not require a special analytic form; moreover, the formulas are only a little harder to evaluate than their static limit.

III. RESULTS FOR GaP AND GaAs

The formulation presented in Sec. II is illustrated with calculations on GaP and GaAs. The calculation makes use of the Corning LDA code for the ground states.[26] The pseudopotentials are found using the method (and

TABLE I. The minimum direct band gaps of GaP and GaAs (at the Γ point) and the spin-orbit splitting. The "self-energy corrected" values are just the LDA values corrected by a shift suggested by *GW* calculations.[23] The parameters of the calculation used throughout the study are given here.

	GaP	GaAs
Lattice constant (nm)	0.5450	0.5652
E_{cut} (hartree)	10	10
No. conduction bands included	100	100
N_{kpt} for BZ integration	28	60
Δ (eV)	0.9	0.8
E_g LDA (eV)	1.71	0.49
E_g self-energy corrected (eV)	2.61	1.29
Expt. (Ref. 39) (eV)	2.89	1.52
Expt. spin-orbit splitting (Ref. 39) (eV)	0.08	0.34

code) of Hamann.[27] The method is substantially the same as used in our earlier studies.[17–20] The direct gaps and spin-orbit splittings for the two materials are given in Table I, along with the parameters of the calculation.

Our principal interest in this paper is GaP, for several reasons. (i) Spin-orbit effects are neglected in this calculation (although scalar relativistic effects are included in the pseudopotential); spin-orbit effects are smaller in GaP. (ii) The Brillouin-zone integral is somewhat more difficult for GaAs due to the region of minimal direct gap near Γ; finite frequency increases the difficulty of the integration. (The integrand becomes singular at $\omega = E_g$ for linear response, or $2\omega = E_g$ for the second-harmonic response.) For now, the implementation is restricted to the nonabsorbing regime. (iii) There is higher quality data on GaP than GaAs due to the work of Levine and Bethea[28] who measured GaP with the intent of creating a single reference standard for second-harmonic measurements throughout the infrared and visible regions. A similar study was performed by Choy and Byer.[29] Most of the frequency-dependent work of GaAs is for the absorbing regime,[30,31] hence it is out of the scope of the present study. Roberts[15] has recently reconsidered the second-harmonic susceptibility measurement, in particular the values of the reference standards, and has resolved an outstanding disagreement between the measurements of Levine and Bethea[28] and Choy and Byer[29] for both GaP and GaAs.

The results for the dielectric function are given in Table II. Only the low-frequency behavior of the optical response functions is given in this study. Two parameters are adequate to describe the electronic part of the linear dielectric response in most semiconductors and insulators over a large fraction of the band gap.[32] The results are of a quality comparable to those we obtained earlier for the frequency dependence of silicon and germanium.[24] The self-energy correction to the LDA is seen to have a beneficial effect on calculated values compared to the LDA itself. Also, the values given by Huang and Ching[33] are presented, which give a good account of the experiments. The disagreement of the present LDA results to the LDA results of Huang and Ching is most likely due to basis-set effects, which cause the band gap of Huang and Ching to be too large by about 0.5 eV.[27] As reported earlier,[17] the present results for the static dielectric constant agree with another high-quality plane-wave calculation[34] to within 0.2 units of ε_∞.

The principal results of this study are given in Figs. 1 and 2, as well as in Table III. Here, a comparison of theory and experiment is given for GaP and GaAs at principal laser lines at frequencies for which $2\omega < E_g$. The open symbols represent the data given by the original authors. Roberts[15] has reconsidered the reference standards used by a number of authors; these are shown in the figures as corresponding solid symbols. There is agreement within 140% of the experimental uncertainties for all rescaled measurements for the two materials. The effect of local-field corrections is given in Table IV. These corrections are moderate, and decrease modestly with frequency. Omission of the local-field corrections would worsen the agreement with the values of $\chi^{(2)}$ suggested by Roberts, though not decisively. The effect of the finite number of **k** points is expected to be only a few percent. A convergence study for GaAs at zero frequency was given earlier.[17] For GaP, using 10 *k* points, the values for *d* are 2–3 % larger than those given in Table III.

Several other authors have calculated the dispersion of $\chi^{(2)}$ for GaP (Refs. 7 and 12) and GaAs.[6,7,11,12] In all cases, these authors have considered a wider frequency

TABLE II. The electronic part of the macroscopic dielectric constant ε_1 and the first nonvanishing derivative $\varepsilon_1''(\omega)$ for the present calculation, other calculations, and experiment. The subscript "00" indicates the neglect of local-field corrections. The parameters of the calculation are stated in Table I. Values for these quantities were calculated using this program at other energy cutoffs in Ref. 17.

