PHYSICAL REVIEW B VOLUME 9, NUMBER 4 15 FEBRUARY 1974

Dielectric-screening matrix and lattice dynamics of Si

C. M. Bertoni, V. Bortolani, and C. Calandra
Istituto di Fisica and Gruppo Nazionale Struttura della Materia, University of Modena, Modena, Italy

E. Tosatti*
Gruppo Nazionale Struttura della Materia del Consiglio Nazionale delle Richerche, Istituto di Fisica Università di Roma, Rome, Italy

(Received 20 February 1973)

The dynamical matrix of a semiconductor is set up by evaluating the full dielectric matrix $\epsilon(\vec{q}+\vec{G},\vec{q}+\vec{G}')$ to the lowest order in a Brillouin-Wigner perturbation scheme. The introduction in the dynamical matrix of the nondiagonal elements of the dielectric matrix, which account for the lattice effects on the electron screening, gives rise to noncentral forces between the atoms and restores the correct $q\rightarrow 0$ limit for the longitudinal acoustic frequencies through the fulfillment of the acoustic sum rule. A calculation of the phonon frequencies of Si is presented. The results, which are in reasonable agreement with experimental data, show that the off-diagonal elements are responsible for the stability of the crystal against shear and give real values for transverse acoustic frequencies. It is also shown that a third-order Rayleigh-Schrödinger perturbation theory, with some correction to include higher-order effects, can provide a good description of phonon spectrum of Si, but poorly reproduces the experimental elastic constants.

I. INTRODUCTION

The microscopic description of the dynamical matrix of crystals can be given in terms of the electron-ion potential and of the inverse dielectric matrix of the electron gas.[1,2] In nearly-free-electron (NFE) materials the screening properties of the electron gas are usually accounted for by the Lindhard dielectric function for metals[3] and by the Penn dielectric function for semiconductors.[4] In both cases the electron gas is considered as homogeneous gas, i.e., translational invariance is assumed. As is well known,[5] this is a good approximation for simple metals, but there are a number of cases where it is bound to fail.

The breakdown of the uniform translational invariance due to the lattice periodicity requires the replacement of the dielectric function with a dielectric matrix (DM) $\epsilon(\vec{q}+\vec{G},\ \vec{q}+\vec{G}')$, with $\vec{G}$ and $\vec{G}'$ reciprocal-lattice vectors.[6] The effects associated with the off-diagonal elements of the dielectric matrix (ODEDM) are particularly important in the calculation of the local field corrections to optical spectra,[6] of plasmon bands in metals,[7] and of valence-electron density[8] and phonon frequencies in semiconductors.[9,10] In this paper we are concerned with the last problem, dealing with the influence of the ODEDM on the lattice vibrations in semiconductors.

The usual pseudopotential formulation of lattice dynamics of diamond-structure semiconductors with the Penn dielectric function has failed for two reasons: First, the incomplete screening gives longitudinal-acoustic (LA) phonon frequencies which do not go to zero at long wavelength; second, the transverse-acoustic (TA) branches turn out to be imaginary for any reasonable choice of the potential, giving rise to an instability against shear. One can circumvent the first difficulty by adding an extra charge on the ion sites, in order to restore the charge neutrality of the crystal, so that one obtains the correct $\vec{q}\rightarrow 0$ limit for LA. The second point is more serious and it appears that, without the introduction of non-central forces between pairs of atoms, the TA branches remain imaginary. This fact is not surprising, if one considers the difficulties of reproducing the experimental phonon frequencies by a Born-von Karman fitting[11] or a shell model[12] with a reasonable number of parameters.

Both of these problems are solved, if the dynamical matrix is set up, by including the effects of the ODEDM in the dielectric screening. The study of these effects is the primary purpose of this paper.

Indeed the nondiagonal screening in the dynamical matrix gives rise, in the direct space, to noncentral forces, which are responsible for the real values of the TA frequencies, and restore the crystal stability. Moreover we find that the introduction of the ODEDM can restore the correct $\vec{q}\rightarrow 0$ limit for the LA frequencies, through the fulfillment of the acoustical sum rule derived by Sham.[13]

In order to prove that the use of DM explains the lattice dynamics of diamond-structure semiconductors, we have to evaluate the ODEDM. This is a formidable task since it requires a detailed knowledge of the electron band structure. To avoid this difficulty various phenomenological approaches have been proposed. By adopting the bond-charge model,[14] Martin[10] simulated the effects of the exact ODEDM in the dynamical matrix by the introduction of point charges, located midway between neighbors, which interact via Coulomb potential among themselves and with the ions, their charge being appropriate to neutralize the system. The results obtained in this paper indicate the deep connection between the ODEDM and the pileup of the charge along the bonds in covalent semiconductors.

In the approach proposed by Sinha, Gupta, and Price,[15] a separable expression of DM is taken, leading to a parametrization, which reproduces either the shell model or the bond-charge picture according to the values of the parameters. More recently, Soma and Morita[16] carried out the calculation of the total energy and the phonon frequencies of Si and Ge using a perturbative approach. In their calculation, terms up to the fourth order in the pseudopotential are retained, in order to reproduce the contribution to the band gap coming from second-order perturbation theory, i.e., the effective enhancement of the $V(220)$ component, caused by the $V(111)$ one.[17] Even if in this approach the bond-charge model is introduced explicitly in the effective potential, it seems that such an analysis is closer to the microscopic description of lattice vibration in terms of dielectric matrix and electron-ion pseudopotential, than the previous one.

In this paper, by assuming that Si can be treated in the NFE scheme, we obtain an explicit form of the ODEDM as a perturbation expansion in the electron-ion potential. In order to handle the degeneracy at the zone boundary, which, as is well known, makes finite the dielectric function of semiconductors for $\vec{q} \to 0$, we use a Brillouin-Wigner expansion. To render the problem feasible, the energy denominators have been approximated with a simple energy-band model, as discussed in Sec. II. A further simplification is introduced with a $\vec{q}\cdot\vec{p}$ expansion around the point $\vec{q}=0$. In this way simple analytical expressions for the ODEDM at large wavelength, as well as for the acoustic sum rule, are obtained. For comparison, the form taken by the dielectric matrix in the case of metals, i.e., when a Rayleigh-Schrödinger (RS) perturbation expansion is used instead of the Brillouin-Wigner theory, is also indicated.

We give in Sec. III the dynamical matrix, set up in this approach, both in the case of semiconductor and of metal. By a careful analysis of the nature of the ion-ion forces, which originate in a complete treatment of electron screening, we show that the ODEDM introduce non-central-type forces in the dynamical matrix. The connection between the bond-charge model and the microscopic theory is also indicated.

