VOLUME 58, NUMBER 18 PHYSICAL REVIEW LETTERS 4 MAY 1987

Green's-Function Approach to Linear Response in Solids

Stefano Baroni
Dipartimento di Fisica Teorica, Università di Trieste, I-34014 Trieste, Italy

Paolo Giannozzi
Institut de Physique Théorique, Université de Lausanne, BSP-Dorigny, CH-1015 Lausanne, Switzerland

and

Andrea Testa
Institut de Physique Théorique, Ecole Polytechnique Fédérale de Lausanne, PHB-Ecublens, CH-1015 Lausanne, Switzerland

(Received 24 February 1987)

We present a new scheme to study the linear response of crystals which combines the advantages of the dielectric-matrix and supercell ("direct") approaches yet avoids many of their drawbacks. The numerical complexity of the algorithm is of the same order as that of a self-consistent calculation for the *unperturbed* system. The method is not restricted to local perturbations as is the dielectric-matrix one nor to short wavelength as is the direct one. As an application, we calculate the long-wavelength optical phonons in Si and GaAs, both transverse and longitudinal, using norm-conserving pseudopotentials, and without any use of supercells.

PACS numbers: 63.10.+a, 71.10.+x, 77.90.+k

The availability of accurate first-principles techniques to study the electronic structure of crystals has renewed the interest in the study of their response to external perturbations. These include impurities and phonons,[1] as well as macroscopic perturbations such as homogeneous strains[2] or electric fields.[3,4]

Two approaches are currently in use to cope with external perturbations in periodic systems. In the first one, which is often referred to as the "direct" method,[5] the perturbed system is handled on the same footing as the unperturbed one, and the response is obtained from the comparison of the properties of the two systems. In the second approach, the response to perturbations is given by the inverse of the dielectric matrix (DM)[1] which is obtained from the eigenfunctions and energy levels of the unperturbed system. The relative merits of the two methods can be summarized as follows.

The direct approach allows us to study both linear and nonlinear effects; linear effects, however, are not directly accessible, and some interpolation scheme must be used to extract them. This method is rather straightforward computationally; in fact, total-energy computer codes written for perfect crystals can be easily extended to deal with supercell geometries which allow an explicit account of this perturbation. As a consequence, no particular problem arises when the perturbation is described by a nonlocal potential. Another advantage is that, since only the knowledge of valence states is required, the self-consistent calculation can be carried out by efficient iterative-diagonalization[6] or global-optimization[7] techniques. These techniques allow us to keep the numerical complexity of the problem at the level of $\sim MN^2$ floating-point operations (FPO),[8] where M is the number of electrons per unit (super)cell, and N is the size of the electron-wave-function basis set. The main limitation is that, in order to maintain the numerical effort at a reasonable level, only those perturbations which can be described by supercells as small as a few elementary cells are accessible to this technique. In simple semiconductors, long-wavelength perturbations such as longitudinal phonons and/or macroscopic electric fields have been recently treated within the direct approach[4,5]; the price to be paid, however, is to simulate their effect by use of large supercells. Large supercells are a source of numerical errors and constitute a severe limitation in the applicability of this method to systems different from the simplest semiconductors.

In the DM approach only quantities from the unperturbed ground state are necessary to calculate the response. As a consequence, the method is not restricted to perturbations of any given periodicity, since the calculation of DM's at any point of the Brillouin zone (BZ) is not much more complex than that at the Γ point. Of course, only linear effects are accessible to this technique: This is not an inconvenience in many instances, such as, e.g., lattice dynamics, where experimentally measured quantities can be exactly expressed in terms of linear-response functions. The real limitation of this approach is that the calculation of dielectric matrices requires cumbersome summations over conduction bands at several $\mathbf{k}$ points in the BZ. Furthermore, the fact that the *whole* spectrum of the unperturbed self-consistent Hamiltonian is needed makes iterative techniques to diagonalize matrices useless, and forces us to use more conventional tridiagonalization-bisection algorithms which require a number of FPO's $\sim N^3$.[9] The

