Ab initio static dielectric matrices from the density-functional approach. I. Formulation and application to semiconductors and insulators

Mark S. Hybertsen and Steven G. Louie
Department of Physics, University of California, Berkeley, California 94720
(Received 22 September 1986)

As a ground-state expectation value, the static dielectric response function can be obtained exactly within the density-functional approach. This approach is developed in the present paper within the local-density approximation. The *ab initio* pseudopotential method is used, extending the techniques that give excellent structural properties to the calculation of dielectric response functions. In particular, the full dielectric matrix is calculated, and so complete information about local fields is obtained. In contrast to recently proposed direct methods for obtaining the dielectric response matrices, the present approach is based on the usual perturbation formulation for the independent-particle polarizability. The results agree well with results obtained with use of direct methods. The advantage of the perturbative approach is that it allows calculation of the response matrices on a systematic grid of points in the Brillouin zone without significant extra computation or loss of accuracy for points of low symmetry. The response matrices for such a grid are required to describe the response to arbitrary perturbations, e.g., a local change in the potential due to an impurity or defect. The role of exchange and correlation is carefully developed and the relation of the response functions calculated within the density-functional approach to the usual random-phase approximation is illustrated. Results from first principles for the full static dielectric matrices are given for a series of semiconductors and insulators: diamond, Si, Ge, and LiCl. Comparison is made to previous results based on empirical potentials. The importance of local fields is illustrated for the macroscopic dielectric function and by using the concept of the dielectric band structure. Sufficient details of the method and results are included to serve as a reference for development of the dielectric matrix as a tool to be used in other applications. In particular, the additional terms in the long-wavelength dielectric matrix due to nonlocal terms in the ionic pseudopotential are presented.

I. INTRODUCTION

The linear-response approach to screening electrodynamical perturbations has been used widely in condensed-matter applications. For the case of crystals, it has long been recognized that the microscopic response of the electrons can be substantially different than the average or macroscopic response. The density of electrons in a crystal is, in general, inhomogeneous on the scale of the bond lengths or interatomic spacings. It is therefore intuitively clear that the screening fields will also vary on this scale. These are the so-called local fields.

The local fields are reflected in the dielectric response function through its matrix form. In general, it has the form $\epsilon(\mathbf{r},\mathbf{r}';\omega)$, depending explicitly both on the location of the probe $\mathbf{r}$ and the perturbation $\mathbf{r}'$. For a crystal, the periodicity can be exploited to Fourier transform this to reciprocal space, in which case a matrix in the discrete reciprocal-lattice vectors results for each $\mathbf{q}$ in the first Brillouin zone: $\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q},\omega)$. For a system with infinitesimal translation invariance (e.g., the "jellium" model), the dielectric matrix simplifies substantially, having the form $\epsilon(|\mathbf{r}-\mathbf{r}'|;\omega)$. In reciprocal space, the dielectric matrix is then diagonal. It is the off-diagonal elements of the dielectric matrix in reciprocal space that contain the information about the inhomogeneity of the microscopic response of the electrons, the local fields.

Although the response function of simple metals can often be approximated by that of jellium, it is well known that local fields are essential in many applications. One example is the dielectric matrix formulation of the phonon problem.[1,2] Here, the acoustic sum rule requires that the off-diagonal elements of the dielectric matrix be nonzero for semiconductors and insulators. A second example is the screening of impurity potentials in semiconductors where models for the local fields have been employed.[3] A third example is the screened Coulomb interaction required in the calculation of the electron self-energy operator. It has been shown that inclusion of local fields in the screening is essential for predicting the correct quasiparticle spectrum in semiconductors and insulators.[4,5]

Despite the importance of local fields, calculations of the full dielectric response matrix have been rare. This is due in part to the complexity of the calculation required for a real material. There are two aspects to this. Firstly, techniques for band-structure calculations required as input to the dielectric matrix had to be developed. These have been well established for some years now. Secondly, there has been uncertainty as to a reasonable approximate Hamiltonian upon which to base the calculation. We will argue that the density-functional approach[6] using the local-density approximation[7] (LDA) provides a sound basis for proceeding with the calculation of the dielectric

matrix (at least for the static response). The other factor that has hindered inclusion of local-field effects in studies of various physical situations is the difficulty in modeling the off-diagonal elements of the dielectric matrix. Without the guide of first-principles calculations, models have been scarce.[8] The most successful model in application has been the bond-charge model for calculation of the phonon frequencies, which represents the effect of local fields in this case directly instead of by approximating the dielectric matrix itself.[9,10]

Earlier realistic calculations have been based on empirical potentials fitted to reproduce optical properties. This is appealing given that the average gap is related to the dielectric constant. However, use of empirical potentials does not have any *a priori* justification. The full dielectric matrix for $\mathbf{q}\rightarrow\mathbf{0}$ was calculated as a function of frequency to obtain the local-field effect on the optical response of semiconductors. These calculations were carried through in the random-phase approximation[11,12] (RPA) as well as including the ladder bubble diagrams (exchange-correlation effects) in the screening.[13,14] For the static response function, the calculation of $\epsilon_{\mathbf{GG}'}(\mathbf{q}\rightarrow\mathbf{0})$ has been extensively developed.[15,16] Recently, direct methods for obtaining the static dielectric matrix have been proposed. The RPA response function has been discussed[17] and the dielectric matrix within the density-functional approach can be obtained.[18] As we will discuss in detail below, this direct approach is based on a supercell calculation along the lines of the frozen-phonon approach for lattice dynamics. This restricts the calculation to the case of $\mathbf{q}$ at points of reasonably high symmetry. Also, to date, these calculations have been based on empirically derived local pseudopotentials.

These comments illustrate the need for well-founded dielectric matrices calculated from first principles. The static dielectric response matrix is a ground-state expectation value. As such it can, in principle, be obtained within the density-functional approach. We have developed this approach using the perturbative technique of Adler[19] and Wiser[20] for the independent-particle polarizability and the local-density approximation to include the exchange-correlation effects. Both the RPA response function as well as the response function for a test charge probe including exchange-correlation effects are considered. We emphasize that this procedure is well founded theoretically. The only significant approximation is the use of the LDA for the exchange-correlation part of the functional. We use the *ab initio* (nonlocal) pseudopotential to represent the electron-ion interaction.[21] As the pseudopotential is derived from atomic calculations carried out with the LDA, this approach yields a calculated response function consistently obtained within the density-functional approach. The present formulation is equivalent to the proposed direct approach[18] within the linear-response regime, provided that the same pseudopotential is used. Our approach has the added advantage that the dielectric matrices can be readily calculated on a regular grid of $\mathbf{q}$ through the first Brillouin zone as required in many applications. Recently, Baroni and Resta have applied a similar approach to obtain the static dielectric constant ($\mathbf{q}\rightarrow\mathbf{0}$) of Si.[22]

We have calculated $\epsilon^{-1}_{\mathbf{GG}'}(\mathbf{q},\omega=0)$ on a regular grid of $\mathbf{q}$ in the first Brillouin zone for the semiconductors diamond, Si, and Ge as well as the ionic insulator LiCl. (For LiCl, we obtain only the electronic contribution. The phonon contribution is not considered here.) These results are based on well-converged band-structure calculations in each case. Indeed, the present work is based on the same calculational techniques that give excellent structural properties.[23] In addition, the calculated elements of the dielectric matrix are well converged with respect to Brillouin-zone summations and the sum over empty states required in the Adler-Wiser formulation. The present dielectric matrices have been used previously to obtain the electron self-energy operator and the quasiparticle spectrum in these materials.[5] The results for selected columns of the dielectric matrices have also been compared to direct calculations we have done along the lines previously suggested. The agreement is excellent indicating that the present dielectric matrices are numerically well converged. The only number which can be directly compared to experiment is the calculated macroscopic dielectric constant ϵ_0. The result is about 10% too large for diamond, Si, and LiCl (in agreement with the Si result of Baroni and Resta[22]). The discrepancy for Ge is larger. The results depend significantly on the proper treatment of local fields and explicit inclusion of exchange-correlation effects in the LDA. The discrepancy may be attributable to the use of the LDA, a point which may be addressed in the future.

In the present paper, our approach for calculating the static dielectric matrices is given. In Sec. II, we outline the density-functional approach for the dielectric matrix, placing emphasis on clearly delineating the role of exchange-correlation and the relation to the usual RPA response function. The Adler-Wiser formulation is discussed and the direct approach briefly outlined. The symmetry properties of the dielectric matrix are summarized. In Sec. III, sufficient details of the actual method of calculation are given so that other workers can develop similar calculations. Some of the details are described in the Appendixes. The perturbative and direct approaches are shown to give the same results. The resulting dielectric matrices are described in Sec. IV. For reference, the small $\mathbf{G},\mathbf{G}'$ portion of the dielectric matrices is tabulated for selected $\mathbf{q}$. The present dielectric matrices are compared to previous results based on empirical potentials. The idea of the dielectric band structure[24] is exploited to give a concise summary of the present results as well as a direct indication of the effects of local fields. The explicit effect of exchange-correlation on the response functions is illustrated as well as the role of local fields on the macroscopic dielectric function $\epsilon_M(\mathbf{q}+\mathbf{G},\omega=0)$. Concluding remarks are given in Sec. V.

In a subsequent paper, we will illustrate the effect of local fields in screening various external perturbations. In particular, we will give the screening charge induced by a point-charge perturbation. This can be done because the dielectric matrices are now available on a fine grid of $\mathbf{q}$. Systematic trends will be illustrated.

II. FORMULATION OF THE DIELECTRIC MATRIX

In this section, the formulation of the static dielectric response function is reviewed. We outline these basic results to clearly establish definitions, give a coherent presentation, and firmly establish the density-functional approach in the perturbative form used here.[25] Emphasis is placed on the response function in the density-functional approach, the role of exchange-correlation contributions (including distinction of various response functions), and the equivalence of direct and perturbative approaches in the linear-response regime.

A. General framework

The polarizability of the electrons relates a small external potential perturbation $\delta V_{\rm ext}(\mathbf{r})$ to the resulting change in the electron density

$$\delta\rho_{\rm ind}(\mathbf{r})=\int d\mathbf{r}'\,\chi(\mathbf{r},\mathbf{r}')\delta V_{\rm ext}(\mathbf{r}') \ . \tag{1}$$

We adopt the polarizability χ as the basic response function and use it to derive the dielectric function below. Note that the perturbation is assumed to be local. The perturbation caused by a phonon in the *ab initio* pseudopotential approach includes terms that are explicitly nonlocal operators. This case requires an extension of the approach described here and will be discussed in a future publication.[26]

The density-functional approach provides the framework for further discussion. Briefly, the total electronic energy is a unique functional of the electron density with the extremal property that the functional is minimized by the correct density. For a given external potential (electron-ion interaction)[6]

$$E[\rho]=\int d\mathbf{r}\,\rho(\mathbf{r})V_{\rm ext}(\mathbf{r})+F[\rho] \ . \tag{2}$$

The functional F is written in the form

$$F[\rho]=T_0[\rho]+\frac{e^2}{2}\int d\mathbf{r}\int d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}+E_{\rm xc}[\rho] \ , \tag{3}$$

where T_0 is the kinetic energy of noninteracting electrons at the same density, the second term is the average electrostatic (Hartree) energy, and the final term contains all the contribution from exchange and correlation effects. The extremal property leads to the Kohn-Sham equations:[7]

$$\left[\frac{p^2}{2m}+V_{\rm ext}+V_H+V_{\rm xc}\right]\phi_i(\mathbf{r})=\varepsilon_i\phi_i(\mathbf{r}) \ . \tag{4}$$

The exchange-correlation potential is obtained as the functional derivative $V_{\rm xc}(\mathbf{r})=\delta E_{\rm xc}/\delta\rho(\mathbf{r})$. The effective one-electron orbitals give the electron density as

$$\rho(\mathbf{r})=\sum_i f_i\,|\,\phi_i(\mathbf{r})\,|^2, \quad 0\le f_i\le 1 \ . \tag{5}$$

The occupation factor f_i may only be fractional for $\varepsilon_i=\mu$, at the chemical potential.