The phonon frequencies calculated with a suitable pseudopotential are given in Sec. IV. The calculations are in good agreement with experimental values, especially at small q, the discrepancies at the zone boundary owing to the use of the $\vec{q}\cdot\vec{p}$ approximation. The key result is that the introduction of the whole screening matrix makes real the TA branches and also, through the fulfillment of the acoustic sum rule, provides the correct $\vec{q} \to 0$ limit for the LA branches.

In order to complete our analysis, and investigate more precisely the effect of the band gap on lattice dynamics, we have also calculated the phonon frequencies of Si with the DM appropriate to a metal. This amounts to evaluating the total electron energy up to the third order in a RS pseudopotential expansion. Although this approach does not reproduce correctly the screening in a semiconductor at long wavelength, we expect that it gives a good approximation to the total electron energy and therefore to the dynamical matrix. In this case the acoustical sum rule is automatically satisfied and non-central-type forces are introduced via the ODEDM.

A comparison between the two calculations shows that for the LA frequencies in the $\vec{q} \to 0$ region and for the elastic constants c_{11} and c_{12}, the results depend critically on the inclusion of the band gap in the electron energies. In fact, we find better agreement with experiment by using the semiconductor screening. For the TA branches and c_{44}, which depend only upon the U processes ($\vec{G}$ and $\vec{G}' \neq 0$), the use of a metallic or a semiconductor dielectric matrix makes very little difference in the results.

We conclude that the correct understanding of the lattice vibrations in semiconductors can be obtained only by including lattice effects in the dielectric matrix.

II. DIELECTRIC MATRIX

The random-phase-approximation (RPA) expression for the static DM can be written in terms of the electron Bloch energies $E_{\vec{k}l}$ and of the periodic part of the wave function[6] $b_{\vec{k}l}(\vec{r})$ as

$$\epsilon(\vec{q}+\vec{G},\ \vec{q}+\vec{G}') = \delta_{\vec{G}\vec{G}'} - v_C(\vec{q}+\vec{G})\pi(\vec{q}+\vec{G},\ \vec{q}+\vec{G}')\ ,$$

$$\pi(\vec{q}+\vec{G},\ \vec{q}+\vec{G}') = \sum_{\vec{k}\,l\,l'} \frac{f_0(E_{\vec{k}+\vec{q}\,l'}) - f_0(E_{\vec{k}l})}{E_{\vec{k}+\vec{q}\,l'} - E_{\vec{k}l}} \times (b_{\vec{k}+\vec{q}\,l'}|\, e^{-i\vec{G}\cdot\vec{r}}\,|\, b_{\vec{k}l}) \times (b_{\vec{k}l}|\, e^{i\vec{G}'\cdot\vec{r}}\,|\, b_{\vec{k}+\vec{q}\,l'})\ , \tag{2.1}$$

where l and l' are band indices, $f_0(E_{\vec{k}l})$ is the Fermi occupation number, and $v_C(q) = 4\pi e^2/q^2\Omega$ is the Fourier transform of the Coulomb potential, Ω being the normalization volume, which will be taken equal to the unit-cell volume.

In a metal both the interband ($l \neq l'$) and the intraband ($l = l'$) contributions are present; in a semiconductor only the former survives and the denominator $E_{\vec{k}+\vec{q}\,l'} - E_{\vec{k}l}$ does not vanish. The difference in the screening properties is evident if the small-$\vec{q}$ behavior of the DM is considered. By expanding in a power series in $\vec{q}$ the matrix elements in Eq. (2.1), one obtains

$$\pi(\vec{q},\vec{q})=\sum_{\vec{k}ll'}\frac{f_0(E_{\vec{k}+\vec{q}l'})-f_0(E_{\vec{k}l})}{E_{\vec{k}+\vec{q}l'}-E_{\vec{k}l}}\left[\delta_{ll'}+(1-\delta_{ll'})\left|(\nabla_{\vec{k}}b_{\vec{k}l'}|b_{\vec{k}l})\right|^2q^2+O(q^4)\right],$$

$$\pi(\vec{q},\vec{q}+\vec{G})=\sum_{\vec{k}ll'}\frac{f_0(E_{\vec{k}+\vec{q}l'})-f_0(E_{\vec{k}l})}{E_{\vec{k}+\vec{q}l'}-E_{\vec{k}l}}\left[\delta_{ll'}(b_{\vec{k}l}|e^{i\vec{G}\cdot\vec{r}}|b_{\vec{k}l})+\vec{q}\cdot(\nabla_{\vec{k}}b_{\vec{k}l'}|b_{\vec{k}l})(b_{\vec{k}l}|e^{i\vec{G}\cdot\vec{r}}|b_{\vec{k}l'})(1-\delta_{ll'})+\cdots\right],$$

$$\pi(\vec{q}+\vec{G},\vec{q})=\pi^*(\vec{q},\vec{q}+\vec{G}), \tag{2.2}$$

$$\pi(\vec{q}+\vec{G},\vec{q}+\vec{G}')=\sum_{\vec{k}ll'}\frac{f_0(E_{\vec{k}+\vec{q}l'})-f_0(E_{\vec{k}l})}{E_{\vec{k}+\vec{q}l'}-E_{\vec{k}l}}\left[(b_{\vec{k}l'}|e^{-i\vec{G}\cdot\vec{r}}|b_{\vec{k}l})(b_{\vec{k}l}|e^{i\vec{G}'\cdot\vec{r}}|b_{\vec{k}l'})+O(q^2)\right].$$

For a semiconductor only the interband terms survive and $\epsilon(\vec{q},\vec{q})$ goes to a finite limit, $\epsilon(\vec{q},\vec{q}+\vec{G})$ and $\epsilon(\vec{q}+\vec{G},\vec{q})$ go as $1/\vec{q}$ and $\vec{q}$, respectively, while $\epsilon(\vec{q}+\vec{G},\ \vec{q}+\vec{G}')$ gives a finite limit. For the case of metals, $\epsilon(\vec{q},\vec{q})$ and $\epsilon(\vec{q},\vec{q}+\vec{G})$ go to infinity as $1/q^2$, whereas all the other matrix elements go to a finite limit.