Material	Condition	$\varepsilon_{00}(0)$	$\varepsilon_1(0)$	$\varepsilon_{00}''(0)$	$\varepsilon_1''(0)$
GaP	LDA/OLCAO (Ref. 33)	9.29		0.8	
	LDA	11.1	10.5	1.32	1.25
	$\Delta = 0.9$ eV	9.3	8.8	0.67	0.63
	Experiment (Ref. 40)		9.1		0.69
GaAs	Pseudofunction (Ref. 11)	9		2.0	
	LDA/LCAO (Ref. 33)	11.21		2.2	
	LDA	14.2	13.5	3.86	3.78
	$\Delta = 0.8$ eV	11.5	11.0	1.43	1.38
	Experiment (Ref. 41)		10.9		1.51

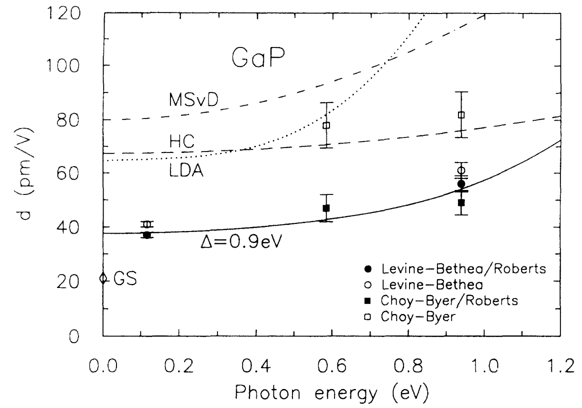

FIG. 1. Nonlinear susceptibility for second-harmonic generation, $d=\frac{1}{2}\chi^{(2)}$ for GaP. The solid line is our calculation with a self-energy correction of $\Delta=0.9$ eV. The dotted line (marked LDA) is our calculation in the LDA. The parameters of the calculation are given in Table I. The dashed line (marked HC) is the calculation of Huang and Ching (Ref. 12). The dash-dotted line (marked MSvD) is the calculation of Moss, Sipe, and van Driel (Ref. 7). The diamond (marked GS) is the calculation of Ghahramani and Sipe (Ref. 35). The circles are the data of Levine and Bethea (Ref. 28). The squares are the data of Choy and Byer (Ref. 29). The open symbols are the data as given by the authors; the solid symbols are the values suggested by Roberts by reconsidering the reference standards used in a variety of second-harmonic generation experiments (Ref. 15).

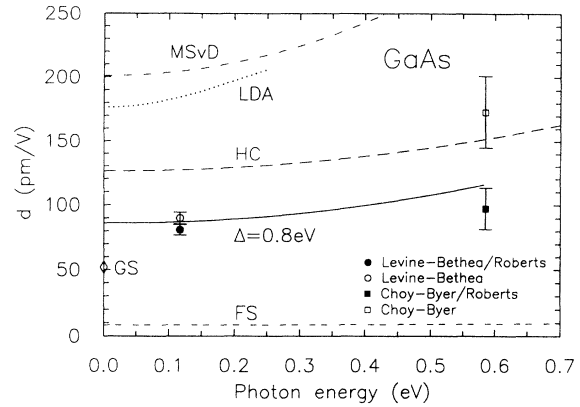

FIG. 2. Nonlinear susceptibility for second-harmonic generation, $d=\frac{1}{2}\chi^{(2)}$ for GaAs. The parameters of the calculation are given in Table I. The double-dash-dotted line is the calculation of Fong and Shen (Ref. 6). The other symbols are defined in Fig. 1.

range than is considered in the present work. (This, incidentally, accounts for why the other papers use many more **k** points than the present one: the evaluation of $\chi^{(2)}$ is more demanding numerically when it is complex; the same phenomenon occurs in linear response.) In addition, a static value for $\chi^{(2)}$ is given by Ghahramani and Sipe.[35] The study of Moss, Sipe, and van Driel[7] used empirical matrix elements; these authors noted that the earlier calculation of Fong and Shen[6] was in error due principally due to matrix element effects rather than local-field corrections. (A more recent calculation from this group[35] used a revised method for obtaining matrix elements which leads to a factor of 3–4 reduction in the predicted values for $\chi^{(2)}$.) Ma *et al.*[11] have also considered local-field corrections using a Matsubara formalism, implemented with the pseudofunction method. The local-field correction of magnitude 3%, comparable to the present result of −7%, is shown in Table IV. Their value of the static dielectric constant for GaAs reported in Table III, namely nine, is quite low. Their values for $d=\frac{1}{2}\chi^{(2)}$, which range from 550 pm/V at 0 frequency to 760 pm/V at 0.585 eV, are substantially higher than other calculations and the experimental results. Considering all the calculations, it is clear that a "consensus value" has yet to emerge.