© 1987 The American Physical Society

subsequent inversion of the DM needs then an analogous numerical work.[9] Since modern norm-conserving pseudopotentials are relatively hard core,[10] the number of plane waves necessary to describe even a simple semiconductor is as high as some hundreds: This implies that a technique allowing us to keep the numerical labor at the level of $\sim N^2$ FPO's would be of extreme value. Another drawback of the DM approach is that, in order to study the response to one particular perturbation, one has to calculate large DM's which contain far more information than physically needed. Yet the information contained in conventional DM's is not enough to study nonlocal perturbations, and one should, in this case, resort to even more complex response functions.

In this paper we show that the advantages of the above two schemes may be combined in a new method which is the solid-state analog of the self-consistent Sternheimer equation for atomic polarizabilities.[11] Its main features are the following: The response is studied via perturbation theory as in the DM approach; however, the response to the total (bare+electronic) potential is not obtained by inversion of DM's, but by iteration of the calculation up to self-consistency. The time-consuming sums over conduction bands are avoided through a Green's-function technique which keeps the numerical complexity of the calculation at the MN^2 level. The method is not restricted to perturbations written in terms of local potentials as is the case in the traditional DM approach. This is of particular interest in lattice dynamics, when norm-conserving potentials are used to describe the electron-ion interaction. No use of supercells is made, and the calculation can be performed, for perturbations of arbitrary wavelength, with use of only valence energy bands and wave functions of the unperturbed crystal.

Within density-functional theory,[12] the electronic (number) density $n(\mathbf{r})$ is determined by a one-body potential, $V_{\mathrm{SCF}}(\mathbf{r})$, which is the sum of the external potential acting on the electrons, $V_{\mathrm{bare}}(\mathbf{r})$, plus an effective electron-electron potential which depends on the density itself:

$$V_{\mathrm{SCF}}(\mathbf{r}) = V_{\mathrm{bare}}(\mathbf{r}) + e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{XC}(n(\mathbf{r})), \tag{1}$$

where the local-density approximation[12] for exchange and correlation is assumed. When a perturbation ΔV_{bare} is superimposed on the external potential, the self-consistent potential is modified accordingly: $V_{\mathrm{SCF}} \rightarrow V_{\mathrm{SCF}} + \Delta V_{\mathrm{SCF}}$. If ΔV_{SCF} is supposed to be known, the linear variation in the electron density Δn is easily obtained by first-order perturbation theory:

$$\Delta n(\mathbf{q}+\mathbf{G}) = -\frac{4}{V} \sum_{k \in \mathrm{BZ}} \sum_{c,v} \frac{\langle v,\mathbf{k} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | c,\mathbf{k}+\mathbf{q} \rangle \langle c,\mathbf{k}+\mathbf{q} | \Delta V_{\mathrm{SCF}} | v,\mathbf{k} \rangle}{E_c(\mathbf{k}+\mathbf{q}) - E_v(\mathbf{k})}, \tag{2}$$

where $\Delta n(\mathbf{q}+\mathbf{G})$ is the Fourier transform of $\Delta n(\mathbf{r})$, V indicates the crystal volume, and v and c denote the valence and conduction bands, respectively. Of course, Eq. (2) is analogous to the expression for the independent-electron polarizability in the random-phase approximation given by Adler and Wiser.[13] The input to Eq. (2) is obtained by our linearizing Eq. (1),

$$\Delta V_{\mathrm{SCF}}(\mathbf{r}) = \Delta V_{\mathrm{bare}}(\mathbf{r}) + e^2 \int \frac{\Delta n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \Delta n(\mathbf{r}) \left(\frac{dv_{XC}}{dn} \right)_{n=n_0(\mathbf{r})}, \tag{3}$$

where n_0 is the unperturbed electron density. Equations (2) and (3) can be used directly to calculate the self-consistent linear response to a given external perturbation, without inverting nor even calculating any dielectric matrix. Equation (2) is of the same complexity as the evaluation of just one single row (or column) of a DM. In contrast to the DM approach, in the present scheme one calculates just one row of the matrix as many times as necessary to achieve self-consistency. Since the number of iterations is usually far smaller than the dimensions of the DM (a few versus some hundreds), the self-consistent approach is found to be much more convenient. We remark that exchange-correlation effects are easily and naturally treated through Eq. (3), while in the DM approach their inclusion further requires the inversion and multiplication of large matrices.[3]

