PHYSICAL REVIEW B VOLUME 54, NUMBER 23 15 DECEMBER 1996-I

Linear-response theory and lattice dynamics: A muffin-tin-orbital approach

S. Y. Savrasov*
Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1, D-70569 Stuttgart, Germany

(Received 8 March 1996; revised manuscript received 13 September 1996)

A detailed description of a method for calculating static linear-response functions in the problem of lattice dynamics is presented. The method is based on density-functional theory and it uses linear muffin-tin orbitals as a basis for representing first-order corrections to the one-electron wave functions. This makes it possible to greatly facilitate the treatment of the materials with localized orbitals. We derive variationally accurate expressions for the dynamical matrix. We also show that large incomplete-basis-set corrections to the first-order changes in the wave functions exist and can be explicitly calculated. Some useful hints on the **k**-space integration for metals and the self-consistency problem at long wavelengths are also given. As a test application we calculate bulk phonon dispersions in Si and find good agreement between our results and experiments. As another application, we calculate lattice dynamics of the transition-metal carbide NbC. The theory reproduces the major anomalies found experimentally in its phonon dispersions. The theory also predicts an anomalous behavior of the lowest transverse acoustic mode along the $(\xi\xi0)$ direction. Most of the calculated frequencies agree within a few percent with those measured. [S0163-1829(96)05348-9]

I. INTRODUCTION

Response of electrons to a static external field is one of the important characteristics of a solid which can be uniquely determined within density functional theory (DFT).[1] The use of the local density approximation (LDA) (Refs. 2 and 3) in the linear-response problem has by now become a common and well-established method for determining various ground-state properties of real materials. These, first of all, include static dielectric, structural, and vibrational properties of semiconductors and insulators such as the screening response to point charges and electric fields, effective charges, dielectric and piezoelectric constants, as well as whole phonon spectra.[4–10] *Ab initio* calculations of the wave-vector dependent lattice-dynamical properties for metallic systems have also recently been performed.[11–17] Among them, transition metals, their alloys, and compounds provide one of the fascinating areas in the study of phonons in crystal lattice. This is because, in addition to the richness and variety of structure of their phonon dispersion curves, these materials also often exhibit lattice instabilities and relatively high-temperature (8 K–23 K) superconductivity, in which phonons play a fundamental role. Here, density-functional based linear-response calculations provide an important first step in studying such phenomena as electron-phonon interactions and transport properties[15] which describe phonon-limited electrical and thermal resistivities, renormalization of specific heat (electron-mass enhancement) as well as superconducting transition temperatures. These properties are connected with the real low-energy excitation spectrum of a metal and will be discussed in detail in a publication[18] followed by this paper.

Initially, two methods have been developed to deal with the perturbations which break the periodicity of the original lattice. The first one, known as the direct or supercell approach,[19] considers the perturbation with wave vector **q** which is periodical in the supercell structure. This is possible if the wave vector is commensurate with the reciprocal lattice of the supercell and only tractable computationally if the size of the supercell is not large. This limits the applications to high-symmetry wave vectors. The same technique can be applied to calculate the interplanar force constants in direct space.[20] The dynamical matrix is found for any **q** using the Fourier transform provided that the interatomic interactions of a solid are of short range. Despite the severe computational-cost restrictions, the supercell approach has two important advantages: (i) the electronic response and lattice dynamics can be studied using programs for self-consistent band-structure calculations which are standardly used in condensed-matter physics, and, as a consequence, (ii) all non-linear-response coefficients are easily obtained. Note, however, that third-order nonlinear coefficients can also be accessed within the linear-response approach[21] just like forces are found within the density-functional total-energy method. We shall discuss this statement in more detail later in this paper.

The second method to deal with the perturbations is known as the perturbative approach. If the external perturbation is weak one can use standard perturbation theory and expand the first-order corrections to the one-electron wave functions in the unperturbed Bloch states of the original crystal. Usually it is done by introducing the so-called independent-particle polarizability function in terms of which the screened perturbation is found by inverting the static dielectric matrix of a crystal.[22] Previously, due to a rapid progress made in the microscopic theory of the phonon spectra in free-electron-like metals through the development and application of the pseudopotential technique, a plane-wave representation was used for all the relevant quantities in the calculation. However, already in the case of covalent semiconductors with sufficiently weak pseudopotential, the convergency of the polarizability with respect to a number of plane waves becomes slow and there are only a few attempts to compute phonon spectra within this framework.[23] The situation becomes worse for materials with localized orbitals. The most time-consuming step in this approach is connected with the problem of summation over high-energy states which at least requires their calculation. Another problem is connected with the inversion of a large dielectric matrix.