These results are quite general and apply to any kind of material, but in the following we restrict our attention to the case of those semiconductors and metals for which a pseudopotential description of the electrons is adequate. Then we can write both the energy and wave function as a Brillouin-Wigner expansion in power of the electron-ion self-consistent pseudopotential:

$$b_{\vec{k}\vec{Q}}(\vec{r})=e^{i\vec{Q}\cdot\vec{r}}+\sum_{\vec{Q}'\neq\vec{Q}}\frac{(e^{-i\vec{Q}'\cdot\vec{r}}|u|b_{\vec{k}\vec{Q}})}{E_{\vec{k}\vec{Q}}-E^0_{\vec{k}\vec{Q}'}}e^{i\vec{Q}'\cdot\vec{r}}, \tag{2.3}$$

$$E_{\vec{k}\vec{Q}}=E^0_{\vec{k}\vec{Q}}+\sum_{\vec{Q}'\neq\vec{Q}}\frac{|(e^{-i\vec{Q}'\cdot\vec{r}}|u|b_{\vec{k}\vec{Q}})|^2}{E_{\vec{k}\vec{Q}}-E^0_{\vec{k}\vec{Q}'}}, \tag{2.4}$$

where u is the screened total potential of the ions seen by an electron, $\vec{Q}$ and $\vec{Q}'$ are reciprocal-lattice vectors, and $E^0_{\vec{k}\vec{Q}}$ are free-electron energies $|\vec{k}+\vec{Q}|^2$.

To lowest order in u, the polarizability is

$$\pi(\vec{q}+\vec{G},\vec{q}+\vec{G})=\sum_{\vec{k}\vec{Q}}\frac{f_0(E_{\vec{k}+\vec{Q}+\vec{q}+\vec{G}})-f_0(E_{\vec{k}+\vec{Q}})}{E_{\vec{k}+\vec{Q}+\vec{q}+\vec{G}}-E_{\vec{k}+\vec{Q}}}+\Theta(u^2)$$

$$\pi(\vec{q}+\vec{G},\vec{q}+\vec{G}')=2u(\vec{G}-\vec{G}')\sum_{\vec{k}\vec{Q}}\left(\frac{f_0(E_{\vec{k}+\vec{Q}+\vec{q}+\vec{G}})-f_0(E_{\vec{k}+\vec{Q}})}{(E_{\vec{k}+\vec{Q}+\vec{q}+\vec{G}}-E_{\vec{k}+\vec{Q}})(E_{\vec{k}+\vec{Q}}-E^0_{\vec{k}+\vec{Q}+\vec{G}-\vec{G}'})}+\text{term with }\vec{G}\rightleftharpoons\vec{G}'\right), \tag{2.5}$$

where $u(\vec{G})$ is the Fourier component of the potential consistently screened. It is evident that the ODEDM are first-order terms in the potential, while corrections in the diagonal terms appear only in second order.

In the case of simple metals the approximation, obtained by replacing the true Bloch energies by their free-electron values $E^0_{\vec{k}}$, is adequate; this means that one can use Rayleigh-Schrödinger perturbation theory instead of the Brillouin-Wigner one. In this way one obtains analytical expressions for the off-diagonal elements in first order in $u(\vec{G})$.[18]

For semiconductors, however, it is essential to retain the true Bloch energies in order to reproduce the correct $\vec{q}\to 0$ limit. In this case, by inserting Eqs. (2.3) and (2.4) into Eq. (2.2) and retaining the leading terms in the $\vec{q}$ expansion of the polarization part, one has for the ODEDM

$$\pi(\vec{q},\vec{q}+\vec{G})=\pi^*(\vec{q}+\vec{G},\vec{q})\equiv\vec{q}\cdot\vec{P}(\vec{G}),$$

$$\vec{P}(\vec{G})=u^*(\vec{G})\sum_{\vec{k}}\frac{\nabla_{\vec{k}}(E_{\vec{k}+\vec{G}}-E^0_{\vec{k}})[f_0(E_{\vec{k}+\vec{q}})-f_0(E_{\vec{k}})]}{(E_{\vec{k}+\vec{G}}-E^0_{\vec{k}})(E_{\vec{k}}-E^0_{\vec{k}+\vec{G}})^2},$$

$$\pi(\vec{q}+\vec{G},\vec{q}+\vec{G}')=\pi(\vec{G},\vec{G}')+O(q^2), \tag{2.6}$$

$$\pi(\vec{G},\vec{G}')=2u(\vec{G}-\vec{G}')\sum_{\vec{k}}\left(\frac{f_0(E_{\vec{k}+\vec{G}})-f_0(E_{\vec{k}})}{(E_{\vec{k}+\vec{G}}-E_{\vec{k}})(E_{\vec{k}}-E^0_{\vec{k}+\vec{G}-\vec{G}'})}+\text{the same with }\vec{G}\rightleftharpoons\vec{G}'\right),$$

where the sums over $\vec{k}$ are now extended to the whole space so that $f_0(E_{\vec{k}})$ is equal to 1, only inside the Jones zone.

To evaluate Eqs. (2.6) we have to specify the energy $E_{\vec{k}}$. For the diagonal terms the main contribution arises from the upper valence band and from the lowest conduction band, so that a two-band model is appropriate. As shown in the work of Walter and Cohen,[19] $\epsilon(\vec{q},\vec{q})$ depends weakly upon the direction of $\vec{q}$, so that we approximate the diagonal part of the DM by the two-band isotropic model for semiconductors, owing to Penn.[4]

For the off-diagonal terms, we approximate the Jones zone by a sphere of the same volume, as in the Penn model, and take the energy as

$$E_{\vec{k}} = k^2 + \Theta(k - k_0)E_g \,, \tag{2.7}$$

where Θ is the step function, k_0 is the sphere radius, and E_g is the energy gap. This is the simplest approximation which accounts for the gap between occupied and unoccupied states and leads to the correct $\vec{q} \to 0$ limit in the DM.

Up to this point we have not given any expression for the self-consistent screened potential u. If the validity of a perturbative approach is assumed, the self-consistent potential can be written as a power series in the bare electron-ion potential $V(\vec{G})$ multiplied by the structure factor $S(\vec{G})$. In our expression we retain only lowest-order terms for the expansion and write the self-consistent potential as

$$u(\vec{G}) = V(\vec{G})S(\vec{G})/\epsilon(\vec{G}) \,, \tag{2.8}$$

where $\epsilon(\vec{G})$ is the diagonal dielectric function of the Penn model.

Since the self-consistent approach does not take into account effects of exchange and correlation, we have to modify the theory to allow for an evaluation of these effects. This is an important point since previous work on both metals and semiconductors shows that phonon frequencies depend considerably upon these corrections.[20] The necessary formalism to evaluate the contribution of exchange and correlation in the total energy has been outlined, in a quite general way, in Ref. 21. In this paper we shall adopt the equivalent approach based on the work of Singwi *et al.*[22] According to this approximation, the dielectric matrix which screens the electron-ion interaction may be taken as in Eq. (2.1), except that the Coulomb potential v_C is modified in the following way:

$$v_C(\vec{q}+\vec{G}) \to v_C(\vec{q}+\vec{G})[1 - f(\vec{q}+\vec{G})] \,, \tag{2.9}$$

where

$$f(\vec{q}) = A(1 - e^{-B(q/k_0)^2}) \,, \tag{2.10}$$

A and B being constants depending on the electron density.