The sum over conduction states appearing in Eq. (2) requires the calculation of the full spectrum of the unperturbed Hamiltonian. This can be very time-consuming since the complete diagonalization of a matrix of order N necessarily requires $\sim N^3$ FPO's. In order to avoid such a sum, we write Eq. (2) as follows:

$$\Delta n(\mathbf{q}+\mathbf{G}) = -(4/V) \textstyle\sum_{\mathbf{k}} \sum_v \langle v,\mathbf{k} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} P_c G_0(E_v(\mathbf{k})) P_c \Delta V_{\mathrm{SCF}} | v,\mathbf{k} \rangle, \tag{4}$$

where P_c is the projector over the conduction-state manifold, and G_0 is the Green's function of the unperturbed system described by the self-consistent Hamiltonian H_{SCF}. To evaluate Eq. (4), we further rewrite it as

$$\Delta n(\mathbf{q}+\mathbf{G}) = -(4/V) \textstyle\sum_{\mathbf{k}} \sum_v \langle v,\mathbf{k} | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | \tilde{v},\mathbf{k} \rangle, \tag{5a}$$

where

$$[H_{\mathrm{SCF}} - E_v(k)] \,|\, \tilde{v},\mathbf{k}\rangle = P_c \Delta V_{\mathrm{SCF}} \,|\, v,\mathbf{k}\rangle. \tag{5b}$$

In the literature, several algorithms are known to solve iteratively linear systems[9] which require a number of FPO's proportional to the square of the size of the matrix times the number of right-hand sides to be considered,[8] whereas more conventional factorization algorithms[9] require a number of FPO's proportional to the cube of the size of the matrix. Iterative algorithms to solve Eq. (5b) offer the further advantage that, in practical self-consistent calculations, a good guess of the solution is available from the previous self-consistent iteration. We conclude that the self-consistent scheme proposed here allows us to calculate the response of a crystal to external perturbations of *arbitrary wavelength* with a numerical effort growing as the square of the basis set times the number of valence bands of the *unperturbed system.*

Much interest has been devoted in recent years to the *ab initio* calculation of the phonon spectrum of simple semiconductors.[1,5] So far, such calculations have been performed either in the direct or in the DM approach. One of the challenging features of phonon spectra is the effects related to the existence of a macroscopic electric field in the long-wavelength longitudinal modes of polar materials. In the DM approach, such effects can be related to the analytical properties of the dielectric matrix[14] and dealt with directly. Unfortunately, the impossibility of the use of norm-conserving pseudopotentials limits the accuracy achievable by this method. Vice versa, nonlocal pseudopotentials are easily treated in the direct approach, but difficulties arise because the potential associated with a homogeneous electric field does not obey Born–von Kármán boundary conditions. As a consequence, only finite-$\mathbf{q}$ calculations are possible and large supercells are needed to recover the $\mathbf{q} \rightarrow 0$ limit. Because of these difficulties, the theoretical results reported so far for Si and GaAs are rather scattered.[15] In particular, several calculations gave a nonvanishing value for the effective charge of Si and/or a nonzero value for the sum of the effective charges in GaAs.