Based directly on the extremal property of the functional in Eq. (2), the polarizability can be obtained. This has been shown by Wendel and Martin[27] as well as in Ref. 8. The result is

$$\delta V_{\rm ext}(\mathbf{r})=-\int d\mathbf{r}'\frac{\delta^2 F}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}\delta\rho_{\rm ind}(\mathbf{r}') \ . \tag{6}$$

With reference to the explicit form of the functional $F[\rho]$, we have

$$\chi^{-1}(\mathbf{r},\mathbf{r}')=-\frac{\delta^2 T_0}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}-V_C(\mathbf{r}-\mathbf{r}')-K_{\rm xc}(\mathbf{r},\mathbf{r}') \ , \tag{7}$$

where the usual Coulomb interaction V_C enters and we have defined $K_{\rm xc}(\mathbf{r},\mathbf{r}')=\delta^2 E_{\rm xc}/\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')$. From Eq. (6), it is clear that we obtain the inverse of the polarizability defined by Eq. (1). These results show explicitly that the polarizability is a functional (in general unknown) of the electron density.

An alternative way to make the connection to the density-functional approach clear parallels the direct formulation of the dielectric matrix to be discussed below. A self-consistent calculation of the electron density is done. Then the external potential is varied by the addition of $\delta V_{\rm ext}$. A second self-consistent calculation of the electron density is done with the new (perturbed) external potential. Comparison of the two densities yields χ directly by reference to Eq. (1). This analysis demonstrates that the polarizability can be obtained as the difference between two ground-state calculations showing once again that the density-functional approach is a formally well-founded formulation for the dielectric matrix.

In the perturbative approach adopted here,[28] one starts by calculating the independent particle polarizability χ_0. This is defined as the response to the total perturbing potential:

$$\delta\rho(\mathbf{r})=\int d\mathbf{r}'\,\chi_0(\mathbf{r},\mathbf{r}')\delta V_{\rm tot}(\mathbf{r}') \ . \tag{8}$$

Here, the electrons respond to the total change in the effective potential in Eq. (4):

$$\delta V_{\rm tot}(\mathbf{r})=\delta V_{\rm ext}(\mathbf{r})+\delta V_{\rm scr}(\mathbf{r}) \ , \tag{9a}$$

$$\delta V_{\rm scr}(\mathbf{r})=e^2\int d\mathbf{r}'\frac{\delta\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}+\int d\mathbf{r}'\,K_{\rm xc}(\mathbf{r},\mathbf{r}')\delta\rho(\mathbf{r}') \ . \tag{9b}$$

Based on this analysis, the full polarizability χ is related to the independent-particle polarizability χ_0 by

$$\chi=(1-\chi_0 V_C-\chi_0 K_{\rm xc})^{-1}\chi_0 \ . \tag{10}$$

The same expression results from Eq. (7) when the independent-particle polarizability is identified with the functional derivative of T_0. Matrix notation has been adopted here so that Eq. (10) requires a matrix inversion. This and subsequent expressions are also valid in a reciprocal-space basis to be used below, provided that factors of the crystal volume are properly taken into account.

Finally, we give the connection to the dielectric response function. It is defined by $\delta V_{\rm tot}=\epsilon^{-1}\delta V_{\rm ext}$. However, care must be given to the portion of the screening potential included, depending on the nature of the *probe*.[29] Implicitly, we have assumed that the source of the external potential was distinguishable from the electrons. If the probe of the screening potential is a test particle, then it is only affected by the electrostatic term in

Eq. (9b) and the resulting dielectric matrix is given by

$$\epsilon_{\mathrm{TC}}^{-1}=1+V_C\chi \ . \tag{11}$$

This is the so-called test-particle—test-particle response function. However, if the probe is the electrons themselves, then the whole screening potential in Eq. (9b) is effective and one has

$$\epsilon_e^{-1}=1+(V_C+K_{\mathrm{xc}})\chi \ . \tag{12}$$

One can also envision the response to an external charge and distinguish the case of an electronic perturbation from a test charge perturbation (instead of the generic external potential used above). We do not pursue that further here.[29] The usual RPA response function is obtained by setting to zero the exchange-correlation contribution to χ. The result is

$$\epsilon_{\mathrm{RPA}}^{-1}=1+V_C(1-\chi_0 V_C)^{-1}\chi_0 \ , \tag{13}$$

which is equivalent to the more common form: $\epsilon_{\mathrm{RPA}}=1-V_C\chi_0$. The question of whether the RPA response function so obtained would in principle be the same as that derived from a standard many-body treatment is subtle. It is the full response function χ which is provided exactly within the density-functional approach. Derivation of χ_0 from that result depends on interpreting V_{xc} as a mean field. It further depends explicitly on the approximation employed for E_{xc}. Within that framework, derivation of the RPA response function within the density-functional approach is well defined, at least operationally.

For the crystalline case, we exploit the lattice translation symmetry to transform these expressions to a reciprocal-space basis with the following convention used:

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}')=\frac{1}{\Omega}\sum_{\mathbf{q},\mathbf{G},\mathbf{G}'} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q})e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'} \ . \tag{14}$$

Here, Ω denotes the crystal volume. As mentioned in the Introduction, the information about local fields is contained in the off-diagonal elements of the matrix in $\mathbf{G},\mathbf{G}'$. The average, or macroscopic, response to an applied field is determined by the diagonal elements of ϵ^{-1}, not by ϵ. Thus the definition of the macroscopic dielectric function is

$$\epsilon_M(\mathbf{q}+\mathbf{G})=1/\epsilon_{\mathbf{G}\mathbf{G}}^{-1}(\mathbf{q}) \ . \tag{15}$$

In particular, the macroscopic dielectric constant is given by $\epsilon_0=\lim_{\mathbf{q}\to 0}\epsilon_M(\mathbf{q})$, (or ϵ_∞ in the case of ionic materials). Because it is derived from the inverse matrix, the presence of nonzero off-diagonal elements of ϵ changes the macroscopic dielectric constant. As we will show explicitly, the local fields in semiconductors have an important effect on the macroscopic screening (typically 10—20% in ϵ_0).

B. Perturbative approach

Application of first-order perturbation theory to the effective one-particle equations (4) yields the standard result

$$\chi_0(\mathbf{r},\mathbf{r}')=\sum_{i,j}\frac{f_i(1-f_j)}{\varepsilon_i-\varepsilon_j}[\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\phi_j^*(\mathbf{r}')\phi_i(\mathbf{r}')+\mathrm{c.c.}] \ . \tag{16}$$

Transforming to reciprocal space, the resulting expression for χ_0 is

$$\chi_{\mathbf{G}\mathbf{G}'}^0(\mathbf{q})=\frac{2}{\Omega}\sum_{n,n',\mathbf{k}}\frac{f_{n,\mathbf{k}}(1-f_{n',\mathbf{k}+\mathbf{q}})}{\varepsilon_{n,\mathbf{k}}-\varepsilon_{n',\mathbf{k}+\mathbf{q}}}(\langle n,\mathbf{k}\,|\,e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\,|\,n',\mathbf{k}+\mathbf{q}\rangle\langle n',\mathbf{k}+\mathbf{q}\,|\,e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}\,|\,n,\mathbf{k}\rangle+\mathrm{c.c.}) \ . \tag{17}$$

The one-particle states are labeled by Bloch wave vector $\mathbf{k}$ and band index n. The sum over spin has been taken into account explicitly yielding the factor of 2. Conservation of crystal momentum has also been used in the matrix elements. This is just the usual Adler-Wiser formulation.[19,20] For the present purposes, Eq. (17) can be further simplified for systems with a gap between empty and occupied states. For semiconductors and insulators, the result is

$$\chi_{\mathbf{G}\mathbf{G}'}^0(\mathbf{q})=\frac{4}{\Omega}\sum_{c,v,\mathbf{k}}\frac{\langle v,\mathbf{k}\,|\,e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}\,|\,c,\mathbf{k}+\mathbf{q}\rangle\langle c,\mathbf{k}+\mathbf{q}\,|\,e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}\,|\,v,\mathbf{k}\rangle}{\varepsilon_{v,\mathbf{k}}-\varepsilon_{c,\mathbf{k}+\mathbf{q}}} \ . \tag{18}$$

Here, c (v) represents the index for the conduction (valence) states. This is the basic expression required for the calculations presented in this paper.

Once χ_0 has been calculated, the exchange-correlation contribution must be included to get $\epsilon_{\mathrm{TC}}^{-1}$. In the LDA used here,

$$K_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')=\frac{\delta^2 E_{\mathrm{xc}}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')}=\left.\frac{dV_{\mathrm{xc}}}{d\rho}\right|_{\rho(\mathbf{r})}\delta(\mathbf{r}-\mathbf{r}') \ . \tag{19}$$

The ordinary differentiation with respect to ρ of the local density expression for V_{xc} is required. As a result of the δ function, the reciprocal-space expression for K_{xc} is independent of $\mathbf{q}$:

$$K_{\mathbf{G}\mathbf{G}'}^{\mathrm{xc}}(\mathbf{q})=K_{\mathrm{xc}}(\mathbf{G}-\mathbf{G}') \ . \tag{20}$$

This is an artifact of the LDA in particular, *not* of the density-functional approach in general. From Eq. (20) we see that the diagonal part of K_{xc} is constant in the LDA. This behavior is, generally speaking, incorrect, even in the electron gas.[30] However, one should note that K_{xc} always enters with χ_0 in a product, e.g., in Eq. (10). Thus the $\mathbf{q}$-dependence of χ_0 damps the large-$\mathbf{q}$ contribution of the combination $\chi_0 K_{\mathrm{xc}}$, although more slowly than is expected for the correct K_{xc}. This ameliorates the severity of

the LDA for K_{xc} in practice. It has not been possible to go beyond the LDA in this regard in a nontrivial way in actual calculations. The required functional derivatives for K_{xc} in, for instance, the weighted density approximation are significantly more complicated than the tractable expressions for V_{xc}. It has been shown[30] that the weighted density approximation yields generally better large-$\mathbf{q}$ behavior for K_{xc}. Progress in going beyond the LDA may be possible using direct methods.

The explicit expressions for the dielectric matrix in reciprocal space are straightforward to derive from the general expressions given in Sec. II A. Using $V_C(\mathbf{q}+\mathbf{G})=4\pi e^2/\Omega\,|\mathbf{q}+\mathbf{G}|^2$ and properly accounting for factors of the crystal volume, the matrix which needs to be inverted in Eq. (10) is

$$A_{\mathbf{G}\mathbf{G}'}(\mathbf{q})=\delta_{\mathbf{G}\mathbf{G}'}-\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q})\frac{4\pi e^2}{|\mathbf{q}+\mathbf{G}'|^2}-\sum_{\mathbf{G}''}\chi^0_{\mathbf{G}\mathbf{G}''}(\mathbf{q})K_{xc}(\mathbf{G}''-\mathbf{G}')\ . \tag{21}$$

The matrix A must be inverted numerically. Then

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q})=\sum_{\mathbf{G}''}A^{-1}_{\mathbf{G}\mathbf{G}''}(\mathbf{q})\chi^0_{\mathbf{G}''\mathbf{G}'}(\mathbf{q})\ , \tag{22}$$

and

$$\epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})=\delta_{\mathbf{G}\mathbf{G}'}+\frac{4\pi e^2}{|\mathbf{q}+\mathbf{G}|^2}\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q})\ . \tag{23}$$

This is the expression for the test-charge response function. Corresponding expressions can be derived for the electron dielectric function as well as the RPA dielectric matrix. The latter follows simply by eliminating K_{xc} from Eq. (21).

C. Direct approach

Recently, direct approaches for obtaining response functions inspired by the success of frozen-phonon techniques have been demonstrated.[17,18] We briefly describe this approach and the key features of its implementation for comparison to the present perturbative technique. We use the direct method below as a check on the more complete results obtained perturbatively.