III. DYNAMICAL MATRIX

Within the Born-Oppenheimer and the harmonic approximations, the effective potential of the ions consists of two parts: (i) a direct Coulomb interaction between the ions, and (ii) an indirect interaction via the electrons. In this scheme, the dynamical matrix which describes the ionic motion becomes

$$D_{\alpha\beta}(\vec{q};\, ss') = D^{I}_{\alpha\beta}(\vec{q};\, ss') + D^{E}_{\alpha\beta}(\vec{q};\, ss') \,, \tag{3.1}$$

where s and s' are indices which label the ions inside the unit cell, and the superscripts I and E indicate the direct and indirect interactions, respectively.

The Coulomb term $D^{I}_{\alpha\beta}(\vec{q}, ss')$ can be easily evaluated by Ewald's method and is explicitly given for the case of two atoms per unit cell in Ref. 23.

The $D^{E}_{\alpha\beta}(\vec{q}, ss')$ term contains the screening effect of the electron gas on lattice vibrations. It can be expressed in terms of the inverse dielectric matrix $\epsilon^{-1}(\vec{q}+\vec{G}, \vec{q}+\vec{G}')$ and the bare electron-ion form factor $V_s(\vec{q})$ by

$$(M_s M_{s'})^{1/2} D^{E}_{\alpha\beta}(\vec{q};\, ss') = \sum_{\vec{G}\vec{G}'} \Big[(\vec{q}+\vec{G})_\alpha (\vec{q}+\vec{G}')_\beta \Phi^{ss'}(\vec{q}+\vec{G}, \vec{q}+\vec{G}') e^{i\vec{G}\cdot\vec{R}_s} e^{-i\vec{G}'\cdot\vec{R}_{s'}} - \delta_{ss'} \sum_{s''} \vec{G}_\alpha \vec{G}'_\beta \Phi^{ss''}(\vec{G}, \vec{G}') e^{i\vec{G}\cdot\vec{R}_s} e^{-i\vec{G}'\cdot\vec{R}_{s''}} \Big] \,, \tag{3.2}$$

with

$$\Phi^{ss'}(\vec{q}+\vec{G}, \vec{q}+\vec{G}') = \frac{\Omega_0 |\vec{q}+\vec{G}|^2}{4\pi e^2} \, \frac{V_s(\vec{q}+\vec{G}) V_{s'}(\vec{q}+\vec{G}')}{1 - f(\vec{q}+\vec{G})} \, [\epsilon^{-1}(\vec{q}+\vec{G}, \vec{q}+\vec{G}') - \delta_{\vec{G}\vec{G}'}] \,, \tag{3.3}$$

where Ω_0 is the cell volume, M_s and $\vec{R}_s$ are, respectively, the mass and the position of the sth ion in the cell.

To evaluate Eq. (3.3) we need an explicit form of the inverse dielectric matrix. To invert the dielectric matrix we follow the procedure given in the work of Pick, Cohen, and Martin.[2] These authors subdivide the DM into two square matrices R and S, whose elements are, respectively, $\epsilon(\vec{q}, \vec{q})$ and $\epsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}')$, with $\vec{G}$ and $\vec{G}' \neq 0$, and row and column matrices, respectively given by $\epsilon(\vec{q}, \vec{q}+\vec{G})$ and $\epsilon(\vec{q}+\vec{G}, \vec{q})$. In this way, they separate blocks of the DM having different analytical behavior for large wavelength. By performing block inversions they are able to obtain for the elements of the inverse DM, the following expressions:

$$
\begin{aligned}
\epsilon^{-1}(\vec{q},\vec{q}) &= \Big[\epsilon(\vec{q},\vec{q}) - \sum_{\vec{G}\vec{G}'}{}' \epsilon(\vec{q},\vec{q}+\vec{G}) \\
&\qquad \times S^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}')\epsilon(\vec{q}+\vec{G}',\vec{q})\Big]^{-1}, \\
\epsilon^{-1}(\vec{q},\vec{q}+\vec{G}) &= -\epsilon^{-1}(\vec{q},\vec{q})\sum_{\vec{G}'}{}' \epsilon(\vec{q},\vec{q}+\vec{G}') \\
&\qquad \times S^{-1}(\vec{q}+\vec{G}',\vec{q}+\vec{G}), \\
\epsilon^{-1}(\vec{q}+\vec{G},\vec{q}) &= -\epsilon^{-1}(\vec{q},\vec{q})\sum_{\vec{G}'}{}' \epsilon(\vec{q}+\vec{G}',\vec{q}) \\
&\qquad \times S^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}'), \\
\epsilon^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}') &= S^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}') \\
&\qquad + \epsilon^{-1}(\vec{q}+\vec{G},\vec{q})[\epsilon^{-1}(\vec{q},\vec{q})]^{-1}\epsilon^{-1}(\vec{q},\vec{q}+\vec{G}').
\end{aligned}
\tag{3.4}
$$

Since the S matrix is obtained by striking out the elements of DM which are not analytical for $\vec{q}\to 0$, its off-diagonal terms, which arise in first order in the perturbation expansion, are small in comparison with the diagonal part. By writing any element of the inverse matrix S^{-1} as a power series in the pseudopotential and retaining the leading terms, we have

$$
S^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}') = \frac{\delta_{\vec{G}\vec{G}'}}{\epsilon(\vec{q}+\vec{G})} - (1-\delta_{\vec{G}\vec{G}'})\frac{\epsilon(\vec{q}+\vec{G},\vec{q}+\vec{G}')}{\epsilon(\vec{q}+\vec{G})\epsilon(\vec{q}+\vec{G}')}\,. \tag{3.5}
$$

By inserting Eq. (3.5) into Eqs. (3.4) and neglecting higher-order terms, we get, in the small-$\vec{q}$ limit,

$$
\begin{aligned}
\epsilon^{-1}(\vec{q},\vec{q}) &= 1/\epsilon(q), \\
\epsilon^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}') &= \frac{\delta_{\vec{G}\vec{G}'}}{\epsilon(\vec{q}+\vec{G})} + (1-\delta_{\vec{G}\vec{G}'})\frac{v_C(\vec{q}+\vec{G})\pi(\vec{G},\vec{G}')}{\epsilon(\vec{q}+\vec{G})\epsilon(\vec{q}+\vec{G}')} \\
&\qquad + v_C(q)\frac{\vec{q}\cdot\vec{P}^*(\vec{G})}{\epsilon(\vec{q}+\vec{G})}\,\frac{\vec{q}\cdot\vec{P}(\vec{G}')}{\epsilon(\vec{q}+\vec{G}')} \\
&\qquad \times v_C(\vec{q}+\vec{G})\,\frac{1}{\epsilon(q)}, \\
\epsilon^{-1}(\vec{q}+\vec{G},\vec{q}) &= \frac{\vec{q}\cdot\vec{P}^*(\vec{G})}{\epsilon(q)\epsilon(\vec{q}+\vec{G})}\,v_C(q+G), \\
\epsilon^{-1}(\vec{q},\vec{q}+\vec{G}) &= \frac{\vec{q}\cdot\vec{P}(\vec{G})}{\epsilon(q)\epsilon(\vec{q}+\vec{G})}\,v_C(q)\,.
\end{aligned}
\tag{3.6}
$$