In the following, we present our results for the long-wavelength optical phonons in Si and GaAs. The calculations are performed by our displacing one atom per unit cell, and screening the bare perturbation along the lines expounded above. Energy differences have been calculated by our expanding the crystal energy up to second order in the atomic displacement.[16] Harmonic forces could be obtained along similar lines. Effective charges are obtained from the screened electric field induced by the phonon. In GaAs this is done by separate calculations of the response to the displacements of each of the two atoms in the unit cell. Macroscopic electric fields are treated in the very same way as in Ref. 3.[17] The ionic pseudopotentials used here are the same as in Ref. 3, for Si, and are taken from Ref. 10 for GaAs. The kinetic-energy cutoffs for the plane-wave basis set are 14 and 16 Ry for Si and GaAs, respectively. The assumed lattice parameters are 10.26 and 10.50 a.u., respectively. Other technical details of the calculation are as in Ref. 3.

In Table I we report the results of our calculations for Si. The entries "10 points" refer to calculations made using the (8,8,8) Monkhorst-Pack mesh for BZ integration,[18] while the entry "28 points" refers to the (12,12,12) mesh. The value of the macroscopic dielectric constant coincides with the result of Ref. 3 obtained by dielectric matrices and so does the transverse frequency previously obtained by the direct method. We refer to Ref. 3 for a discussion of the slow convergence of ϵ_∞ with respect to the number of integration points in the BZ. We notice that the value obtained with 28 points is within ~0.5% of the converged value, while the value with 10 points is ~7% off. The slow convergence of the screening of a macroscopic field also affects the values of the effective charges. These are zero as a result of the acoustic sum rule (ASR)[19] which imposes a complete screening of the bare ionic polarization. Since the convergence of the macroscopic screening is very slow, it is not surprising that the convergence of Z_T^* is slow as well. We conclude that, in order to satisfy the ASR, a high number of special points is necessary.

In Table II, the same data are summarized as calculated for GaAs. As was the case for Si, the local-density approximation predicts a value of dielectric constant ~10% too high with respect to experiment. A thorough discussion of this partial failure of the local-density approximation is given in Ref. 3. The two entries $\omega_{\mathrm{LO}}^{\mathrm{Ga}}$ and $\omega_{\mathrm{LO}}^{\mathrm{As}}$ refer to calculations in which the Ga and As atoms respectively are displaced. Their values differ somewhat when a "low" number of special points is used for BZ integrations, as a consequence of the violation of the ASR. This happens because, as a result of the violation of the ASR, the frequency of the longitudinal acoustic mode is different from zero.[14] Since the mixed optic-acoustic mode corresponding to the displacement of a single ion (per unit cell) contains a different component of the acoustic mode according to which ion is moved, the resulting energy variations are different. The absolute

TABLE I. Calculated values of the macroscopic dielectric constant, ϵ_∞, transverse optic frequency, ω_{TO}, longitudinal optic frequency, ω_{LO} (THz), and effective charge, Z_T^*, in Si. The number of special points for the BZ integration in Eq. (2) is indicated. Experimental data as in Ref. 3.

	ϵ_∞	ω_{TO}	ω_{LO}	Z_T^*
10 points	13.5	15.4	15.4	7×10^{-2}
28 points	12.7	15.4	15.4	3×10^{-3}
Expt.	11.4	15.5	15.5	0

TABLE II. Calculated values of the high-frequency dielectric constant, ϵ_∞, static dielectric constant, ϵ_0, transverse optic frequency, ω_{TO}, longitudinal optic frequency, ω_{LO} (THz), and effective charges, Z_T^*, in GaAs. The labels Ga and As in ω_{LO} refer to the ion which is actually displaced in the calculation (see text). The number of special points for the BZ integration in Eq. (2) is indicated. Experimental data as in Ref. 20.

	ϵ_∞	ϵ_0	ω_{TO}	$\omega_{\text{LO}}^{\text{Ga}}$	$\omega_{\text{LO}}^{\text{As}}$	$Z_T^{*\text{Ga}}$	$Z_T^{*\text{As}}$
10 points	13.4	15.3	8.14	8.65	8.73	2.01	−2.15
28 points	12.3	14.0	8.14	8.70	8.70	2.02	−2.02
Expt.	10.9	12.9	8.06	8.75	8.75	2.16	−2.16

values of the effective charges differ accordingly. When the number of special points is high enough to guarantee convergence in the screening of the macroscopic field, the longitudinal frequency does not depend on which atom is moved, and the absolute values of the effective charges coincide, as they should. The present calculation is the first accurate enough to yield equal absolute values of the effective charges.