The above mentioned drawbacks of the perturbative approach have been circumvented using the solid-state generalization[24,25] of the Sternheimer method.[26] In this reformulated linear-response method, the first-order corrections to the unperturbed wave functions are found by solving the Sternheimer equation (which is the Schrödinger equation to linear order) directly without using the expansions over unperturbed states. This avoids the summation problem of the perturbation theory. The screening of the external perturbation is calculated self-consistently within DFT in close analogy with what is done in standard band-structure calculations. This avoids the inversion problem. The present formulation is thus very efficient computationally which is demonstrated by an increasing number of applications to the problem of lattice dynamics in recent years.[6–17]

In order to solve the Sternheimer equation, one has to construct a rapidly convergent basis set for representing the first-order perturbations. This is important because these corrections as well as the unperturbed wave functions oscillate in the core region. In the free-electron-like metals, broad band semiconductors, and insulators, this problem can be eliminated by the pseudopotential approximation and the majority of the applications performed so far use plane-wave basis sets.[6,8,10,14,17] Unfortunately, with decreasing bandwidth, the plane-wave expansion of the pseudo-wave-functions converges more slowly and it becomes less advantageous to employ pseudopotentials. Indeed, until most recently,[11–13,15–17] the literature contains no *ab initio* calculations of phonon dispersions for transition-metal systems.

In the present paper we describe an efficient all-electron generalization of the linear-response approach introduced in Refs. 24 and 25. (A brief report of this work has been published already[12].) The first-order corrections are represented in terms of the muffin-tin (MT) basis sets which greatly facilitate the treatment of localized valence wave functions. While the approach developed in the paper is general and can be applied to any kind of localized-orbital representation, we use the linear-muffin-tin-orbital (LMTO) method[27] as a framework of such all-electron formulation.

There are two problems addressed in this paper which are connected with the use of MT-basis functions in a linear-response method. The first problem concerns the construction of a variational solution of the Sternheimer equation. This is necessary because the unperturbed energy bands and wave functions are obtained within the LMTO method by applying the Rayleigh-Ritz variational principle. They are, therefore, not exact solutions of the one-electron Schrödinger equation. As first shown by Pulay,[28] the use of variational solutions gives rise to the incomplete-basis-set (IBS) corrections in force calculations. The IBS corrections must be carefully accounted for to get accurate forces in the LMTO method.[29] These corrections also exist and must be taken into account when calculating the first-order changes in the wave functions and the dynamical matrix within the linear-response approach.

The second problem discussed in the paper is connected with finding the change in the basis functions due to the perturbation. Since the one-electron wave function in the LMTO method is represented by the expansion coefficients in the MT-basis set, under static external perturbation, such as the displacement of a nucleus from its equilibrium position, the change in the wave function will be described by both the change in the expansion coefficients and the change in the basis set. The contribution from the change in the basis set is important because the original basis set is tailored to the one-electron potential and must therefore be reconstructed to account for the specifics of the perturbation. It should be noted that this contribution is not taken into account in the standard perturbation theory, where only the change in the expansion coefficients is taken into account.

The rest of the paper is organized as follows. The variational formulation of the linear-response approach is described in Sec. II. Implementation of the theory in the framework of the LMTO method is described in Sec. III. Application of the method to phonon spectra in Si and NbC is given in Sec. IV. Section V concludes the paper.

II. THEORY

A. Density-functional linear response

Within density-functional theory, the problem of calculating the lattice dynamics essentially amounts to finding the change in the electronic charge density induced by the presence of a phonon with wave vector $\mathbf{q}$. Consider a lattice with a few atoms in the unit cell given by the positions $\mathbf{R}+\mathbf{t}$, where $\mathbf{R}$ are the basis vectors and $\mathbf{t}$ are the translation vectors. Suppose that the atoms are displaced from their equilibrium positions by a small amount:

$$\Delta\mathbf{t}_R=\mathbf{Q}_R\exp(+i\mathbf{qt})+\mathbf{Q}_R^*\exp(-i\mathbf{qt}), \tag{1}$$

where $\mathbf{Q}_R$ is a complex (infinitesimal) polarization vector and $\mathbf{q}$ is the phonon wave vector. The presence of such a displacement field changes the bare Coulomb potential as follows:

$$\tilde{V}_{\text{ext}}(\mathbf{r})=\sum_{R,t}\frac{-Z_Re^2}{|\mathbf{r}-\mathbf{R}-\mathbf{t}-\Delta\mathbf{t}_R|}, \tag{2}$$