If $\mathbf{q}$ is a simple rational fraction of a reciprocal-lattice vector, then a supercell can be found such that $\mathbf{q}$ is itself a reciprocal-lattice vector for the new lattice. Then a monochrome perturbation can be added to the electron-ion interaction:

$$\delta V_{ext}(\mathbf{r})=V_0 e^{i(\mathbf{q}+\mathbf{G}_0)\cdot\mathbf{r}}\ . \tag{24}$$

(The notation is such that the wave vectors all refer to the original lattice.) The self-consistent calculation of the charge density is repeated. The perturbation leads to a change in the charge density for wave vectors $\mathbf{q}+\mathbf{G}$. This is computed by direct comparison to the self-consistent charge density without the perturbation. From this a whole column of the polarizability or dielectric matrix can be obtained:[18]

$$\delta\rho(\mathbf{q}+\mathbf{G})=V_0\chi_{\mathbf{G}\mathbf{G}_0}(\mathbf{q})\ , \tag{25a}$$

$$\delta V_H(\mathbf{q}+\mathbf{G})=V_0[\epsilon^{-1}_{\mathbf{G}\mathbf{G}_0}(\mathbf{q})-\delta_{\mathbf{G}\mathbf{G}_0}]\ . \tag{25b}$$

Exchange-correlation effects are automatically included. By examining the change in the electrostatic (Hartree) potential, the test-charge dielectric matrix is obtained. The electron case can also be considered by including the change in the exchange-correlation potential. For sufficiently small perturbation V_0, only the linear-response regime is probed leading to the usual dielectric matrix. The independent-particle polarizability χ_0 can similarly be obtained.[17]

There are several advantages to the direct approach. Exchange-correlation effects are included through the usual self-consistent procedure in the density-functional approach. Problems with the summation over conduction-band states in the perturbative approach are circumvented. This approach also has several drawbacks. First, a self-consistent calculation is required for each independent column of the dielectric matrix. (Symmetry properties described below limit the number of independent elements of the matrix depending on $\mathbf{q}$.) Second, the size of the supercell required rapidly increases as the $\mathbf{q}$ of interest departs from symmetry points or lines in the Brillouin zone. This is a very serious problem because the size of the eigenvalue problem required in the self-consistent calculations becomes prohibitively large. The cost for the standard matrix diagonalization scales as the cube of the size. Thus, one expects that cases where $\mathbf{q}$ has low symmetry may not be treatable with the direct approach or only by sacrificing the accuracy of the calculation. In practice, the method has been demonstrated using the numerically simpler local pseudopotentials for the Γ, X, and L points in the Brillouin zone and simple fractional points in between, e.g., $\Delta(\frac{1}{2})$ halfway between Γ and X. However, applications where a single localized perturbation (e.g., a point charge) is screened require $\epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$ for a regular grid of $\mathbf{q}$ through the Brillouin zone. Finally, the macroscopic dielectric constant ϵ_0 is not readily obtainable using the direct approach. The perturbative approach applies straightforwardly. Other alternative methods for calculating ϵ_0 have been proposed.[31,32]

In the perturbative approach, Eq. (18) requires a sum over all the virtual (conduction-band) states. Concerns about convergence with respect to the number of conduction-band states included and the perceived difficulty with this have been a major factor inhibiting widespread use of the perturbative approach for calculating the dielectric matrix. They are also a motivating factor in the development of direct techniques. Indeed, one requires a large number of states in the summation. We find typically the number to be of order twice the dimension of χ_0 being calculated for numerical convergence of the elements of χ_0 corresponding to the largest $\mathbf{G},\mathbf{G}'$ included.[33] However, with modern computational facilities available, this does not present an excessive burden. In particular, this aspect of the calculation scales at worst as roughly N^2, where N is the number of plane waves used in the expansion of the wave functions. This is to be compared with N^3 scaling for the standard eigenvalue problem associated with self-consistent calculations required in the direct methods. Moreover, the perturbative approach

applies equally to all the points in a regular $\mathbf{q}$ grid with no loss of numerical accuracy. As will be seen below, for $\mathbf{q}$ of lower symmetry, the calculation is longer, but only scales roughly linearly as the symmetry is reduced.

D. Symmetry considerations

For completeness, and use below, we summarize the symmetry properties[34] of ϵ^{-1}. In real space, the dielectric matrix has the space-group symmetry of the crystal, as applied to both arguments. The periodicity has already been exploited:

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}')=\epsilon^{-1}(\mathbf{r}+\mathbf{T},\mathbf{r}'+\mathbf{T}) , \tag{26}$$

where $\mathbf{T}$ is a lattice translation vector. The point symmetry also applies:

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}')=\epsilon^{-1}(\underline{R}\mathbf{r}+\tau_R,\underline{R}\mathbf{r}'+\tau_R) , \tag{27}$$

where $\underline{R}$ is a rotation matrix and τ_R is the nonprimitive translation vector which may be required for operations in a nonsymmorphic group (e.g., the group for the diamond structure). Equation (27) defines the convention used in this paper for the action of the symmetry operations.

As noted above, the lattice periodicity leads to the simplified reciprocal-space form of ϵ^{-1}. In reciprocal space, the point symmetry has the following consequence. Define

$$\mathbf{q}_1=\underline{R}\mathbf{q}+\mathbf{G}_R , \tag{28}$$

where both $\mathbf{q}$ and $\mathbf{q}_1$ are taken to be in the first Brillouin zone so that an auxiliary reciprocal-lattice vector $\mathbf{G}_R$ is in general required, depending on $\underline{R}$. Then it is straightforward to show that Eq. (28) implies that

$$\epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q}_1)=e^{i(\mathbf{G}'-\mathbf{G})\cdot\tau_R}\epsilon^{-1}_{\mathbf{G}_1\mathbf{G}_1'}(\mathbf{q}) . \tag{29}$$

Here, $\mathbf{G}_1=\underline{R}^{-1}(\mathbf{G}+\mathbf{G}_R)$, with the inverse of the rotation matrix being required and the phase depending on the nonprimitive translation associated with $\underline{R}$. Equation (29) has two consequences. First, given the dielectric matrix for $\mathbf{q}$, the corresponding dielectric matrix for any $\mathbf{q}_1$ related to $\mathbf{q}$ by symmetry can be easily obtained. Second, if $\mathbf{q}_1=\mathbf{q}$, then Eq. (29) provides a relation among different elements of the *same* dielectric matrix. These constraints can be used to reduce the number of elements of the dielectric matrix which need actually be calculated. The saving achieved by only computing those elements which are independent can be substantial for $\mathbf{q}$ of high symmetry, e.g., the Γ, X, or L points in the Brillouin zone for the diamond structure. As noted above, this is also exploited in the direct approach.

III. TECHNICAL DETAILS

In this section we give the details required to implement the calculation of the dielectric matrix as described in the preceding section. We briefly describe the parameters for the underlying *ab initio* pseudopotential band-structure calculations in Sec. III A. The details of the perturbative calculations based on Eq. (18) are given in Sec. III B and in Appendixes A and B. Finally, the results of the perturbative approach are checked against results from the direct approach in Sec. III C.

A. Band calculations

The LDA eigenvalues and wave functions are obtained from a self-consistent *ab initio* pseudopotential calculation[35] carried out in the plane-wave basis.[36] The correlation data from the electron gas calculation of Ceperely and Alder were used.[37] The reference configurations for generation of the norm-conserving pseudopotentials were ionic and potentials for the s, p, and d wave were generated in all cases considered here. For Ge, scalar relativistic effects were included.[38] Transferability to configurations nearby in energy including the atomic ground state was of order 1 mRy for C, Si, Ge, and Cl and of order 5 mRy for Li. The s-wave potentials were separated out as the local part of the potential, leaving nonlocal p and d wells on each site. It is important to note that the calculations were done for the experimental lattice constant. For diamond, Si, and Ge, the calculated lattice constant is within 1% of experiment so there is little difference. However, for LiCl, the nonlinear core correction approach[39] was not used in the pseudopotential for Li, so the calculated lattice constant is several percent smaller than the experimental value, introducing a larger difference in this case.[40] Calculations for each material were carefully converged with respect to the number of plane waves in the basis and the number of special points used in the Brillouin-zone summation for the charge density.[41] These data are summarized in Table I. Scalar relativistic effects are included in the Ge band structure through the pseudopotential. The spin-orbit coupling is neglected here. We note that the calculated macroscopic dielectric constant is particularly sensitive to the convergence of the band structure. For Si and LiCl, a larger energy cutoff has been employed for that special case (17 and 30 Ry, respectively). The rest of the dielectric matrix is less sensitive. However, because of the large number of conduction-band states required in the sum over bands in Eq. (18), one must still diagonalize a large Hamiltonian to obtain these states.

B. Implementation of the perturbative approach

The difficult part of the calculation is in producing $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$ from Eq. (18). Subsequent construction of $\epsilon^{-1}_{\mathrm{RPA}}$ or $\epsilon^{-1}_{\mathrm{TC}}$ is straightforward. The only exception is the case of $\mathbf{q}\rightarrow\mathbf{0}$ where special care must be taken with the nonanalytic portion of the matrix. The forms for ϵ^{-1} given in Eqs. (11)–(13) are well suited to careful treatment of the $\mathbf{q}\rightarrow\mathbf{0}$ case. Further details for this case are given in Appendix B. In this section, we concentrate on details associated with calculation of χ_0.

The symmetry properties of the dielectric matrix summarized in Sec. II D apply to χ_0. In particular, χ_0 is generally Hermitian. For the materials considered explicitly here, the inversion symmetry further leads to real symmetric matrices. Thus, only the lower triangle of $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q})$ need be calculated. In addition, Eq. (29) can be exploited to reduce the number of elements of the polarizability that need be calculated. As an example, for the case of $\mathbf{q}=\mathbf{q}_X$ in Si there are only 659 independent elements in the lower

TABLE I. Atomic configurations and cutoff radii used to generate the *ab initio* pseudopotentials are listed. The cubic lattice constant, the cutoff in plane-wave kinetic energy used in the self-consistent calculations, the approximate corresponding number of plane waves, and the number of **k** points in the irreducible part of the Brillouin are also given for each bulk case.

	Configuration	r_c (a.u.) s p d	a (a.u.)	E_{max} (Ry)	N	**k** points
C	$2s^2 2p^{0.8} 3d^{0.2}$	1.11,1.01,1.36	6.742	50.0	450	28
Si	$3s^2 3p^{0.8} 3d^{0.2}$	1.58,1.93,1.93	10.260	14.0	320	10
Ge	$4s^2 4p^{0.8} 4d^{0.2}$	1.54,1.98,2.42	10.686	20.0	460	10
Li	$2s^{0.15} 2p^{0.05} 3d^{0.05}$	2.59,2.72,2.72	9.700	25.0	600	10
Cl	$3s^2 3p^{3.8} 3d^{0.2}$	1.18,1.31,1.52				

triangle for a 126×126 matrix. Of course, for **q** of lower symmetry, the number of independent elements increases.

The most tedious aspect of calculating χ_0 is the Brillouin-zone summation and the associated plane-wave matrix elements which must be calculated for each **k**. Matrix elements must be computed connecting to all the large number of conduction-band states required in Eq. (18). In a usual band calculation, one only needs to consider **k** in an irreducible wedge of the Brillouin zone. That is because the charge density has the full symmetry of the crystal. Here, the problem is somewhat more complicated. From Eq. (29), it follows that $\chi^0_{GG'}(\mathbf{q})$ only has the symmetry from the little group of **q**. Therefore the degree of reduction achievable in the Brillouin-zone summation depends on **q**. Furthermore, care must be taken to ensure that all phases are correctly accounted for in that reduction. The detailed expressions are given in Appendix A. The important result is that the number of separate **k** that must be considered are those that fall in the wedge of the Brillouin zone found by reducing according to the symmetry operations in the little group of **q**, BZ_q. Most of the computation time is devoted to obtaining the plane-wave matrix elements required. Therefore, for a **q** with any symmetry, this results in a significant saving in computation time because the matrix elements need only be calculated for those **k** in BZ_q. As an example, consider $\mathbf{q}=\mathbf{q}_X$ in the Brillouin zone for the diamond structure. Represent the sum over the usual irreducible wedge of the Brillouin zone by a sum over the ten special **k** points.[41] These correspond to 256 **k** points in the full Brillouin zone. However, it turns out that there are only 20 **k** points in BZ_q for $\mathbf{q}=\mathbf{q}_X$.