Inserting Eqs. (3.6) into Eqs. (3.2) and (3.3), the long wavelength expression for the dynamical matrix is obtained:

$$
\begin{aligned}
(M_sM_{s'})^{1/2}D^E_{\alpha\beta}(\vec{q};ss') &= -V_s(q)V_{s'}(q)\frac{q_\alpha q_\beta q^2\Omega_0}{4\pi e^2[1-f(q)]} + \frac{v_C(q)}{\epsilon(q)}\sum_{\lambda\mu}\left(\frac{V_s(q)}{v_C(q)}\delta_{\alpha\mu} + \sum_{\vec{G}\neq 0}\frac{V_s(\vec{q}+\vec{G})}{\epsilon(\vec{q}+\vec{G})}(\vec{q}+\vec{G})_\alpha P^*_\mu(\vec{G})e^{i\vec{G}\cdot\vec{R}_s}\right)q_\mu q_\lambda \\
&\quad \times\left(\frac{V_{s'}(q)}{v_C(q)}\delta_{\beta\lambda} + \sum_{\vec{G}\neq 0}\frac{V_{s'}(\vec{q}+\vec{G})}{\epsilon(\vec{q}+\vec{G})}(\vec{q}+\vec{G})_\beta P_\lambda(\vec{G})e^{-i\vec{G}\cdot\vec{R}_{s'}}\right) \\
&\quad + \sum_{\vec{G}\vec{G}'}{}'\Big[\delta_{\vec{G}\vec{G}'}(\vec{q}+\vec{G})_\alpha(\vec{q}+\vec{G})_\beta V_s(\vec{q}+\vec{G})V_{s'}(\vec{q}+\vec{G})\left(\frac{1}{\epsilon(\vec{q}+\vec{G})}-1\right)\frac{1}{v_C(\vec{q}+\vec{G})} \\
&\quad + (1-\delta_{\vec{G}\vec{G}'})(\vec{q}+\vec{G})_\alpha(\vec{q}+\vec{G}')_\beta\,\frac{V_s(\vec{q}+\vec{G})V_{s'}(\vec{q}+\vec{G}')}{\epsilon(\vec{q}+\vec{G})\epsilon(\vec{q}+\vec{G}')}\pi(\vec{G},\vec{G}')\Big]e^{i\vec{G}\cdot\vec{R}_s}e^{-i\vec{G}'\cdot\vec{R}_{s'}} \\
&\quad - \sum_{\vec{G}\vec{G}'}{}'\delta_{ss'}\sum_{s''}(\text{the same with } q=0 \text{ and } s'\to s'')e^{i\vec{G}\cdot\vec{R}_s}e^{-i\vec{G}'\cdot\vec{R}_{s''}}\,.
\end{aligned}
\tag{3.7}
$$

In order to obtain, in our model, the fact that the longitudinal-acoustic frequencies vanish for $\vec{q}\to 0$, we have to impose the extra condition

$$
Z^s_{\alpha\mu} = \sum_{\vec{G}\neq 0}\frac{V_s(\vec{q}+\vec{G})}{\epsilon(\vec{q}+\vec{G})}(\vec{q}+\vec{G})_\alpha P_\mu(\vec{G})e^{-i\vec{G}\cdot\vec{R}_s} = Z^s\delta_{\alpha\mu}\,. \tag{3.8}
$$

This condition is the acoustic sum rule, derived in a general way by Sham.[13] It shows the necessity of including the ODEDM in the calculation of phonon frequencies of semiconductors. In fact, a pure diagonal screening does not give rise to LA modes which vanish at large wavelength, because of the finite limit of $\epsilon(\vec{q},\vec{q})$. We note that the Eq. (3.8) can be viewed as a relation, which must be satisfied by the elements $\pi(\vec{q},\vec{q}+\vec{G})$ of the polarization part, given in Eq. (2.6). In the actual calculations, formula (3.8) has been used as a tool to ensure general consistency for our DM and to overcome the approximations involved in the model band structure (2.9) and in the choice of the pseudopotential. For this purpose, we multiply $\vec{P}(\vec{G})$ by a constant factor α and require that it satisfy the acoustic sum rule. With this choice of the elements $\pi(\vec{q},\vec{q}+\vec{G})$, we are led to multiply by the same factor the off-diagonal elements $\pi(\vec{q}+\vec{G},\vec{q}+\vec{G}')$, so that the elements of π, evaluated to the same order of perturbation theory and using the same approximations, are consistent among them, as Eq. (2.5) requires. This amounts to multiplying both the effective charge tensor $Z^{s}_{\alpha\mu}$ and the double sums in Eq. (3.7) by the same factor. As we discussed previously,[9] we find that the choice of a multiplicative constant gives good results and allows us to overcome, to a large extent, the approximations used in the evaluation of our model DM.

In the case of a metal, where we can use RS perturbation expansion in the evaluating of DM, the off-diagonal matrix elements are smaller than the diagonal function, so that the block separation is not necessary. By retaining the leading term in u, the elements of the inverse matrix are given by

$$\epsilon^{-1}(\vec{q}+\vec{G},\vec{q}+\vec{G}')=\frac{1}{\epsilon(\vec{q}+\vec{G})}\delta_{\vec{G}\vec{G}'}-(1-\delta_{\vec{G}\vec{G}'})\frac{\epsilon(\vec{q}+\vec{G},\vec{q}+\vec{G}')}{\epsilon(\vec{q}+\vec{G})\epsilon(\vec{q}+\vec{G}')}\,. \tag{3.9}$$