where Z_R are the nuclei charges. By expanding this expression to first order in the displacements, we obtain that the crystal is perturbed by the static external field:

$$\Delta^{(1)}V_{\text{ext}}(\mathbf{r})=\sum_R\mathbf{Q}_R\sum_t e^{+i\mathbf{qt}}\nabla\frac{Z_Re^2}{|\mathbf{r}-\mathbf{R}-\mathbf{t}|}+\sum_R\mathbf{Q}_R^*\sum_t e^{-i\mathbf{qt}}\nabla\frac{Z_Re^2}{|\mathbf{r}-\mathbf{R}-\mathbf{t}|}, \tag{3}$$

which is represented as a superposition of two traveling waves with wave vectors $+\mathbf{q}$ and $-\mathbf{q}$, i.e.,

$$\Delta^{(1)}V_{\text{ext}}(\mathbf{r})=\sum_R\mathbf{Q}_R\frac{\delta^+V_{\text{ext}}(\mathbf{r})}{\delta\mathbf{R}}+\sum_R\mathbf{Q}_R^*\frac{\delta^-V_{\text{ext}}(\mathbf{r})}{\delta\mathbf{R}}. \tag{4}$$

To shorten the notations, we will sometimes omit $\delta\mathbf{R}$ from this definition and, therefore, $\Delta^{(1)}V_{\text{ext}}=\Sigma\mathbf{Q}\delta^+V_{\text{ext}}+\Sigma\mathbf{Q}^*\delta^-V_{\text{ext}}$. Both components δ^+V_{ext} and δ^-V_{ext} have the property $[\delta^{\pm}V_{\mathrm{ext}}]^*=\delta^{\mp}V_{\mathrm{ext}}$ and translate like Bloch waves in the original crystal: $\delta^{\pm}V_{\mathrm{ext}}(\mathbf{r}+\mathbf{t})=\exp(\pm i\mathbf{qt})\delta^{\pm}V_{\mathrm{ext}}(\mathbf{r})$.

The first-order change in the charge density, $\Delta^{(1)}\rho$, induced by the perturbation (3) is represented in the same form as (4), i.e., $\Delta^{(1)}\rho=\Sigma\mathbf{Q}\delta^{+}\rho+\Sigma\mathbf{Q}^*\delta^{-}\rho$ and it is expressed in terms of the one-electron wave functions $\psi_{\mathbf{k}j}$ and their first-order corrections $\delta^{+}\psi_{\mathbf{k}j}$ and $\delta^{-}\psi_{\mathbf{k}j}$ as follows:

$$\delta^{\pm}\rho=\sum_{\mathbf{k}j} f_{\mathbf{k}j}(\delta^{\pm}\psi^*_{\mathbf{k}j}\psi_{\mathbf{k}j}+\psi^*_{\mathbf{k}j}\delta^{\pm}\psi_{\mathbf{k}j}),\tag{5}$$

where $f_{\mathbf{k}j}$ are the occupation numbers, $\mathbf{k}$ lies in the first Brillouin zone, and j is the band index. The first-order correction $\delta^{\pm}\psi_{\mathbf{k}j}\equiv|\delta^{\pm}\mathbf{k}j\rangle=(\delta^{\mp}\psi_{\mathbf{k}j})^*$ is a Bloch wave with vector $\mathbf{k}\pm\mathbf{q}$ and it is the solution of the Sternheimer equation, which is the Schrödinger equation to linear order:

$$(-\nabla^2+V_{\mathrm{eff}}-\epsilon_{\mathbf{k}j})|\delta^{\pm}\mathbf{k}j\rangle+\delta^{\pm}V_{\mathrm{eff}}|\mathbf{k}j\rangle=0.\tag{6}$$

Here, V_{eff} is the effective DFT potential and $\delta^{\pm}V_{\mathrm{eff}}$ is the change in the potential which is the external perturbation screened by the induced charge density:

$$\delta^{\pm}V_{\mathrm{eff}}=\delta^{\pm}V_{\mathrm{ext}}+e^2\int\frac{\delta^{\pm}\rho}{|\mathbf{r}-\mathbf{r}'|}+\frac{dV_{\mathrm{xc}}}{d\rho}\delta^{\pm}\rho,\tag{7}$$