Convergence of the elements of $\chi^0_{GG'}(\mathbf{q})$ with respect to the Brillouin-zone summation and number of conduction bands included must be addressed. We have found the ten-**k**-point scheme to be adequate for all materials considered here.[42] The actual number used depends on **q**. The exception to this is the case of $\mathbf{q}\to\mathbf{0}$. Here, the $\mathbf{G}=\mathbf{G}'=\mathbf{0}$ element (the "head" of the dielectric matrix) requires many more **k** points to achieve convergence. From the expressions in Appendix B, this comes from the extra powers of the energy denominator that enter the expression for the head of the dielectric matrix. The convergence properties of the head of the matrix for $\mathbf{q}\to\mathbf{0}$ are illustrated in Table II. Evidently the number of **k** points required varies depending essentially on the size of the gap between the empty and occupied states. At least 60 **k** points are required for good numerical convergence in the semiconductors with even more necessary for Ge. For the insulators, ten **k** points seems to be quite adequate. The case of Ge is more difficult because of the very small gap in the LDA spectrum. The so-called "wings" (**G** or **G'** equal to zero) of the $\mathbf{q}\to\mathbf{0}$ dielectric matrix present intermediate convergence properties. We found ten **k** points to be adequate. (The exception to this was Ge where the low-**G** elements of the wings required 28 **k** points.) These results are consistent with previous work.[15,22]

The difficulty of convergence with respect to the number of conduction bands included is largely a problem for the large **G**,**G'** elements of the polarizability. Thus the number of bands required depends on the size of the dielectric matrix of interest. We summarize the sizes considered here as well as the number of bands included in Table III. A spherical cutoff is used to determine the size of the dielectric matrices: $|\mathbf{q}+\mathbf{G}|<G_{max}$. With matrices of dimension 130 to 230, we find that such properties as the macroscopic dielectric constant are well converged with respect to the size of the matrix. We have found these sizes to be quite adequate for screening local perturbations in various applications.[5] As discussed by Baldereschi and Tosatti,[15] the large **G**,**G'** components of the dielectric matrix are dominated by matrix elements connecting the valence bands to high conduction bands. For this reason, we find that the number of conduction bands included must be of order double the size of the dielectric matrix to achieve numerical convergence of all the elements of the matrix to within a few percent. Representative data are shown for Si in Table IV to illustrate this point. For small **G**,**G'** it is evident that χ_0 converges rapidly. However, for the largest momentum

TABLE II. Convergence of $\epsilon_{00}(\mathbf{q}\to\mathbf{0})$ in the RPA is illustrated for each of the crystals considered as a function of the number of **k** points used in the irreducible wedge of the Brillouin zone.

	10 points	28 points	60 points	110 points
Diamond	6.06	6.06		
Si	14.57	13.71	13.61	
Ge	28.30	23.14	21.88	21.56
LiCl	3.40	3.40		

TABLE III. For each crystal considered, the reciprocal-space cutoff used to determine the size of the dielectric matrices is given along with the corresponding average size of the matrices calculated and the number of bands included in the sum over conduction-band states.

	G_{max}	N	Number of bands
Diamond	5.55	220	388
Si	3.1	140	196
Ge	3.5	220	420
LiCl	3.8	210	388

transfer, the convergence is still at the level of approximately 10%.

Finally, the case of $\mathbf{q}\rightarrow\mathbf{0}$ presents other special difficulties. Generally, $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}\rightarrow\mathbf{0})$ depends on the direction of $\mathbf{q}$. Thus it has the symmetry associated with that direction in the Brillouin zone, *not* the full point symmetry usually associated with $\mathbf{q}=\mathbf{0}$.[43] There are generally nonanalytic contributions to the elements of ϵ^{-1}. This follows along the lines described by Pick, Cohen, and Martin.[1] The macroscopic dielectric constant for cubic materials, such as those considered here, is a scalar independent of the orientation of $\mathbf{q}$. However, the wings of the matrix depend explicitly on that orientation.[44] In the calculation for $\mathbf{q}\rightarrow\mathbf{0}$, the matrix elements with $\mathbf{G}=\mathbf{0}$ or $\mathbf{G}'=\mathbf{0}$ must be carefully evaluated. We employ a $\mathbf{k}\cdot\mathbf{p}$-type expansion of $|c,\mathbf{k}+\mathbf{q}\rangle$. Due to the nonlocal part of the pseudopotential, extra terms enter. This is described in more detail in Appendix B. Furthermore, in forming the dielectric matrices, the nonanalyticities of the head and wings must be carefully taken into account. Further details are given in Appendix B with special attention given to the case of ϵ^{-1}_{TC}.

The $\mathbf{q}\rightarrow\mathbf{0}$ response function is required to obtain the macroscopic dielectric constant. There are applications where the $\mathbf{q}=\mathbf{0}$ response function is of interest.[45] This is also well defined and has the full point symmetry of the crystal. Charge conservation requires the wings to be zero in this case. The comparison of these two cases is already fully developed in previous work.[15] Our interest is primarily in the $\mathbf{q}\rightarrow\mathbf{0}$ case.

C. Comparison to direct calculations

In order to check the results of the perturbative calculation, the direct approach has been used to evaluate selected columns of χ and $\epsilon^{-1}_{\mathrm{TC}}$ at $\mathbf{q}=\mathbf{q}_\Gamma$, $\mathbf{q}=\mathbf{q}_{\Delta(1/2)}$, and $\mathbf{q}=\mathbf{q}_X$. The latter two calculations require a supercell of volume four and two times the usual cell volume, respectively. The required self-consistent calculations were carried out with the same ionic pseudopotentials as used in the perturbative calculations. The cutoffs were scaled back somewhat to accommodate the necessary matrices in the supercell calculation. We present results here for the case of Si with $\mathbf{q}=\mathbf{q}_X$. Plane waves of kinetic energy up to 9 Ry were included. Perturbations of 2 and 4 mRy were examined and the results found to be linear.

For simplicity, only the cosine part of the perturbation was considered. As a result, it is easily seen that the direct calculation yields

$$\frac{\delta V(\mathbf{q}_X+\mathbf{G})}{\delta V_{\mathrm{ext}}}=\epsilon^{-1}_{\mathbf{G}\mathbf{G}_0}(\mathbf{q}_X)+\epsilon^{-1}_{\mathbf{G},(00\bar{2})-\mathbf{G}_0}(\mathbf{q}_X)\ . \tag{30}$$

The *umklapp* $(00\bar{2})$ brings $\mathbf{q}_X$ to $-\mathbf{q}_X$ in the Brillouin zone. A similar result holds for χ given by $\delta\rho/\delta V_{\mathrm{ext}}$. The results for Si are compared to the perturbative calculation in Table V. The polarizability χ and test-charge dielectric matrix $\epsilon^{-1}_{\mathrm{TC}}$ are shown for the case of $\mathbf{G}_0=\mathbf{0}$. The perturbative results include exchange-correlation effects in the LDA evaluated as indicated in Sec. II. It is seen that the agreement is excellent. The convergence with respect to number of conduction bands and size of the dielectric matrix is evidently adequate.

IV. DIELECTRIC MATRICES FOR DIAMOND, Si, Ge, AND LiCl

The dielectric matrices have been calculated on a regular grid of $\mathbf{q}$ points in the Brillouin zone for diamond, Si, Ge, and LiCl, using the techniques described in Sec. III. In this section we summarize and illustrate the results of those calculations. Particular emphasis is placed here on tabulating representative results for reference and illustrating the importance of the off-diagonal elements of the dielectric matrix as well as exchange-correlation effects. This is done by displaying selected elements of the dielectric matrices (Sec. IV A), showing the eigenvalues of the symmetric dielectric matrices (the dielectric band structure in Sec. IV B) and by showing the results for the macroscopic dielectric function (Sec. IV C).

A. Representative results

With a firm theoretical foundation in the density-functional approach, we feel the present results will serve as a useful reference for other groups developing the dielectric matrix approach as a tool for other work. For

TABLE IV. Convergence as a function of the number of conduction bands included is illustrated for the case of Si for selected elements of $\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}=\mathbf{q}_X)$. Units for χ_0 are electrons per cell per Ry.

G	G′	68 bands	164 bands	196 bands
$(\bar{1}\bar{1}\bar{1})$	$(00\bar{2})$	−0.003 66	−0.003 70	−0.003 70
$(\bar{1}\bar{1}\bar{1})$	$(\bar{1}\bar{1}\bar{1})$	−0.035 42	−0.035 61	−0.035 61
$(\bar{2}\bar{2}\bar{2})$	$(\bar{2}\bar{2}\bar{2})$	−0.007 71	−0.011 99	−0.012 18
$(\bar{2}22)$	$(\bar{2}\bar{2}\bar{2})$	−0.000 98	−0.001 06	−0.001 06
(004)	(004)	−0.001 95	−0.005 78	−0.006 53

TABLE V. Comparison of results from the direct and perturbative approaches for Si with $\mathbf{q}=\mathbf{q}_X$ and for $\mathbf{G}_0=\mathbf{0}$. As described in the text, response to a cosine perturbation is shown which in general involves the sum of two components of the response function χ for $\delta\rho$ and ϵ^{-1} for $\delta V_{\rm tot}$. The units used for $|\mathbf{q}+\mathbf{G}|^2$ are a.u.^{2} and for the polarizability are electrons per cell per Ry.

		$\delta\rho(\mathbf{q}+\mathbf{G})/\delta V_{\rm ext}$		$\delta V_{\rm tot}(\mathbf{q}+\mathbf{G})/\delta V_{\rm ext}$	
G	$\lvert\mathbf{q}+\mathbf{G}\rvert^2$	Pert.	Dir.	Pert.	Dir.
(000)	0.375	−0.0108	−0.0107	0.274	0.281
$(1\bar{1}\bar{1})$	0.750	−0.0011	−0.0011	−0.037	−0.037
$(2\bar{2}\bar{2})$	3.375	0.0012	0.0013	0.009	0.010
(004)	9.376	−0.0003	−0.0004	−0.0008	−0.0010

that reason, we tabulate some of the elements of the dielectric matrix here.

Table VI gives the first few *independent* elements of $\tilde{\epsilon}_{GG'}(\mathbf{q}\to\mathbf{0})$ in Si. (The body of the matrix has the symmetry of $\mathbf{q}=\mathbf{0}$.) The results presented are for the symmetric dielectric matrix defined by

$$\tilde{\epsilon}_{GG'}(\mathbf{q})=(\,|\mathbf{q}+\mathbf{G}|\,/\,|\mathbf{q}+\mathbf{G}'|\,)\epsilon_{GG'}(\mathbf{q})\ . \tag{31}$$

This has the advantage of eliminating the singularities associated with ϵ and is readily compared to the empirical pseudopotential results tabulated by Baldereschi and Tosatti.[15] Results for the RPA response function as well as the test-charge response function are shown. First comparing the RPA results to the results of Baldereschi and Tosatti (also in the RPA), the $\mathbf{G}=\mathbf{G}'=\mathbf{0}$ component is substantially larger in the present calculation. This difference seems to be considerably less outside of the first Brillouin zone. Also, the off-diagonal elements are larger in magnitude, indicating larger local fields in the present results. Inclusion of exchange-correlation effects (the test-charge case) leads to more effective screening. Both the diagonal and off-diagonal elements of the dielectric matrix are enhanced. As will be seen more clearly in Sec. IV B, the inclusion of exchange-correlation in fact leads to larger local fields. The off-diagonal elements are larger relative to the diagonal elements. This is in agreement with observations in previous work.[46]

The inverse matrices are considered in Table VII for $\mathbf{q}\to\mathbf{0}$ in Si. For this case, the entire matrix has the symmetry of the orientation of $\mathbf{q}$ as discussed in Appendix B. Hence, more independent elements appear. Comparison of the RPA to the test-charge response function yields similar results to those noted above. The diagonal elements are substantially smaller in the test-charge case, indicative of more effective screening. This is straightforward to understand with reference to Eqs. (21)—(23) in Sec. II. The diagonal elements of χ_0 are negative as is $K_{\rm xc}(\mathbf{G}=\mathbf{0})$. Therefore, considering diagonal elements only for the moment, the presence of $K_{\rm xc}$ reduces A which enhances χ in Eq. (22). But diagonal elements of χ are negative and ϵ^{-1} comes about as the difference from unity [Eq. (23)] so that an enhanced χ leads to reduced diagonal elements of ϵ^{-1} corresponding to more effective screening.[47] Also the off-diagonal elements in general are *larger* in the test-charge case so that it is apparent that relative to the diagonal elements they are more important.