TABLE III. The frequency-dependent nonlinear susceptibility for second-harmonic generation $d_{123}=\frac{1}{2}\chi^{(2)}_{123}$ in pm/V for GaP and GaAs. Data are given only for frequencies such that $2\omega<E_g$, i.e., the regime in which direct absorption is forbidden. The frequencies are chosen to correspond to common laser wavelengths, which are also presented. The term "rescaled" refers to the new scale for the measurements of Refs. 28 and 29 suggested by Ref. 15. The parameters of the calculation are given in Table I.

ω (eV)		0	0.117	0.585	0.94	1.17
λ (μm)		∞	10.6	2.12	1.32	1.06
GaP	LDA	64.6	65.2	82.1	142.7	
	$\Delta=0.9$ eV	37.6	37.8	42.8	54.1	69.9
	Expt. (Ref. 28) (rescaled)		37±2		56±6	
	Expt. (Ref. 28)		41±2		61±6	
	Expt. (Ref. 29) (rescaled)			47±10	49±9	
	Expt. (Ref. 29)			78±17	82±17	
GaAs	LDA	177	185			
	$\Delta=0.8$ eV	86.2	87.2	116.0		
	Expt. (Ref. 28) (rescaled)		81±5			
	Expt. (Ref. 28)		90±5			
	Expt. (Ref. 29) (rescaled)			97±14		
	Expt. (Ref. 29)			173±28		

TABLE IV. The effect of local-field corrections on the frequency-dependent nonlinear susceptibility for second-harmonic generation $d_{123}=\frac{1}{2}\chi^{(2)}_{123}$ in pm/V for GaP and GaAs; the conditions of the calculation are the same as those of Table III. Results are given for $\Delta=0.9$ eV for GaP and $\Delta=0.8$ eV for GaAs. The local-field correction is seen to decrease slightly with increasing frequency. The parameters of the calculation are given in Table I.

ω (eV)		0	0.117	0.585	0.94	1.17
λ (μm)		∞	10.6	2.12	1.32	1.06
GaP	long wave	42.1	42.3	47.7	59.8	76.7
	total	37.6	37.8	42.8	54.1	69.9
	local field	−11.2%	−11.2%	−11.1%	−11.1%	−11.0%
GaAs	long wave	92.3	93.3	123.0		
	total	86.2	87.2	116.0		
	local field	−7.2%	−7.0%	−6.0%		

One may consider the results in Figs. 1 and 2 for the LDA and for the LDA with a self-energy correction of 0.9 eV (for GaP) or 0.8 eV (for GaAs) to be two curves in a family, parametrized by Δ. (The LDA corresponds to $\Delta=0$.) The magnitude of the static value of $d=\frac{1}{2}\chi^{(2)}$ and the curvature at low frequency will be correlated. Curiously, the calculations of both Moss, Sipe, and van Driel[7] and Huang and Ching[12] display higher static values than the ones found here, but with less curvature. Hence, our pictures are substantially different, and the differences cannot be lumped into a single "energy gap" parameter.

In the early days of second-harmonic generation, Miller made the conjecture known as "Miller's rule" that the quantity

$$\Delta_M(\omega)=\frac{\chi^{(2)}(-2\omega;\omega,\omega)}{\chi^{(1)}(2\omega)[\chi^{(1)}(\omega)]^2}$$

would vary little as a function of frequency, and account for most of the variation from one material to another within a given class.[36] In the present formulation, there is some support for a frequency-independent Δ_M. Terms 2 and 3 of Eq. (2) (after restriction to *ccv* terms) which represent the largest contribution to $\chi^{(2)}$,[5,17] do have the same energy denominator as the product $\chi^{(1)}(2\omega)[\chi^{(1)}(\omega)]^2$ as may be seen by comparison with Eq. (A17) of Ref. 24. However, other terms are present with a variety of frequency dependencies.