where the exchange-correlation effects are taken into account in the local density approximation. In Eq. (6) we have dropped the term with the first-order corrections to the one-electron energies: $\delta^{\pm}\epsilon_{\mathbf{k}j}=\langle\mathbf{k}j|\delta^{\pm}V_{\mathrm{eff}}|\mathbf{k}j\rangle$, which are equal to zero if $\mathbf{q}\neq 0$. The Eqs. (5)–(7) must be solved self-consistently, i.e., (i) one has to solve (6) with the external perturbation $\delta^{\pm}V_{\mathrm{ext}}$, or the one screened by some guessed $\delta^{\pm}\rho$; (ii) find the induced charge density according to (5); and (iii) find a new $\delta^{\pm}V_{\mathrm{eff}}$ after (7). Steps (i)–(iii) are repeated until input and output $\delta^{\pm}\rho$ coincide within a required accuracy. This is analogous to finding the unperturbed quantities $\psi_{\mathbf{k}j}$, ρ, and V_{eff} in standard band-structure calculations.

B. Dynamical matrix

We must now solve two problems in order to calculate the lattice dynamics: to develop a method for solving the Sternheimer equation (6), and to find an expression for the dynamical matrix. The general strategy employed in the following is to consider the dynamical matrix as a functional of the first-order perturbations. Expanding $\delta^{\pm}\psi_{\mathbf{k}j}$ in terms of the MT-basis functions will lead, under the stationarity condition, to a set of matrix equations which represent a variational solution to Eq. (6).

The variational formulation of the linear-response problem is required because the original states $\psi_{\mathbf{k}j}$ are normally found not as exact but variational solutions of the one-electron Schrödinger equation. In an all-electron method such as the LMTO method, the wave function is expanded in terms of the basis set $|\chi^{\mathbf{k}}_{\alpha}\rangle$:

$$|\mathbf{k}j\rangle=\sum_{\alpha}|\chi^{\mathbf{k}}_{\alpha}\rangle A^{\mathbf{k}j}_{\alpha}.\tag{8}$$

The total energy is then considered as a functional of only the expansion coefficients $A^{\mathbf{k}j}_{\alpha}$, which are found by applying the Rayleigh-Ritz variational principle. This leads to the following matrix eigenvalue problem:

$$\sum_{\alpha}\langle\chi^{\mathbf{k}}_{\beta}|-\nabla^2+V_{\mathrm{eff}}-\epsilon_{\mathbf{k}j}|\chi^{\mathbf{k}}_{\alpha}\rangle A^{\mathbf{k}j}_{\alpha}=0,\tag{9}$$

which, in particular, allows all nonspherical terms in the potential to be taken explicitly into account.

In the problem of lattice dynamics the second-order change in the total energy must be found. It is obtained by expanding the total energy with respect to the change in the external potential (nuclei displacements) up to second order, i.e., $E=E_0+\Delta^{(1)}E+\Delta^{(2)}E$. The LDA total energy is given by the standard expression

$$E=\sum_{\mathbf{k}j} f_{\mathbf{k}j}\epsilon_{\mathbf{k}j}-\int\rho V_{\mathrm{eff}}+\int\rho V_{\mathrm{ext}}+\frac{e^2}{2}\int\frac{\rho\rho}{|\mathbf{r}-\mathbf{r}'|}+\int\rho\epsilon_{\mathrm{xc}},\tag{10}$$

where ϵ_{xc} is the density of exchange-correlation energy as given by LDA. The expression for the first-order change in the total energy has been discussed many times in the literature (for example, see Ref. 29). Since our purpose is to proceed with the second-order variation, we give a brief description. $\Delta^{(1)}E$ is obtained by expanding Eq. (10) with respect to the displacements. It is given by

$$\begin{aligned}\Delta^{(1)}E={}&\sum_{\mathbf{k}j}\{\Delta^{(1)}f_{\mathbf{k}j}\epsilon_{\mathbf{k}j}+f_{\mathbf{k}j}\Delta^{(1)}\epsilon_{\mathbf{k}j}\}-\int\Delta^{(1)}\rho V_{\mathrm{eff}}\\&-\int\rho\Delta^{(1)}V_{\mathrm{eff}}+\int\Delta^{(1)}\rho V_{\mathrm{ext}}+\int\rho\Delta^{(1)}V_{\mathrm{ext}}\\&+e^2\int\frac{\rho\Delta^{(1)}\rho}{|\mathbf{r}-\mathbf{r}'|}+\int\Delta^{(1)}\rho V_{\mathrm{xc}},\end{aligned}\tag{11}$$

where the first-order change in the eigenvalues is given by

$$\Delta^{(1)}\epsilon_{\mathbf{k}j}=\langle\mathbf{k}j|\Delta^{(1)}V_{\mathrm{eff}}|\mathbf{k}j\rangle+\{\langle\Delta^{(1)}\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\mathbf{k}j\rangle+\mathrm{c.c.}\}.\tag{12}$$