The full density-functional response function gives considerably more effective screening for small momentum transfer than does the response function based on the empirical-pseudopotential-method (EPM) band structure. This may appear to be obvious. The EPM band structure reproduces the experimental gap in the spectrum, while the gap in the LDA eigenvalues is well known to be too small. The actual situation is subtler. First, simply opening up the gap with a "scissor" type operation drastically alters the $\mathbf{G}=\mathbf{G}'=\mathbf{0}$ element of the dielectric matrix yielding far too small a macroscopic dielectric constant. This sensitivity may be traced to the third power of the energy denominator in Eq. (B2). Second, the required ma-

TABLE VI. The present results for $\tilde{\epsilon}(\mathbf{q}\to\mathbf{0})$ in Si calculated in RPA and including exchange-correlation effects in the LDA are compared to previous results based on the EPM band structure ($\mathbf{q}\to\mathbf{0}$ along [100]). Results are for the symmetric dielectric matrix defined in the text.

G	G′	RPA	TC (LDA)	RPA[a] (EPM)
(000)	(000)	13.610	15.399	11.305
(111)	(000)	−0.441	−0.659	−0.392
(111)	(111)	1.780	2.139	1.718
$(\bar{1}11)$	(111)	0.013	0.033	0.007
$(1\bar{1}1)$	$(\bar{1}11)$	0.009	0.047	0.025
$(1\bar{1}\bar{1})$	$(\bar{1}11)$	−0.186	−0.372	−0.156

[a]Reference 15.

TABLE VII. The present results for $\tilde{\epsilon}^{-1}(\mathbf{q}\to\mathbf{0})$ in Si calculated in RPA and including exchange-correlation effects in the LDA are compared to previous results based on the EPM band structure. Results are for the symmetric dielectric matrix defined in the text.

G	G′	RPA	TC (LDA)	RPA[a] (EPM)
(000)	(000)	0.082	0.077	0.098
(111)	(000)	0.019	0.022	0.021
(111)	(111)	0.595	0.522	0.610
$(\bar{1}11)$	(111)	0.007	0.011	0.009
$(1\bar{1}1)$	(111)	−0.002	−0.001	−0.001
$(1\bar{1}1)$	$(\bar{1}11)$	−0.002	−0.006	−0.008
$(11\bar{1})$	$(1\bar{1}1)$	0.008	0.007	0.001
$(1\bar{1}\bar{1})$	$(\bar{1}11)$	0.049	0.066	0.044

[a]Reference 15.

TABLE VIII. Results for $\tilde{\epsilon}^{-1}(\mathbf{q}=\mathbf{q}_X)$ are tabulated for the four materials studied here for the first few independent elements of the matrix. Data are for the test-charge response function. Note that $\mathbf{q}=\mathbf{q}_X=(0,0,1)(2\pi/a)$ here.

G	G'	Diamond	Si	Ge	LiCl
(000)	(000)	0.3909	0.2749	0.2822	0.4886
(00$\bar{2}$)	(000)	0.0	0.0	0.0	0.1288
(11$\bar{1}$)	(000)	−0.0420	−0.0259	−0.0305	−0.0156
(11$\bar{1}$)	(11$\bar{1}$)	0.5422	0.4138	0.4360	0.5727
(11$\bar{1}$)	($1\bar{1}\bar{1}$)	0.0	0.0	0.0	0.0300
($1\bar{1}\bar{1}$)	($\bar{1}$1$\bar{1}$)	0.0018	0.0456	0.0330	0.0713

trix elements in Eq. (18) play an equally important role in determining the dielectric matrix. The wave functions and hence matrix elements in the EPM calculation have no particular foundation. Finally, the perturbative formulation was shown to be equivalent to the direct approach involving the difference between two ground-state calculations. Seen in this way, the matrix elements and energy denominator together have meaning in the perturbative approach while the energy denominator that appears has *no spectroscopic interpretation.*

The results for $\tilde{\epsilon}^{-1}$ at $\mathbf{q}=\mathbf{q}_X$ and $\mathbf{q}=\mathbf{q}_L$ are summarized in Tables VIII and IX. The first few independent elements of the test-charge response function are shown for diamond, Si, Ge, and LiCl. The results generally follow the trend one would expect based on the macroscopic dielectric constants of these materials. From least to most effective screening one would rank these in the order LiCl, diamond, Si and Ge. One would expect the ranking for importance of the off-diagonal elements to be just the opposite with local fields being largest in LiCl. This is generally the case with one interesting exception. The diagonal elements of $\tilde{\epsilon}^{-1}$ for Ge are generally somewhat *larger* than for Si, indicating slightly *less* effective screening. On the other hand, the off-diagonal elements in Ge are frequently larger, suggesting local fields at least as large in Ge as in Si. This trend is also apparent in the results tabulated by Baldereschi and Tosatti.[15]

B. Dielectric band structure

It can be quite difficult to elucidate the meaning of the off-diagonal elements of the dielectric matrix in a simple, quantitative fashion. No particular element can be singled out for meaningful discussion and the aggregate effect is only apparent in the local fields generated by a particular perturbation. This latter approach will be further illustrated in the following paper. One solution to this problem is provided by the concept of a dielectric band structure introduced by Baldereschi and Tosatti.[24] One observes that $\tilde{\epsilon}$ introduced in Eq. (31) is a Hermitian matrix. It has a real eigenvalue spectrum. These eigenvalues, or at least the most important ones, then contain the essential information in the dielectric matrix. We adopt this approach here as a concise way of summarizing our results for the dielectric matrix. We refer to Baldereschi and Tosatti[24] for more detailed discussion and interpretation of this approach.

Table X gives the largest eigenvalues of $\tilde{\epsilon}(\mathbf{q}\rightarrow\mathbf{0})$ for Si with $\mathbf{q}\rightarrow\mathbf{0}$ along (001). The RPA and test-charge response functions are considered in comparison to the EPM-based results of Baldereschi and Tosatti.[24] The eigenvalues are labeled according to the irreducible representations of the full space group. However, the wings of this matrix only display the symmetry of the orientation of $\mathbf{q}$. Thus the largest eigenvector does not precisely transform according to Γ_1 but rather Δ_1. Furthermore, the Γ_{15} eigenvalues are split into longitudinal (Δ_1) and transverse (Δ_5) components. For comparison, the corresponding diagonal elements of $\tilde{\epsilon}$ are shown. The largest Γ_1 eigenvalue corresponds to the (000,000) element of the matrix while the other eigenvalues shown correspond to the eight elements equivalent to the (111,111) entry. (One should note that there is an eigenvalue of $\Gamma_{12'}$ symmetry not shown that is slightly larger than the second Γ_1 value shown in Table X. It is associated with the next star of

TABLE IX. Results for $\tilde{\epsilon}^{-1}(\mathbf{q}=\mathbf{q}_L)$ are tabulated for the four materials studied for the first few independent elements of the matrix. Results are for the test-charge response function. Note that $\mathbf{q}=\mathbf{q}_L=(\frac{1}{2},\frac{1}{2},\frac{1}{2})2\pi/a$ here.

G	G'	Diamond	Si	Ge	LiCl
(000)	(000)	0.3894	0.2828	0.2870	0.4763
($\bar{1}\bar{1}\bar{1}$)	(000)	−0.0505	−0.0719	−0.0620	−0.1746
($\bar{1}\bar{1}$1)	(000)	−0.0523	−0.0400	−0.0433	0.0306
($\bar{1}\bar{1}$1)	($\bar{1}\bar{1}\bar{1}$)	−0.0050	−0.0019	−0.0029	−0.0151
($\bar{1}\bar{1}$1)	($\bar{1}\bar{1}$1)	0.6067	0.4822	0.5054	0.6488
($\bar{1}$1$\bar{1}$)	($\bar{1}\bar{1}$1)	−0.0227	0.0089	0.0091	0.0053
($\bar{2}$00)	($\bar{1}\bar{1}$1)	−0.0619	−0.0520	−0.0504	0.0666
($\bar{2}$00)	($1\bar{1}\bar{1}$)	−0.0218	−0.0013	−0.0058	−0.0423

TABLE X. The dielectric band structure of $\tilde{\epsilon}(\mathbf{q}\rightarrow\mathbf{0})$ is compared for the RPA and including exchange-correlation effects in the LDA for the case of Si. Here $\mathbf{q}\rightarrow\mathbf{0}$ along [001] so that the Γ_{15} eigenvalues are split into longitudinal L and transverse T components. This is compared to results based on an empirical pseudopotential band structure. For reference, the corresponding diagonal element of the symmetric dielectric matrix is shown.

	RPA	TC (LDA)	RPA[a] (EPM)
Γ_1	13.79	15.75	
(diag)	(13.60)	(15.40)	
Γ_2'	2.46	3.51	1.98
$\Gamma_{15}(T)$	2.04	2.59	1.95
(L)	1.95	2.39	1.82
Γ_{25}'	1.84	2.12	1.75
Γ_1	1.66	1.93	1.64
(Diag.)	(1.78)	(2.14)	

[a]Reference 24.

TABLE XI. The first few eigenvalues in the dielectric band structure of $\tilde{\epsilon}^{-1}$ at the X and L points are compared for the four materials studied here. The test-charge response function is considered here. The symmetry designations in parenthesis for the X point refer to LiCl. Note that the order of the eigenvalues at the L point is rather different for LiCl.

	Diamond	Si	Ge	LiCl
$X_1(X_4')$	0.2668	0.2303	0.2325	0.2908
(X_1)				0.3016
$X_4(X_5')$	0.4893	0.3113	0.3501	0.3687
$X_1(X_4')$	0.4791	0.3617	0.3793	0.5608
(X_1)				0.6217
$X_3(X_5')$	0.6240	0.5095	0.5369	0.6275
L_1	0.2228	0.1658	0.1780	0.2511
L_2'	0.3226	0.2944	0.2869	0.3052
L_1	0.3824	0.3293	0.3445	0.5257
L_3'	0.5493	0.3843	0.4208	0.6486
L_3	0.5770	0.4209	0.4476	0.3873
L_2'	0.6412	0.5225	0.5439	0.6381

diagonal elements.) The order of the eigenvalues is identical to that reported in Ref. 24 as regards the symmetry.

Two points made in the preceding section are immediately apparent in Table X. First, the test-charge response function gives more effective screening than the RPA response function; the eigenvalues are all larger for the test-charge case. Second, the density-functional response functions give more effective screening than those based on the EPM band structure. The other interesting aspect of the dielectric band structure is the directly apparent role of the off-diagonal elements of the dielectric matrix. By analogy to usual electronic band structure, they open "gaps" and induce shifts in the spectrum. This is apparent upon comparison of the eigenvalues of $\tilde{\epsilon}$ to the corresponding diagonal elements shown in Table X. The Γ_1 eigenvalue is larger than the (000,000) component of the matrix. A perturbation corresponding to this eigenpotential is *more* effectively screened than the simple long-wavelength plane wave. The eight equivalent diagonal elements represented by the (111,111) element are split into several components. The largest eigenvalue is some 30% larger than the diagonal element, indicating a significant contribution from local fields. The splitting is also substantial. Finally, the shifts are larger for the test-charge case than for the RPA case. This is a direct indication that the off-diagonal elements are significantly more important in the test-charge response function.

The *smallest* eigenvalues of the inverse test-charge dielectric matrix are given for diamond, Si, Ge, and LiCl at the X and L points in the Brillouin zone in Table XI. Similar analysis applies to these data upon comparison to the diagonal elements in Tables VIII and IX. The enhanced screening of the eigenpotential over the simple plane wave is clear in so far as the eigenvalues are smaller. Here the trend across the four materials is also apparent. The largest shift is seen for LiCl and C with substantially smaller fractional changes for Si and Ge. Screening for these momentum transfers is also somewhat less effective for Ge than for Si, although Ge exhibits larger shifts indicative of larger local fields. These results give a concise summary of local-field effects in these materials.