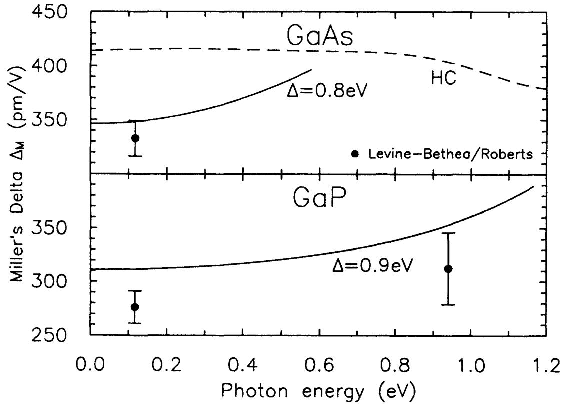

FIG. 3. Miller's Δ_M as a function of frequency. This quantity should be constant if Miller's conjecture is correct. The solid line is the local-field calculation. The parameters of the calculation are given in Table I. The dashed line is from Huang and Ching (Ref. 37). The solid symbols are from the data for second-harmonic generation of Levine and Bethea (Ref. 28) as corrected by Roberts (Ref. 15) with the electronic part of the dielectric function taken from Refs. 40 and 41.

Since Miller's rule remains in use,[15] it is interesting to compare the present calculation for Miller's Δ_M to experiment, and to the calculation of Huang and Ching.[37] For this purpose, only the most credible data, that of Levine and Bethea[28] as corrected by Roberts,[15] is shown in Fig. 3. In GaP, between 0.117 and 0.94 eV, the experimental value of $\chi^{(2)}$ rises by some 51%. The present calculated value for Δ_M rises 13%, compared to a rise of 13±15% for the experiment, and a flat or declining value in the calculation of Huang and Ching.[37] Hence, the comparison is inconclusive, being consistent with both the prediction of the present theory and with a fixed value for Δ_M. This situation may be typical: Roberts[15] remarks "Miller's rule [is] no worse than the 5–20% precision to which the [d] coefficients are usually measured."

IV. CONCLUSIONS

A formula for the frequency-dependent second-harmonic susceptibility of an insulating solid within band theory, including local-field corrections, is given in Eqs. (2) and (3). In the static limit, the formula reduces to a formula presented earlier.[17] The advantages of this form are that there are no apparent divergences as (i) external frequency vanishes, (ii) due to degenerate valence states, or (iii) an equality of the external frequency and the eigenvalue difference of two valence states.

This formula is implemented in the LDA with a self-energy correction in the form of a "scissors" operator (i.e., a rigid shift of the conduction bands) using pseudopotentials in a plane-wave basis. Results are presented for GaP and GaAs. The GaP case in more definitive because the experimental data are better, and because the theoretical prediction is subject to fewer assumptions, such as the neglect of a smaller spin-orbit splitting. There is agreement with the experimental values suggested in a recent review article by Roberts[15] to within experimental uncertainties in many cases, and with 140% of experimental uncertainties in all cases considered.

One of the goals of first-principles calculations is the prediction of new materials with useful properties. The synthesis of β-C_3N_4, which was predicted to be harder than diamond, may be a demonstration of this ability.[38] To accomplish a similar feat for nonlinear optics, materials of interest for nonlinear optical devices will probably operate near resonance, though still in the nonabsorbing regime. The present work which suggests that the frequency dependence of $\chi^{(2)}$ may be calculated accurately may be seen as a small but real step toward a prediction of technological interest.

ACKNOWLEDGMENTS

I thank Professor D. C. Allan and Professor M. P. Teter for providing and supporting the ground-state local-density code used in this study as well as the pseudopotentials. I thank Professor W. Y. Ching for sending large copies of his published figures. This work was supported by the National Science Foundation (DMR-9306318), the Department of Energy—Division of Materials Sciences, and the Cornell Theory Center.

---

[1] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Phys. Rev. Lett. **7**, 118 (1961).

[2] N. Bloembergen, *Nonlinear Optics* (Benjamin, New York, 1965).

[3] B. F. Levine, Phys. Rev. Lett. **22**, 787 (1969); Phys. Rev. B **7**, 2600 (1973).

[4] R. K. Chang, J. Ducuing, and N. Bloembergen, Phys. Rev. Lett. **15**, 415 (1965).

[5] D. E. Aspnes, Phys. Rev. B **6**, 4648 (1972).

[6] C. Y. Fong and Y. R. Shen, Phys. Rev. B **12**, 2325 (1975).

[7] D. J. Moss, J. E. Sipe, and H. M. van Driel, Phys. Rev. B **36**, 9708 (1987).

[8] E. Ghahramani, D. J. Moss, and J. E. Sipe, Phys. Rev. B **43**, 8990 (1991).

[9] S. T. Chui, H. Ma, R. V. Kasowski, and W. Y. Hsu, Phys. Rev. B **47**, 6293 (1993).

[10] A. S. Pine, Phys. Rev. **139**, A901 (1965).

[11] H. Ma, S. T. Chui, R. V. Kasowski, and W. Y. Hsu, Opt. Commun. **85**, 437 (1991).