By inserting Eq. (3.9) in the dynamical matrix and using expression (2.5) for the polarizability, with the true Bloch energies replaced by the free-electron values, we have by keeping only the lowest-order terms in u:

$$\Phi^{ss'}(\vec{q}+\vec{G},\vec{q}+\vec{G}')=\frac{\Omega_0(\vec{q}+\vec{G})^2}{4\pi e^2}\;\frac{V_s(\vec{q}+\vec{G})V_{s'}(\vec{q}+\vec{G}')}{1-f(\vec{q}+\vec{G})}\Bigg[\delta_{\vec{G}\vec{G}'}\left(\frac{1}{\epsilon(\vec{q}+\vec{G})}-1\right)+(1-\delta_{\vec{G}\vec{G}'})\frac{V(\vec{G}-\vec{G}')S(\vec{G}-\vec{G}')}{\epsilon(\vec{q}+\vec{G})\epsilon(\vec{q}+\vec{G}')\epsilon(\vec{G}-\vec{G}')}$$
$$\times 2\sum_{\vec{k}} f_0(E^0_{\vec{k}})\Bigg(\frac{1}{(E^0_{\vec{k}}-E^0_{\vec{k}+\vec{q}+\vec{G}})(E^0_{\vec{k}}-E^0_{\vec{k}+\vec{G}-\vec{G}'})}+\frac{1}{(E^0_{\vec{k}}-E^0_{\vec{k}+\vec{G}'-\vec{G}})(E^0_{\vec{k}}-E^0_{\vec{k}+\vec{G}'+\vec{q}})}+\frac{1}{(E^0_{\vec{k}}-E^0_{\vec{k}+\vec{q}'+\vec{G}})(E^0_{\vec{k}}-E^0_{\vec{k}+\vec{G}'+\vec{q}})}\Bigg)\Bigg]\,. \tag{3.10}$$

The same results may be derived by evaluating the dynamical matrix from the expression of the total energy given by a third-order RS perturbation expansion.[24]

We note that in a metal, by virtue of the behavior of $\epsilon(\vec{q},\vec{q})$ at large wavelength, the acoustic sum rule is automatically satisfied and a pure diagonal screening ensures, by itself, the correct limit for the LA frequencies.

It is convenient at this point to discuss the nature of the ion-ion forces, which originate from the indirect interaction, when electron screening is treated by a DM. By Fourier-transforming Eq. (3.7) or (3.10), it is easily seen that the diagonal part of the DM gives rise in real space to a pairwise central interaction between the ions. On the other hand, the terms which involve three times the electron-ion potential, introduced by the inclusion of the ODEDM, originate in real space two-bodies non-central forces, which arise from the sum of three-body interactions via the electron gas.[25] The inclusion of such non-central forces should not change appreciably the phonon spectra of simple metals, but is very important in a semiconductor with diamond structure. It is, in fact, quite well established by the work of Dolling[12] and Keating,[26] as well as from the calculation performed by Martin,[10] that purely central forces do not give crystal stability, leaving the TA branches imaginary.

Some more insight on the nature of the indirect interaction can be obtained by considering the relation between the DM and the electron charge density. Within the linear-response theory, it can be shown[8] that the bond charge, which is the preeminent feature of covalent crystals, arises from the off-diagonal screening, the diagonal approximation giving simply the superposition of spherically symmetric charge distributions centered on the ions. Since semiconductors are essentially covalently bonded materials, no NFE approach, which neglects off-diagonal screening, adequately describes their phonon frequencies. We can illustrate this point in more detail by noting that a purely diagonal screening with the dielectric function appropriate for a semiconductor does not screen completely the ionic potential, so that a net charge $Z/\epsilon(0)$ remains to be neutralized. One could restore the charge neutrality either by adding the extra charge on the ions or assuming a point charge $Z/2\epsilon(0)$ located midway between nearest neighbors. In both cases the dynamical matrix vanishes for $\vec{q}\to 0$, giving the correct limit to LA branches. The first choice gives rise to purely central forces and therefore leaves the TA branches imaginary. In the second case, one recovers the Phillips-bond charge model[14] for covalent solids, which has been used by Martin[10] in his work on lattice vibrations in Si. In this model the inclusion of bond interaction as well as bond-ion interaction in setting up the dynamical matrix, gives rise to non-central bond-bending forces. Such a model dynamical matrix furnishes a fairly accurate description of lattice vibrations in Si, indicating that the inclusion of bonding charge

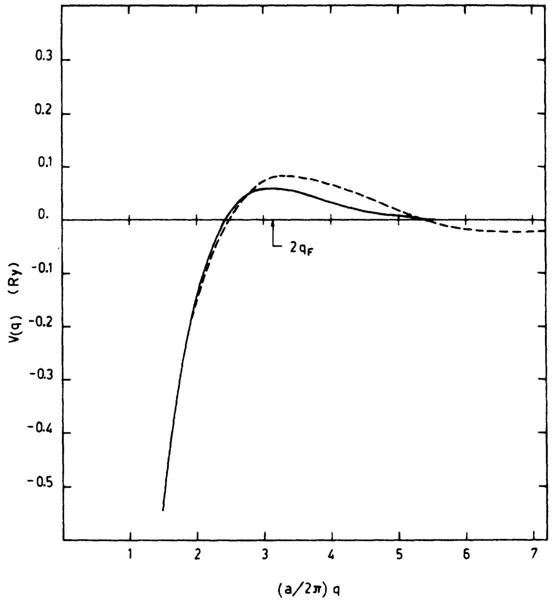

FIG. 1. Plot of the modified Heine-Abarenkov bare form factor (full curve). The dashed curve gives the original HA potential. q_F is the Fermi momentum.

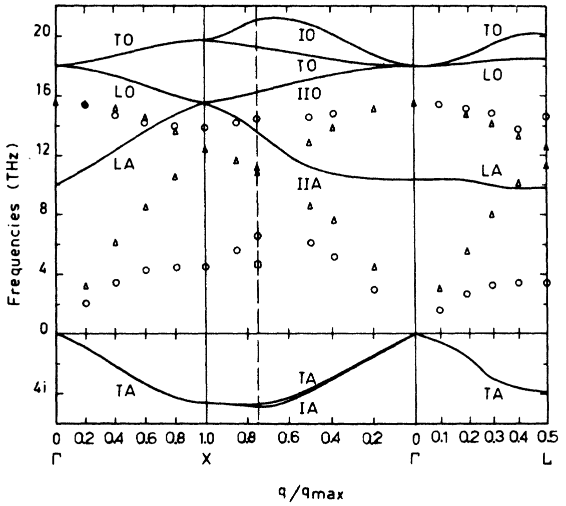

FIG. 2. Phonon dispersion curves of Si as calculated using Penn dielectric function. Note that TA branches are imaginary. The LA $\vec{q} \rightarrow 0$ limit is $\Omega_p/\epsilon(0)$, where Ω_p is the plasma frequency. Experimental points from Ref. 12.

can simulate the effect of the complete treatment of dielectric screening.