These points are illustrated further in Fig. 1. The dielectric band structure for $\tilde{\epsilon}^{-1}$ is plotted along two symmetry directions for Si. The eigenvalues of $\tilde{\epsilon}_{\mathrm{TC}}^{-1}$ were obtained for $\mathbf{q}=\mathbf{q}_L$, $\mathbf{q}=\mathbf{q}_\Gamma$ and $\mathbf{q}=\mathbf{q}_X$ as well as points $\frac{1}{3}$, $\frac{1}{2}$, and $\frac{2}{3}$ of the way in between each. The results for the lowest bands were connected smoothly by a Fourier fit according to symmetry and are displayed in Fig. 1(b). For comparison, the diagonal elements $\tilde{\epsilon}_{\mathbf{GG}}^{-1}(\mathbf{q})$ are plotted according to $\mathbf{G}$ for the first few stars in Fig. 1(a). These results are evaluated for the same points indicated and have been connected by a polynomial fit. This is the analogue of a free-electron band structure. Note that the inverse dielectric bands are bounded by unity and that those bands near 1 are not shown. Comparison of the two figures illustrates how the local fields split and shift the eigenvalues of $\tilde{\epsilon}^{-1}$. The off-diagonal elements are evidently relatively strong. The potentials that are optimally screened are rather different from a single plane wave having several higher $\mathbf{G}$ components. Also, the eigenvalue of $\tilde{\epsilon}^{-1}$ is a considerably weaker function of $\mathbf{q}$ than the corresponding diagonal element (or the macroscopic dielectric function).

C. Macroscopic dielectric function

A macroscopic probe is sensitive to the whole dielectric matrix, but only through its influence on the macroscopic dielectric function defined by Eq. (15). In particular, $\epsilon_0=\epsilon_M(\mathbf{q}\rightarrow\mathbf{0})$ is the only experimental result to which the present results can be directly compared. The macroscopic dielectric constants as calculated here are compared to experiment[48] in Table XII. With reference to Table II, it is seen that local fields decrease ϵ_0 by about 10 to 15%. The origin of the reduction is apparent from Eq. (B11). Furthermore, inclusion of exchange-correlation effects increases the dielectric constant[49] by about 5%. Finally, in comparison to experiment, the density-functional result in

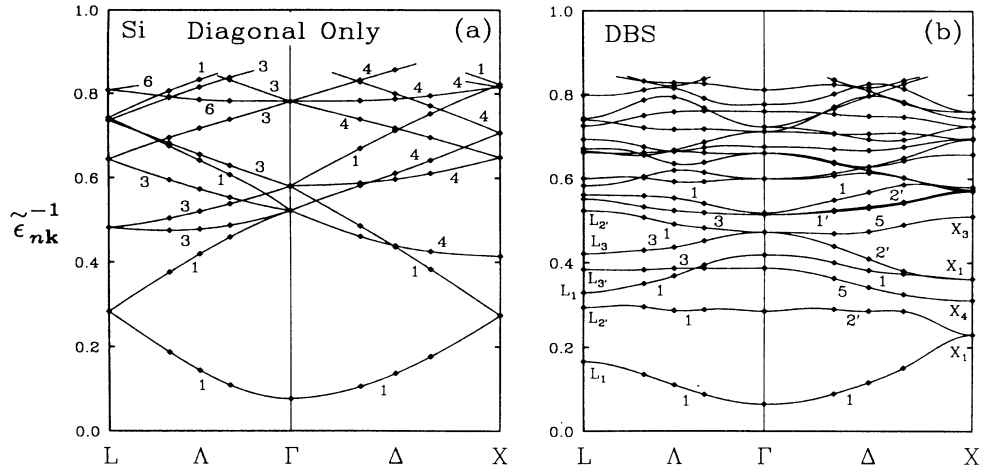

FIG. 1. The lowest bands in the dielectric band structure of Si are plotted along the symmetry directions Λ from L to Γ and Δ from Γ to X for $\tilde{\epsilon}^{-1}$ calculated including exchange-correlation effects. Part (a) corresponds to the diagonal elements only of $\tilde{\epsilon}^{-1}$. Degeneracies are noted. Part (b) shows the results of diagonalizing the full inverse dielectric matrix. Symmetry designations are noted.

the LDA is about 10% too large for diamond, Si, and LiCl. All these results are in close agreement with the recent work of Baroni and Resta for Si.[50] We also note that it is consistent with the atomic calculations of Zangwell and Soven where the calculated static polarizability of the atom is too large in comparison to experiment.[51]

Our results for Ge are about 25% too large. We suggest that Ge is a particularly sensitive case. It has already been shown that inclusion of scalar relativistic effects dramatically reduces the calculated LDA gap.[52] In the present fully converged calculation, the direct gap in Ge at the Γ point is of order 0.05 eV. The LDA conduction and valence bands thus overlap upon inclusion of the spin-orbit splitting. In this situation, even small errors associated with the LDA eigenvalues in comparison to the exact density-functional eigenvalues will have a magnified effect. This is a problem only for the head of the dielectric matrix in the limit of small $\mathbf{q}$. Outside a small region around Γ, the response function is relatively insensitive to the smallest gaps. Previous calculations[31,32] for Ge and GaAs were not based on *ab initio* pseudopotentials and further neglected scalar relativistic effects. As a result, they should be considered less representative of the LDA dielectric constant for these materials.

TABLE XII. The macroscopic dielectric constant as calculated in the RPA and including exchange-correlation effects in the LDA are compared to experiment. Only the electronic part of the dielectric constant (ϵ_∞) is considered for the case of LiCl.

	RPA	TC (LDA)	Expt.[a]
Diamond	5.62	5.90	5.5
Si	12.2	12.9	11.7
Ge	19.2	20.7	15.8
LiCl	2.90	3.07	2.7

[a]Reference 48.

The role of local fields in ϵ_M is further illustrated in Fig. 2. Here $\epsilon_M(\mathbf{q}+\mathbf{G})$ is plotted against $|\mathbf{q}+\mathbf{G}|/q_F$ where q_F is the Fermi wave vector taken for the average density. For comparison, $\epsilon_{GG}(\mathbf{q})$ is also plotted. This corresponds to neglecting the influence of local fields through neglect of the necessary matrix inversion. The test-charge response function is shown for diamond, Si, Ge, and LiCl. The experimental value for ϵ_0 is indicated by the arrow in each case. This illustrates graphically that the calculated dielectric constant is too large. It is also clear that local fields are quantitatively quite important, not only for the macroscopic dielectric constant in the limit of $\mathbf{q}\to\mathbf{0}$, but also for finite $\mathbf{q}$. The effect of local fields is discernible for $\mathbf{q}$ up to roughly the zone edge (q_F) and even beyond in the case of LiCl. For shorter-wavelength perturbations, the screening is weak and the generation of a response at other wavelengths is relatively unimportant. The local fields are seen to be more important for diamond and LiCl than for Si and Ge, as one intuitively expects. Another interesting feature of Fig. 2 is the degree of scatter in the plotted points which are shown for the sets of discrete $\mathbf{q}$ available (8 or 16 $\mathbf{q}$ in the irreducible wedge of the Brillouin zone). The deviations from a smooth curve for small wave vector reflect the degree of directionality of the screening response. This can be of order 15% for some of the data shown. The dependence on orientation becomes negligible outside the first Brillouin zone.

The effect of exchange correlation on the macroscopic dielectric function is shown in Fig. 3 for the four materials considered here. The relative contribution of exchange and correlation given by $(\epsilon_{TC}-\epsilon_{RPA})/\epsilon_{TC}$ is plotted as a function of $|\mathbf{q}+\mathbf{G}|/q_F$. Exchange-correlation effects are important for wave vectors up to of order $2q_F$, with the largest contribution coming in the first Brillouin zone. However, the peak in the relative contribution comes at an

intermediate **q**. The magnitude of the exchange-correlation effects is larger in Si and Ge than in diamond and LiCl. This is consistent with the lower average density in these materials and the fact that, in the LDA, K_{xc} is proportional to $\rho^{-2/3}$ (exchange part). The scatter in the plots is indicative of a directional dependence comparable in magnitude to that found in $\epsilon_M(\mathbf{q}+\mathbf{G})$ itself.

V. CONCLUDING REMARKS

We have shown that it is quite tractable to calculate *ab initio* static dielectric matrices on a regular grid of **q** points in the Brillouin zone for semiconductors and insulators. The present approach is well founded from a theoretical point of view using the density-functional theory. Exchange-correlation effects are explicitly included, with the only significant approximation being the use of the LDA. Technically, use of the *ab initio* pseudopotential band-structure technique allows quite accurate calculations to be done. The present approach has the added advantage over other approaches of not relying on empirical input.[53]

The resulting dielectric matrices have been illustrated in several ways. Local fields and exchange-correlation effects were shown to be quite important in the final dielectric matrix and its screening properties. The calculated macroscopic dielectric constant is found to be generally about 10% larger than experiment. This is indicative of the sensitivity of ϵ_0 to the approximations made, primarily, we believe, the LDA. The role of local fields in the screening of various perturbations will be further illustrated in the following paper.

We suggest that the comparison of the present results to the EPM-based calculations in Sec. IV A and experiment

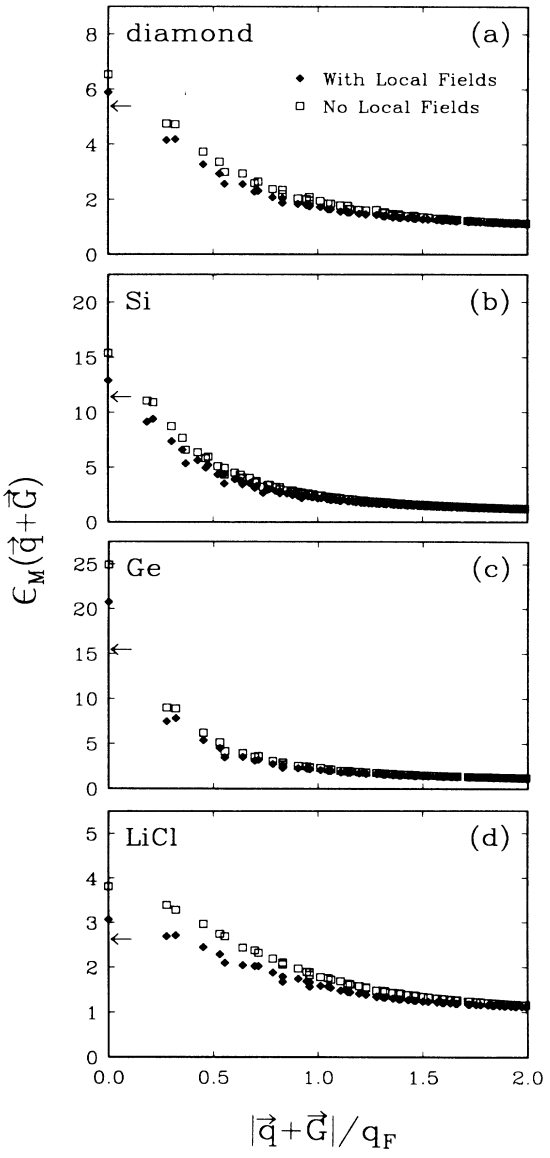

FIG. 2. The macroscopic dielectric function $\epsilon_M(\mathbf{q}+\mathbf{G})$ is plotted as a function of $|\mathbf{q}+\mathbf{G}|/q_F$ (solid symbols). For comparison, $\epsilon_{GG}(\mathbf{q})$ is also shown (open symbols). Results are for the TC response function in this case and are displayed for (a) diamond, (b) Si, (c) Ge, and (d) LiCl. The experimental value of ϵ_0 is indicated by the arrow in each case.