[12] M.-Z. Huang and W. Y. Ching, Phys. Rev. B **47**, 9464 (1993).

[13] Z. H. Levine, Phys. Rev. B **42**, 3567 (1990); **44**, 5981(E) (1991).

[14] D. A. Kleinman, Phys. Rev. **126**, 1977 (1962).

[15] D. A. Roberts, IEEE J. Quantum Electron. QE-**28**, 2057 (1992).

[16] Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **66**, 41 (1991).

[17] Z. H. Levine and D. C. Allan, Phys. Rev. B **44**, 12 781 (1991); **48**, 14 768(E) (1993). There are minor differences between the values reported here and those reported earlier. These are due to the following sources: in the reference, the standard conditions involved an energy cutoff of 12 hartree, compard to 10 hartree in the present study. According to the convergence study in the reference, this leads to changes of about 1%. One hundred conduction bands were retained in the present study, corresponding to a value of about $E_{\rm cut}^{(sp)}$ of 3.4 hartree for GaP and 3.2 hartree for GaAs, whereas all bands were retained in the reference. According to the convergence study in the reference, this leads to changes of about 0.03%. The error in the formula reported in the erratum caused the reported values of the local-field correction to $d=\frac{1}{2}\chi^{(2)}$ to be quantitatively, but not qualitatively, in error. For GaP, a previously reported -12.5% correction is revised to -11.2% in Table IV. For GaAs, the -8.1% correction of the reference is revised to a -7.2% local-field correction in the present work.

[18] H. Zhong, Z. H. Levine, D. C. Allan, and J. W. Wilkins, Phys. Rev. Lett. **69**, 379 (1992); **70**, 1032(E) (1993).

[19] H. Zhong, Z. H. Levine, D. C. Allan, and J. W. Wilkins, Phys. Rev. B **48**, 1384 (1993).

[20] Z. H. Levine and D. C. Allan, Phys. Rev. B **48**, 7789 (1993).

[21] J. Chen, Z. H. Levine, and J. W. Wilkins (unpublished).

[22] A "virtual machine" is a term from computer science used to denote a particular programming style. Elementary operations are defined and may be applied to any appropriate object. English equivalents of the present implementation might be "Apply the operator H_1 to a wave function" or "transform the result from the plane-wave representation to the eigenstate representation."

[23] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10 159 (1988).

[24] Z. H. Levine and D. C. Allan, Phys. Rev. B **43**, 4187 (1991).

[25] X. Gonze, D. C. Allan, and M. P. Teter, Phys. Rev. Lett. **68**, 3603 (1992).

[26] M. P. Teter, M. C. Payne, and D. C. Allan, Phys. Rev. B **40**, 12 255 (1989).

[27] D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).

[28] B. F. Levine and C. G. Bethea, Appl. Phys. Lett. **20**, 272 (1972).

[29] M. M. Choy and R. L. Byer, Phys. Rev. B **14**, 1693 (1976).

[30] D. Bethune, A. J. Schmidt, and Y. R. Shen, Phys. Rev. B **11**, 3867 (1975).

[31] H. Lotem, G. Koren, and Y. Yacoby, Phys. Rev. B **9**, 3532 (1976).

[32] S. H. Wemple and M. DiDomenico, Jr., Phys. Rev. Lett. **23**, 1156 (1969); Phys. Rev. B **3**, 1338 (1971).

[33] M.-Z. Huang and W. Y. Ching, Phys. Rev. B **47**, 9449 (1993).

[34] S. de Gironcoli, S. Baroni, and R. Resta, Phys. Rev. Lett. **62**, 2853 (1989).

[35] E. Ghahramani and J. E. Sipe, Phys. Rev. B **46**, 1831 (1992).

[36] R. C. Miller, Appl. Phys. Lett. **5**, 17 (1964).

[37] M.-Z. Huang and W. Y. Ching, Phys. Rev. B **45**, 8738 (1992).

[38] A. Y. Liu and M. L. Cohen, Science **245**, 841 (1989); Phys. Rev. B **41**, 10 727 (1990); M. L. Cohen, Science **261**, 307 (1993); C. Niu, Y. Z. Lu, and C. M. Lieber, *ibid.* **261**, 334 (1993).

[39] *Semiconductors Group IV Elements and Group III-V Compounds,* edited by O. Madelung (Springer-Verlag, Berlin, 1991), pp. 93 and 101.

[40] E. D. Palik, in *Handbook of Optical Constants of Solids,* edited by E. D. Palik (Academic, New York, 1985), p. 429.

[41] A. Borghesi and G. Guizzetti, in *Handbook of Optical Constants of Solids* (Ref. 40), p. 445.