We believe that the method presented in this work will bring many new interesting phenomena within the reach of first-principles calculations. Among them, we mention here the calculation of phonon spectra at interfaces or in heterostructures; the electron-phonon coupling at an arbitrary point of the BZ; the study of order/disorder in alloys and heterostructure by perturbation theory with respect to the virtual crystal; and all those cases where perturbation theory is expected to yield some insight into complex phenomena.

This work has been partially supported by the Italian Ministry of Education through the Centro Interuniversitario di Struttura Materia, and by the Swiss National Science Foundation under Grant No. 2.683-085, and is part of the collaborative project between Scuola Internazionale Superiore di Studi Avanzati (SISSA) and CINECA. One of us (S.B.) wishes to thank A. Baldereschi and A. Quattropani for the hospitality at the Ecole Polytechnique Féderale de Lausanne, where part of this work has been performed.

---

[1] R. Resta, in *Festkörperprobleme: Advances in Solid-State Physics,* edited by P. Grosse (Vieweg, Braunschweig, 1985), Vol. 25, p. 183.

[2] O. H. Nielsen, and R. M. Martin, Phys. Rev. B **32**, 3780, 3792 (1985).

[3] S. Baroni and R. Resta, Phys. Rev. B **33**, 5969, 7017 (1986).

[4] R. Resta and K. Kunc, Phys. Rev. B **34**, 7146 (1986).

[5] See, for instance, K. Kunc, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter,* edited by J. T. Devreese and P. Van Camp (Plenum, New York, 1985), p. 227.

[6] See, for instance, E. R. Davidson, in *Methods in Computational Molecular Physics,* edited by G. H. F. Diercksen and S. Wilson (Reidel, Dordrecht, 1983), p. 95.

[7] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[8] MN^2 stands here for the number of FPO's necessary to perform M vector-times-matrix multiplications. By the use of the sparseness of the potential- and kinetic-energy matrices and fast-Fourier-transform techniques, this number can be reduced, for large N, to $\sim MN\log N$ (cf. Ref. 7).

[9] J. H. Wilkinson and C. Reinsch, *Handbook for Automatic Computation, Linear Algebra* (Springer-Verlag, Heidelberg, 1971), Vol. 2.

[10] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[11] G. D. Mahan, Phys. Rev. A **22**, 1780 (1980). A thorough discussion of Green's-function techniques for evaluating sums over intermediate states in atoms can be found in S. Baroni and A. Quattropani, Nuovo Cimento D **5**, 89 (1985).

[12] See, for instance, W. Kohn and P. Vashista, in *Theory of the Inhomogeneous Electron Gas,* edited by S. Lundqvist and N. H. March (Plenum, New York, 1983), p. 79.

[13] S. L. Alder, Phys. Rev. **126**, 413 (1962); N. Wiser, Phys. Rev. **129**, 62 (1963).

[14] R. Resta, Phys. Rev. B **27**, 3620 (1980).

[15] K. Kunc and R. Martin, Phys. Rev. Lett. **48**, 406 (1982); R. Resta and A. Baldereschi, Phys. Rev. B **24**, 4839 (1981); see also Refs. 4 and 5.

[16] P. D. DeCicco and F. A. Johnson, Proc. Roy. Soc. London, Ser. A **310**, 111 (1969).

[17] Further details on the study of *macroscopic* perturbations within the present scheme will be given in a forthcoming publication.

[18] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[19] R. M. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B **1**, 910 (1970).

[20] E. Burstein, A. Pinczuk, and R. I. Wellis, in *The Physics of Semimetals and Narrow-Gap Semiconductors* (Pergamon, Oxford, 1971), p. 251.