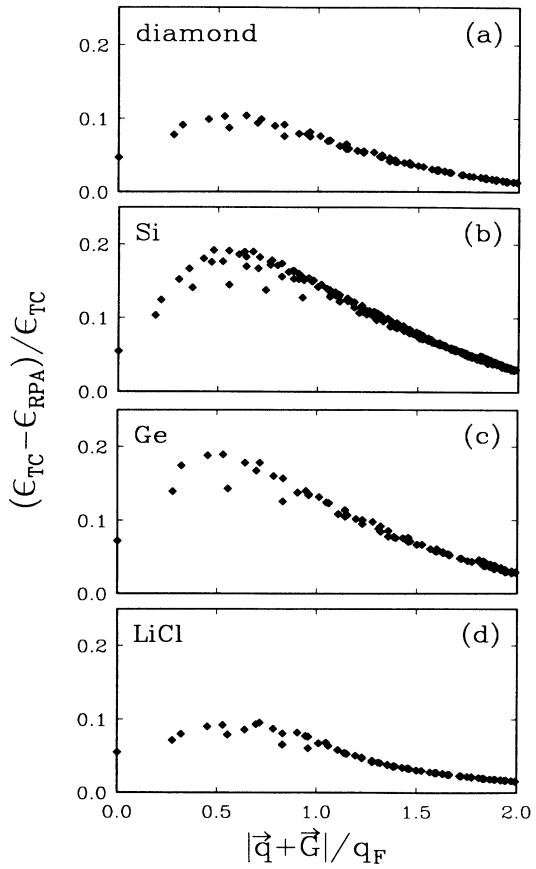

FIG. 3. The exchange-correlation contribution is displayed by plotting $[\epsilon_{TC}(\mathbf{q}+\mathbf{G})-\epsilon_{RPA}(\mathbf{q}+\mathbf{G})]/\epsilon_{TC}(\mathbf{q}+\mathbf{G})$ as a function of $|\mathbf{q}+\mathbf{G}|/q_F$. The macroscopic dielectric function is used so that local-field effects are included. Results are displayed for (a) diamond, (b) Si, (c) Ge, and (d) LiCl.

in Sec. IV C indicate that the LDA overestimates screening for small momentum transfer. Similar conclusions follow upon comparing ϵ_M to simple models fitted to screening for larger $\mathbf{q}$. However, as the discussion in Sec. IV A indicates, care must be taken in making such comparisons. We also feel that this discrepancy is relatively small and that for many applications, use of the LDA to generate dielectric matrices from first principles is the method of choice at this time.

The possibility of calculating the dielectric matrices for a reasonable grid of $\mathbf{q}$ points has already proved crucial in our recently presented theory of quasiparticle energies in semiconductors and insulators.[5] That application required extending the static results to finite frequency using a generalized plasmon pole model. Such a model may be useful in other applications as well. We expect that availability of well-founded static dielectric matrices will be important in many other problems. The calculation of phonon frequencies is one well-known application. Others include investigating systematics of shallow impurity levels in semiconductors and calculating electron-phonon interactions.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation Grant No. DMR-83-19024. The work was facilitated by the IBM Distributed Academic Computing Environment at the University of California, Berkeley. One of us (S.G.L.) wishes to acknowledge support by the Miller Institute.

APPENDIX A: REDUCTION OF BRILLOUIN-ZONE SUMMATIONS

The sum on $\mathbf{k}$ in Eq. (18) includes the full Brillouin zone. The degree to which it can be reduced depends on the symmetry operations in the little group of $\mathbf{q}$. These are defined by

$$\underline{R}\mathbf{q}=\mathbf{q}+\mathbf{G}_R \ . \tag{A1}$$

An *umklapp* may be required for some operations if $\mathbf{q}$ is on the zone boundary. These operations can be used to define an irreducible wedge of the Brillouin zone, BZ_q. The idea is to break the sum on $\mathbf{k}$ into a sum on $\mathbf{k}$ in BZ_q only and a complementary sum over a set of symmetry operations from the little group which regenerate the star of $\mathbf{k}$ in the full Brillouin zone. Thus we have, schematically,

$$\sum_{\mathbf{k}\in \mathrm{BZ}} f(\mathbf{k})=\sum_{\mathbf{k}\in \mathrm{BZ}_q}\sum_{\underline{R}} f(\underline{R}^{-1}\mathbf{k}) \ . \tag{A2}$$

Here, the operations are chosen such that $\underline{R}^{-1}$ generates the star of $\mathbf{k}$ for convenience below. Now, the energies in the denominator in Eq. (18) also have the symmetry of $\mathbf{q}$:

$$\begin{aligned} &\varepsilon_{v,\underline{R}^{-1}\mathbf{k}}=\epsilon_{v,\mathbf{k}} \ , \\ &\varepsilon_{c,\underline{R}^{-1}\mathbf{k}+\mathbf{q}}=\varepsilon_{c,\underline{R}^{-1}(\mathbf{k}+\mathbf{q})-\underline{R}^{-1}\mathbf{G}_R}=\varepsilon_{c,\mathbf{k}+\mathbf{q}} \ . \end{aligned} \tag{A3}$$

The wave functions $\phi_{v,\underline{R}^{-1}\mathbf{k}}$ are also required. In Eq. (18) each wave function enters in a product with its complex conjugate. Therefore, the *overall* phase of the wave function has no effect. Furthermore, all the states are summed over, valence- and conduction-band states separately. In particular, any degenerate complexes of states are summed over. This is important because in generating wave functions for $\mathbf{k}$ related by symmetry, the action of the symmetry operation in the space spanned by the degenerate wave functions must in general be included. However, since the representation matrices may be assumed to be unitary, the states enter in complex conjugate products and the states in the degenerate complex are summed, the representation matrices drop out of the final expression. For these reasons, we may take

$$\phi_{v,\underline{R}^{-1}\mathbf{k}}(\mathbf{r})=\phi_{v,\mathbf{k}}(\underline{R}\mathbf{r}+\tau_R) \ . \tag{A4}$$

The associated nonprimitive translation vector enters on the right. Because we use operations satisfying Eq. (A1), the same can be done for $\phi_{c,\underline{R}^{-1}\mathbf{k}+\mathbf{q}}$.

Using the relations in Eq. (A4), the plane-wave matrix elements required in Eq. (18) may be straightforwardly manipulated to obtain the necessary phases that enter. The final expression is

$$\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q})=\frac{4}{N\Omega_c}\sum_{c,v}\sum_{\mathbf{k}\in \mathrm{BZ}_q}(\varepsilon_{v\mathbf{k}}-\varepsilon_{c\mathbf{k}+\mathbf{q}})^{-1} \times\sum_{\underline{R}\in G_k} e^{i(\underline{R}\mathbf{G}-\underline{R}\mathbf{G}')\cdot\tau_R}\langle v,\mathbf{k}\,|\,e^{-i(\mathbf{q}+\underline{R}\mathbf{G}+\mathbf{G}_R)\cdot\mathbf{r}}\,|\,c,\mathbf{k}+\mathbf{q}\rangle\langle c,\mathbf{k}+\mathbf{q}\,|\,e^{i(\mathbf{q}+\underline{R}\mathbf{G}'+\mathbf{G}_R)\cdot\mathbf{r}'}\,|\,v,\mathbf{k}\rangle \ . \tag{A5}$$

Here, N is the number of points in the summation grid in the full Brillouin zone and Ω_c is the primitive cell volume. The key point is that the plane-wave matrix elements need only be computed once for each $\mathbf{k}\in \mathrm{BZ}_q$ over the $\mathbf{G}$ spanned by the size of χ_0. The accumulation of the sum over symmetry operations only requires including the correct phase and mapping to the correct matrix element.

In Eq. (A5) wave functions for $\mathbf{k}$ outside the usual irreducible wedge of the Brillouin zone are required. These can be easily generated by symmetry using Eq. (A4) with symmetry operations taken as needed from the full space group. As discussed in the text, the calculation of the necessary plane-wave matrix elements has dominated the computer time required. One way to ameliorate that is to observe that the important components of the valence-band wave functions are the first few plane waves. The

large number of plane waves included in the Hamiltonian matrix is required to achieve a well-converged band structure and represent the necessary high conduction-band states. Therefore, we have found that the valence-band wave functions can effectively be truncated by a much lower cutoff at the time the plane-wave matrix elements are calculated. For example, in Si, the valence-band wave functions were truncated with $E_{\max}$ of 7 Ry in comparison to 14 Ry in the band calculation with essentially no change in the resulting polarizability. This gives a factor of 3 in the time required in the matrix element calculation. One could also exploit the fact that the high-conduction-band states should only have a few plane-wave components about a dominant wave vector. We have not explored that further. Finally, it has been pointed out that a fast Fourier-transform technique can be applied to calculation of the plane-wave matrix elements.[46]

APPENDIX B: SPECIAL CASE OF $\mathbf{q}\rightarrow\mathbf{0}$

In this appendix, the special considerations required for the calculation of $\epsilon^{-1}(\mathbf{q}\rightarrow\mathbf{0})$ are outlined.

In Eq. (18) for χ_0, special care must be taken with the matrix elements required when $\mathbf{G}$ or $\mathbf{G}'$ are zero. Otherwise, $\mathbf{q}$ can be taken to be exactly zero. Using first-order perturbation theory for $\mathbf{q}$ small, the wave functions for $|c,\mathbf{k}+\mathbf{q}\rangle$ can be obtained in terms of those for $|n,\mathbf{k}\rangle$. The nonlocal part of the ionic pseudopotential enters explicitly as it does not commute with local functions of $\mathbf{r}$. The result is

$$\langle v,\mathbf{k}|e^{-i\mathbf{q}\cdot\mathbf{r}}|c,\mathbf{k}+\mathbf{q}\rangle = \frac{\langle v,\mathbf{k}|-2i\mathbf{q}\cdot\nabla_{\mathbf{r}}+[V_{\mathrm{NL}},i\mathbf{q}\cdot\mathbf{r}]|c,\mathbf{k}\rangle}{\varepsilon_{c,\mathbf{k}}-\varepsilon_{v,\mathbf{k}}}, \tag{B1}$$

to lowest order in $\mathbf{q}$. The commutator of the nonlocal part of the pseudopotential V_{NL} with $\mathbf{r}$ enters explicitly.[54] Using this result, the polarizability can be conveniently broken down into three types of terms corresponding to the dependence on $\mathbf{q}$. The "head" of the matrix is given, through Eq. (18), by

$$\chi^0_{00}(\mathbf{q}\rightarrow\mathbf{0})=\frac{4}{\Omega}\sum_{v,c,\mathbf{k}}\frac{1}{(\varepsilon_{v,\mathbf{k}}-\varepsilon_{c,\mathbf{k}})^3}\langle v,\mathbf{k}|-2i\mathbf{q}\cdot\nabla_{\mathbf{r}}+[V_{\mathrm{NL}},i\mathbf{q}\cdot\mathbf{r}]|c,\mathbf{k}\rangle\langle c,\mathbf{k}|-2i\mathbf{q}\cdot\nabla_{\mathbf{r}}+[V_{\mathrm{NL}},i\mathbf{q}\cdot\mathbf{r}]|v,\mathbf{k}\rangle . \tag{B2}$$

The head of χ_0 goes to zero as q^2. The "wings" of the matrix are given by

$$\chi^0_{0\mathbf{G}'}(\mathbf{q}\rightarrow\mathbf{0})=-\frac{4}{\Omega}\sum_{v,c,\mathbf{k}}\frac{1}{(\varepsilon_{v,\mathbf{k}}-\varepsilon_{c,\mathbf{k}})^2}\langle v,\mathbf{k}|-2i\mathbf{q}\cdot\nabla_{\mathbf{r}}+[V_{\mathrm{NL}},i\mathbf{q}\cdot\mathbf{r}]|c,\mathbf{k}\rangle\langle c,\mathbf{k}|e^{i\mathbf{G}'\cdot\mathbf{r}'}|v,\mathbf{k}\rangle . \tag{B3}$$

There is a complementary expression for $\chi^0_{\mathbf{G}0}(\mathbf{q}\rightarrow\mathbf{0})$. The wings of χ_0 go to zero as q. Finally, the "body" of the matrix is given straightforwardly from Eq. (18) as

$$\chi^0_{\mathbf{G}\mathbf{G}'}(\mathbf{q}\rightarrow\mathbf{0}) = \frac{4}{\Omega}\sum_{v,c,\mathbf{k}}\frac{\langle v,\mathbf{k}|e^{-i\mathbf{G}\cdot\mathbf{r}}|c,\mathbf{k}\rangle\langle c,\mathbf{k}|e^{i\mathbf{G}'\cdot\mathbf{r}'}|v,\mathbf{k}\rangle}{\varepsilon_{v,\mathbf{k}}-\varepsilon_{c,\mathbf{k}}} . \tag{B4}$$

From Eq. (B4) it is apparent that the body of χ_0 has the symmetry of $\mathbf{q}=\mathbf{0}$ where as the wings and head have the symmetry associated with the direction along which $\mathbf{q}\rightarrow\mathbf{0}$. Furthermore, the wings and head depend on that direction, in general. For cubic symmetry, such as the cases considered here, the head of the matrix is independent of the orientation of $\mathbf{q}$. The wings, however, do depend on orientation. This simply expresses the dependence of the microscopic screening of a constant electric field on the direction of the applied field. The special convergence properties of the head and wings of the matrix have been discussed by Baldereschi and Tosatti,[15] and Baroni and Resta[22] and are illustrated in Sec. III B.