IV. RESULTS AND DISCUSSION

We now present the phonon frequencies of Si, calculated according to the theory of the previous sections. To show the effects of the ODEDM, we present various results which allow for a direct comparison between different types of screening. We carried out all the calculations using a local form of the Heine-Abarenkov (HA) model potential. Since for large values of the momentum transfer q any form factor is rather uncertain, we eliminate the spurious oscillations of the HA potential by making it go smoothly to zero at the second node. The modified form factor together with the original HA are given in Fig. 1. It was found that the final results were rather insensitive to the detailed form of the tail. In this way we also achieved a rapid convergence in the sums over reciprocal lattice vectors, so that we extend our sums up to 338 vectors only.

To estimate the importance of the ODEDM, we present in Fig. 2 the phonon frequencies calculated with the Penn dielectric function without the introduction of the ODEDM. We used the full Penn formula with exchange and correlation corrections. The main feature of these results is that the TA branches are imaginary, leaving the crystal unstable against shear. Moreover, because of the incomplete screening of the ions, the LA frequencies go to a finite limit as $\vec{q} \rightarrow 0$. We note also that the optical modes are higher than the experimental data, the disagreement at small $\vec{q}$ being about 20%.

The phonon frequencies calculated with the DM appropriate to a semiconductor, as given in formulas (2.6) and (2.7), are displayed in Fig. 3. The parameter α, which lets the acoustic sum rule be satisfied, has been evaluated via Eq. (3.8), and turns out to be $\alpha = 3$. No appreciable change of the value of α was found by using different reasonable form factors.

The most important conclusion which can be drawn from the analysis of the calculated phonon frequencies is that the replacement of the Penn dielectric function with the whole DM gives rise to real TA branches, which compare satisfactorily

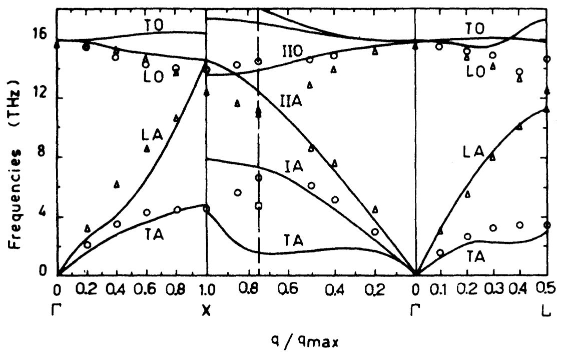

FIG. 3. Phonon dispersion curves given by the dielectric matrix of a semiconductor. The misfit at the zone boundary is due to the small-q expansion.

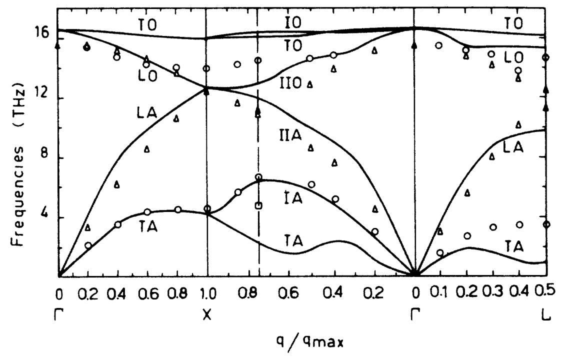

FIG. 4. Phonon dispersion curves obtained by using the dielectric matrix appropriate to a metal, i.e., RS third-order perturbation theory. Exchange and correlation are included.

with experimental data. We note that the agreement with experiment is not substantially washed out even for large q, where the small-q expansion is very rough. This indicates that in spite of the crudeness of the approximations, the essential feature of the electronic screening in semiconductors is still present in our model. We also obtain a substantial improvement for LA frequencies, which compare more favorably with experimental data than the results in Fig. 2. As regards the optical branches, there is an appreciable improvement with respect to previous results, indicating that the ODEDM cannot be neglected even in the calculation of the Raman frequency.

All the sources of the outstanding quantitative discrepancies between experiment and theory, which appear in Fig. 3, can be explained by considering our rough approximation on the energy-band structure, our restriction to a small-q expansion, and to the use of a local pseudopotential. The effects of the $\vec{q}\cdot\vec{p}$ approximation are particularly evident at the zone boundary, where the degenerations due to the symmetry of the crystals are not respected. On the other hand, the $\vec{q}\cdot\vec{p}$ approximation is also responsible for the anomalous curvature of the LA branch in the 100 direction. The high values of α which comes out from our calculation can be imputed partly to the above-mentioned approximations and partly to higher-order contributions, which have been folded back into first-order terms through the fulfillment of the acoustic sum rule. However, the introduction of this multiplicative factor in the ODEDM, beside giving consistency to our DM, allows us to account fairly well for the role of noncentral forces in Si. This shows that our DM with the off-diagonal elements evaluated to first order and scaled by α in order to satisfy the acoustic sum rule, provides a reliable description of the dielectric screening in semiconductors and is appropriate to study the lattice dynamics of Si.

Having established the importance of the ODEDM in the evaluation of the lattice vibrations in Si, it remains to be seen whether these results can be obtained from a calculation of the properties of NFE semiconductors based on RS perturbation theory, i.e., neglecting the band gap. It is generally accepted[5] that naïve perturbation theory can give an excellent approximation to the total energy of a metal, even if it furnishes a bad description of the electron energy-band structure. In order to see whether this argument remains appropriate to NFE semiconductors, we have performed the calculation of the phonon frequencies of Si with the dielectric matrix of a metal. As pointed out before, if one neglects corrections of order u^2 to the elements of the DM, this corresponds to evaluating the total energy up to third order in the pseudopotential. In this way, non-central forces are introduced via the ODEDM, and because of the absence of the band gap, there is no need of the acoustic sum rule.

In the calculation of the DM using RS perturbation theory, we found that in the $\vec{q}\to 0$ limit the ODEDM $\epsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}')$ with $\vec{G}$ and $\vec{G}'$ different from zero, differ within few percent from the corresponding elements of the semiconductor DM. For this reason, we expect the TA branches, which depend only on these matrix elements, to be well described if the ODEDM are multiplied by a factor α, as in the case of the BW calculation. However, while in the BW theory, the value of α is fixed unequivocally by the acoustic sum rule: in this case it remains an adjustable parameter. We found that a reasonable agreement with experiments is ob-

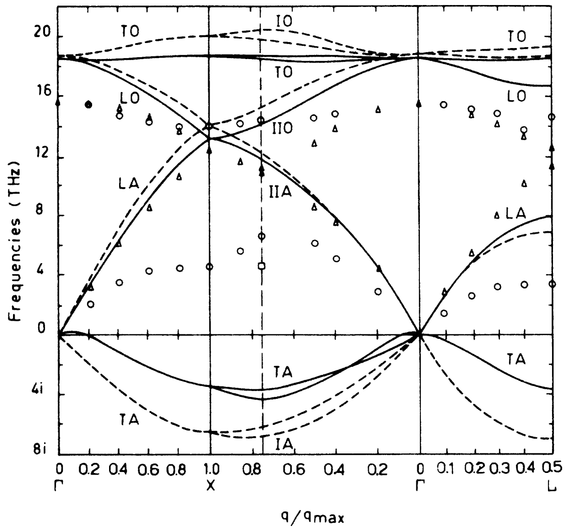

FIG. 5. Full curve: phonon dispersion curves of Si obtained using Lindhard dielectric function. Dashed curve: with the inclusion of exchange and correlation.