In the present plane-wave basis, the first part of the new matrix elements required in Eq. (B1) is straightforward to evaluate. For reference, we outline the part depending on the nonlocal pseudopotential as it is more involved. The simplest approach is to observe that the required matrix element between plane waves can be written as

$$\langle\mathbf{K}|[V_{\mathrm{NL}},i\mathbf{q}\cdot\mathbf{r}]|\mathbf{K}'\rangle=(\mathbf{q}\cdot\nabla_{\mathbf{K}}+\mathbf{q}\cdot\nabla_{\mathbf{K}'})V_{\mathrm{NL}}(\mathbf{K},\mathbf{K}') , \tag{B5}$$

where $\mathbf{K}=\mathbf{k}+\mathbf{G}$ for brevity and

$$V_{\mathrm{NL}}(\mathbf{K},\mathbf{K}')=\langle\mathbf{K}|V_{\mathrm{NL}}|\mathbf{K}'\rangle . \tag{B6}$$

The nonlocal part of the pseudopotential is written in the form

$$V_{\mathrm{NL}}(\mathbf{r})=\sum_{i,j}\sum_{l}V_l(|\mathbf{r}-\mathbf{R}_i-\tau_j|)\hat{P}_l(\mathbf{r}-\mathbf{R}_i-\tau_j) . \tag{B7}$$

The vector $\mathbf{R}_i+\tau_j$ locates atom j in cell i, l runs over the angular momentum quantum number on a given site, and $\hat{P}_l$ projects out the l angular component around the site indicated. The matrix elements indicated in Eq. (B6) are exactly those required in the band-structure calculation. When Eq. (B5) is worked out, the result is

$$\langle \mathbf{K} | [V_{\rm NL}, i\mathbf{q}\cdot\mathbf{r}] | \mathbf{K}' \rangle = \frac{1}{\Omega_c} \sum_j e^{i(\mathbf{K}'-\mathbf{K})\cdot\tau_j} \mathbf{q}\cdot\sum_l \left\{ \frac{P_l'(\hat{\mathbf{K}}\cdot\hat{\mathbf{K}}')}{KK'} \left[\mathbf{K}\left[1-\frac{\mathbf{K}\cdot\mathbf{K}'}{K^2}\right] + \mathbf{K}'\left[1-\frac{\mathbf{K}\cdot\mathbf{K}'}{(K')^2}\right]\right] V_l(K,K') \right.$$
$$\left. + P_l(\hat{\mathbf{K}}\cdot\hat{\mathbf{K}}') \left[\frac{\partial V_l(q,q')}{\partial q}\bigg|_{K,K'} \frac{\mathbf{K}}{K} + \frac{\partial V_l(q,q')}{\partial q'}\bigg|_{K,K'} \frac{\mathbf{K}'}{K'} \right] \right\} . \quad \text{(B8)}$$

Here, P_l is the l Legendre polynomial and P_l' its first derivative. The usual Fourier transform of V_l is given by

$$V_l(q,q') = 4\pi(2l+1) \int dr\, r^2 j_l(qr) j_l(q'r) V_l(|\mathbf{r}|) , \quad \text{(B9)}$$

where j_l are the spherical Bessel functions. The required derivatives can be similarly formulated, although we have found that a finite difference approach is adequate for a reasonably fine grid in q,q'. The expression (B8) is symmetric as one would expect. Finally, we comment that proper account of the nonlocal pseudopotential is essential as it contributes of order 5—10 % to the matrix element in Eq. (B1).

The final issue to be addressed is the proper account of the nonanalyticity of ϵ^{-1}. A complete general account is given in Ref. 1. Here we indicate how to handle the specific case of the test-charge response function through Eq. (11) as explicitly written out in Eqs. (21)—(23).

Assume that $\mathbf{q}\to\mathbf{0}$ along some given direction. For that case, the polarizability has the form

$$\chi_0 = \begin{bmatrix} q^2\chi^0_{00} & q\chi^0_{01} \\ q\chi^0_{10} & \chi^0_{11} \end{bmatrix} . \quad \text{(B10)}$$

Here the head, wings, and body of the matrix are indicated schematically. The matrix A defined by Eq. (21) must be inverted according to the prescription given in Ref. 1. Then when A is combined with χ_0 and the Coulomb interaction in Eqs. (22) and (23), ϵ^{-1} has the correct analytic properties. Here we focus on the resulting macroscopic dielectric constant to show how exchange-correlation effects enter in the LDA. The result is that properly accounting for the singularities at each stage yields

$$\epsilon_0^{\rm TC} = 1 - 4\pi e^2\chi^0_{00} - 4\pi e^2\chi^0_{01}(V^C_{11}+K^{\rm xc}_{11}) \times [1-\chi^0_{11}(V^C_{11}+K^{\rm xc}_{11})]^{-1}\chi^0_{10} . \quad \text{(B11)}$$

The matrix operations are indicated in blocks according to the scheme in Eq. (B10). In particular, V^C_{11} is the nonsingular part of the Coulomb kernel and is diagonal as usual. When $K^{\rm xc}$ is neglected, Eq. (B11) reverts to the usual expression for the effect of local fields on the macroscopic dielectric constant in the RPA. With the inclusion of exchange-correlation effects in the LDA, the dielectric constant is shifted to larger values as shown in Sec. IV.

[1] R. M. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B **1**, 910 (1970).
[2] L. J. Sham, Phys. Rev. **188**, 1431 (1969).
[3] R. Car and A. Selloni, Phys. Rev. Lett. **40**, 1365 (1979).
[4] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. Lett. **45**, 290 (1980); Phys. Rev. B **25**, 2867 (1982).
[5] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985); Phys. Rev. B **32**, 7005 (1985); **34**, 5390 (1986).
[6] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[7] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[8] Several models for the dielectric matrix are evaluated in R. Car, E. Tosatti, S. Baroni, and S. Leelaprute, Phys. Rev. B **24**, 985 (1981).
[9] J. C. Phillips, Phys. Rev. **166**, 832 (1968).
[10] R. M. Martin, Phys. Rev. **186**, 871 (1969).
[11] J. A. Van Vechten and R. M. Martin, Phys. Rev. Lett. **28**, 446, 646(E) (1972).
[12] S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. Lett. **34**, 155 (1975).
[13] W. Hanke and L. J. Sham, Phys. Rev. Lett. **33**, 582 (1974); Phys. Rev. B **12**, 4501 (1975).
[14] W. Hanke and L. J. Sham, Phys. Rev. Lett. **43**, 387 (1979); Phys. Rev. B **21**, 4656 (1980).
[15] A. Baldereschi and E. Tosatti, Phys. Rev. B **17**, 4710 (1978).
[16] R. Resta and A. Baldereschi, Phys. Rev. B **23**, 6615 (1981).
[17] A. Fleszar and R. Resta, Phys. Rev. B **31**, 5305 (1985).
[18] K. Kunc and E. Tosatti, Phys. Rev. B **29**, 7045 (1984).
[19] S. L. Adler, Phys. Rev. **126**, 413 (1962).
[20] N. Wiser, Phys. Rev. **129**, 62 (1963).
[21] D. R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
[22] S. Baroni and R. Resta, Phys. Rev. B **33**, 7017 (1986).
[23] For a review, see, e.g., S. G. Louie, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, edited by J. Devreese and P. van Camp (Plenum, New York, 1985), p. 335.
[24] A. Baldereschi and E. Tosatti, Solid State Commun. **29**, 131 (1979).
[25] Specific references will be given where appropriate. A broad overview and some specific issues are covered by A. R. Williams and U. von Barth, in *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York, 1983), p. 189.
[26] A. Berera, M. S. Hybertsen, and S. G. Louie (unpublished).
[27] Wendel and R. M. Martin, Phys. Rev. B **19**, 5251 (1979).
[28] This type of derivation is given by Williams and von Barth as cited in Ref. 25 as well as by P. Singhal and J. Callaway, Phys. Rev. B **14**, 2347 (1976).

[29]A straightforward discussion of the distinct response functions is given in S. P. Singhal and J. Callaway, Phys. Rev. B **14**, 2347 (1976). See also L. Kleinman, Phys. Rev. **172**, 383 (1968).

[30]In the electron gas K_{xc} is a strong function of $\mathbf{q}$, whereas in the LDA, it is independent of $\mathbf{q}$ for that case. The weighted density approximation restores the $\mathbf{q}$ dependence to K_{xc}. This is illustrated explicitly by Williams and von Barth in Ref. 25.

[31]K. Kunc and R. Resta, Phys. Rev. Lett. **51**, 686 (1983).

[32]J. B. McKitterick, Phys. Rev. B **28**, 7384 (1983).

[33]Strict numerical convergence of the dielectric matrix may not be required for all elements of the matrix in some applications. In this case, the stringent requirements on the sum over empty states may be relaxed.

[34]Similar symmetry relations based on a different convention are given by P. E. Van Camp, V. E. Van Doren, and J. T. Devreese in *Ab initio Calculation of Phonon Spectra,* edited by J. T. Devreese, V. E. Van Doren, and P. E. Van Camp (Plenum, New York, 1983), p. 25.

[35]The algorithm used is from G. P. Kerker, J. Phys. C **13**, L189 (1980).

[36]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).

[37]D. M. Ceperley and B. I. Alder, Phys. Rev. Lett. **45**, 566 (1980), as parametrized in J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[38]L. Kleinman, Phys. Rev. B **21**, 2630 (1980); G. B. Bachelet and M. Schluter, *ibid.* **25**, 2103 (1982). We use the algorithm from Ref. 35.

[39]S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).

[40]The use of the core correction in the pseudopotential was found to be essential for obtaining good structural properties in NaCl; S. Froyen and M. L. Cohen, Phys. Rev. B **29**, 3770 (1984).

[41]D. J. Chadi and M. L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[42]For simplicity, $\mathbf{k}$-point sets will be referred to by the number of points in the usual irreducible wedge of the Brillouin zone, with the understanding that the actual number of points summed over depends on the contents of BZ_q.

[43]As pointed out in Ref. 15, for $\mathbf{G},\mathbf{G}'$ both nonzero, $\chi_0(\mathbf{q}\rightarrow\mathbf{0})$ has the symmetry of $\mathbf{q}=\mathbf{0}$. However, when the inverse dielectric matrix is formed, nonanalytic contributions appear throughout and the whole inverse matrix has the symmetry associated with the orientation of $\mathbf{q}$.

[44]This just distinguishes the microscopic response to an external $\mathbf{E}$ field in different directions, e.g., parallel versus perpendicular to a bond.

[45]See, for example, Resta and Baldereschi in Ref. 16.

[46]A. Fleszar, Ph.D. thesis, ISSA, University of Trieste, 1985 (unpublished).

[47]Baroni and Resta have given a different argument for the sign of the exchange-correlation effect in Ref. 22.

[48]C. Kittel, *Introduction to Solid State Physics,* 5th ed. (Wiley, New York, 1976), p. 309. The data given there are a good representation of the available experimental data as surveyed in *Landolt-Bornstein: Zahlenwerte und Funktionen aus Naturwissenschaften und Technik* (Springer, New York, 1982), Vol. III, Pt. 17a.

[49]This is contrary to previous results which indicated that the exchange-correlation contribution vanished as $\mathbf{q}\rightarrow\mathbf{0}$. P. E. Van Camp, V. E. Van Doren, and J. T. Devreese, Phys. Rev. B **24**, 1096 (1981).

[50]The small numerical difference between the present results for Si and those given in Ref. 22 is attributable to the larger energy cutoff used in the present work (17 versus 14 Ry in Ref. 22). This illustrates the sensitivity of the dielectric constant to numerical convergence.

[51]A. Zangwell and P. Soven, Phys. Rev. A **21**, 1561 (1980).

[52]G. B. Bachelet and N. E. Christensen, Phys. Rev. B **31**, 879 (1985).

[53]For example, the approach in Refs. 4, 15, and 16. See also H. J. Mattausch, W. Hanke, and G. Strinati, Phys. Rev. B **27**, 3735 (1983).

[54]This is derived somewhat differently by Baroni and Resta in Ref. 22.