TABLE I. Elastic constants of Si: (a) experimental data from Ref. 27; (b) calculated with Lindhard dielectric function; (c) Lindhard dielectric function plus exchange and correlation; (d) with the dielectric matrix appropriate to a metal; (e) with the DM of a semiconductor. The units are 10^{11} N/m^2.

	a	b	c	d	e
c_{11}	1.66	1.89	1.97	2.44	1.15
c_{12}	0.64	2.40	2.14	2.33	0.73
c_{44}	0.79	−0.50	0.03	0.59	0.45

tained with $1.6 \leq \alpha \leq 3.0$. The results of the calculations with $\alpha = 1.6$, which is the lowest value giving completely real TA branches, are displayed in Fig. 4. Although the phonon frequencies are modified by increasing the value of α, we found that there is not an appreciable improvement of the agreement with experiments of the whole phonon spectrum.

For comparison we give in Fig. 5 the dispersion curves calculated in the diagonal scheme with the Lindhard[3] dielectric function (dashed curve). The full curve in the same figure shows the results obtained by including the exchange and correlation corrections. Once again the important feature which distinguishes the frequencies obtained from the full calculation with respect to the diagonal approximation, is the presence of real TA branches. We note that the anomalous behavior of the phonon curves calculated with BW perturbation theory, due to the small-$\vec{q}$ expansion, is not present in this calculation, because the $\vec{q}$ dependence of the matrix element is taken into account exactly. We note also that the exchange and correlation correction contribute to lift the imaginary branches, as shown in Fig. 5.

As regards the deviation of the value of α from 1, we think that, even in this case, it represents the effects of higher-order terms in RS perturbation expansion, which can be simulated by our first-order evaluation of the ODEDM. This is in agreement with the remarks of Brovman and Kagan,[25] who indicate that even fourth-order terms may be important in the description of the lattice dynamics and the elastic constants of NFE materials.

To gain some more insight on the effects of the ODEDM on crystal dynamics in silicon, we give in Table I the elastic constants calculated with different forms of screening, together with experimental data. We see that, both in the case of a semiconductor screening and in the case of DM suitable for a metal, the inclusion of the ODEDM improves sensibly the c_{44} elastic constant, leading to a good agreement with experiment. We note that the exchange and correlation corrections give a contribution of the same magnitude as the correction due to the ODEDM.

For the elastic constants c_{11} and c_{12} the screening appropriate to a semiconductor gives better results, while no appreciable improvement is found using a third-order RS perturbative approach. This can be understood, if we remember that while the c_{44} elastic constant depends exclusively upon Umklapp terms involving $\epsilon(\vec{q}+\vec{G}, \vec{q}+\vec{G}')$, with $\vec{G}$ and $\vec{G}'$ different from 0, the c_{11} and c_{12} constants depend also upon $\epsilon(\vec{q}, \vec{q}+\vec{G})$ and $\epsilon(\vec{q}+\vec{G}, \vec{q})$, whose behavior at long wavelength is strongly affected by the presence of the band gap.

---

*Present address: Cavendish Laboratory, Cambridge, United Kingdom.

[1] L. J. Sham, Proc. R. Soc. A **283**, 33 (1965).

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

[3] J. Lindhard, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. **28**, 8 (1954).

[4] D. Penn, Phys. Rev. **128**, 2093 (1962).

[5] V. Heine and D. Weaire, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970), Vol. 24, p. 249.

[6] S. Adler, Phys. Rev. **126**, 413 (1962); N. Wiser, Phys. Rev. **129**, 62 (1963).

[7] N. H. March and M. P. Tosi, Proc. R. Soc. A **330**, 373 (1972).

[8] E. Tosatti, C. Calandra, V. Bortolani, and C. M. Bertoni, J. Phys. C **5**, L299 (1972).

[9] C. M. Bertoni, V. Bortolani, C. Calandra, and E. Tosatti, Phys. Rev. Lett. **28**, 1578 (1972).

[10] R. M. Martin, Phys. Rev. **186**, 871 (1969).

[11] F. Herman, J. Phys. Chem. Solids **8**, 405 (1959).

[12] G. Dolling, in *Proceedings of the Symposium on Inelastic Scattering of Neutrons in Solids and Liquids, Chalk River, Canada, September, 1962* (International Atomic Energy Agency, Vienna, 1963), Vol. II, p. 37.

[13] L. J. Sham, Phys. Rev. **188**, 1431 (1969).

[14] J. C. Phillips, Phys. Rev. **166**, 832 (1968).

[15] S. K. Sinha, R. P. Gupta, and D. L. Pryce, Phys. Rev. Lett. **26**, 1324 (1967).

[16] T. Soma and A. Morita, J. Phys. Soc. Jap. **32**, 38 (1972); A. Morita, T. Soma, and T. Takeda, J. Phys. Soc. Jap. **32**, 29 (1972).

[17] V. Heine and R. O. Jones, J. Phys. C **2**, 719 (1969).

[18] R. Johnson, K. Dan. Vidensk. Selsk. Mat.-Fys. Medd. **37**, 9 (1970).

[19] J. P. Walter and M. L. Cohen, Phys. Rev. B **2**, 1821 (1970).

[20] M. A. Coulthard, J. Phys. C **3**, 820 (1970).

[21] M. H. Cohen, Phys. Rev. **130**, 1301 (1963).

[22] K. S. Singwi, M. P. Tosi, R. H. Land, and A. Sjolander, Phys. Rev. **176**, 589 (1968).

[23] R. W. Shaw and R. Pynn, J. Phys. C **2**, 2071 (1969).

[24] P. Lloyd and C. A. Sholl, J. Phys. C **1**, 1620 (1968).

[25] E. G. Brovman and Y. Kagan, Zh. Eksp. Teor. Fiz. **52**, 557 (1967) [Sov. Phys.-JETP **25**, 365 (1967)].

[26] P. N. Keating, Phys. Rev. **145**, 637 (1966).

[27] H. B. Huntington, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1958), Vol. 7, p. 213.