Here $H=-\nabla^2+V_{\mathrm{eff}}$ and the consequence of the orthonormalization property of the wave functions has been used, i.e.,

$$\langle\Delta^{(1)}\mathbf{k}j|\mathbf{k}j\rangle+\langle\mathbf{k}j|\Delta^{(1)}\mathbf{k}j\rangle=0.\tag{13}$$

We now insert Eq. (12) to Eq. (11) and sum over the occupied states in the first contribution in (12). The obtained contribution will cancel the integral with $\rho\Delta^{(1)}V_{\mathrm{eff}}$ in (11). According to the electron-number-conservation condition, which always holds when using the tetrahedron method for the integration over the Brillouin zone, the term containing the change in the occupation numbers vanishes, i.e.,

$$\begin{aligned}\sum_{\mathbf{k}j}\Delta^{(1)}f_{\mathbf{k}j}\epsilon_{\mathbf{k}j}&=\sum_{\mathbf{k}j}\delta(\epsilon_F-\epsilon_{\mathbf{k}j})\epsilon_{\mathbf{k}j}\{\Delta^{(1)}\epsilon_F-\Delta^{(1)}\epsilon_{\mathbf{k}j}\}\\&=\epsilon_F\Delta^{(1)}N_{\mathrm{val}}=0.\end{aligned}\tag{14}$$

Therefore, we arrive at the following expression for the first-order change in the total energy

$$\Delta^{(1)}E=\sum_{\mathbf{k}j} f_{\mathbf{k}j}\{\langle\Delta^{(1)}\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\mathbf{k}j\rangle+\text{c.c.}\}+\int\rho\Delta^{(1)}V_{\text{ext}}+\int\Delta^{(1)}\rho\left\{V_{\text{ext}}+e^2\int\frac{\rho}{|\mathbf{r}-\mathbf{r}'|}+V_{\text{xc}}-V_{\text{eff}}\right\}. \quad (15)$$

If the electron density is self-consistent, the expression in curly brackets of the integral with $\Delta^{(1)}\rho$ is equal to zero, and the obtained result is known as the Hellmann-Feynman force (second contribution) plus the incomplete-basis-set (or Pulay) correction [first contribution in (15)].

$\Delta^{(1)}E$ is equal to zero if the lattice is in the equilibrium. The second-order change, $\Delta^{(2)}E$, which is expressed via the dynamical matrix of a solid, is found by performing one more variation of Eq. (11):

$$\begin{aligned}\Delta^{(2)}E=&\sum_{\mathbf{k}j}\{f_{\mathbf{k}j}\Delta^{(2)}\epsilon_{\mathbf{k}j}+2\Delta^{(1)}f_{\mathbf{k}j}\Delta^{(1)}\epsilon_{\mathbf{k}j}+\Delta^{(2)}f_{\mathbf{k}j}\epsilon_{\mathbf{k}j}\}\\&-\int\rho\Delta^{(2)}V_{\text{eff}}-2\int\Delta^{(1)}\rho\Delta^{(1)}V_{\text{eff}}+\int\rho\Delta^{(2)}V_{\text{ext}}\\&+2\int\Delta^{(1)}\rho\Delta^{(1)}V_{\text{ext}}+e^2\int\frac{\Delta^{(1)}\rho\Delta^{(1)}\rho}{|\mathbf{r}-\mathbf{r}'|}\\&+\int\Delta^{(1)}\rho\Delta^{(1)}V_{\text{xc}}+\int\Delta^{(2)}\rho\left\{V_{\text{ext}}+\int\frac{\rho}{|\mathbf{r}-\mathbf{r}'|}\right.\\&\left.+V_{\text{xc}}-V_{\text{eff}}\right\},\end{aligned} \quad (16)$$

where the last integral with $\Delta^{(2)}\rho$ can be dropped since the expression in curly brackets is equal to zero if the unperturbed charge density is self-consistent.

It is easy to prove that all the contributions with change in the occupation numbers vanish if the wave vector $\mathbf{q}$ of the perturbation is not equal to zero. In this case, as we have mentioned already, first-order changes in the eigenvalues are equal to zero. It is directly seen from the standard perturbation theory expression:

$$\begin{aligned}\Delta^{(1)}\epsilon_{\mathbf{k}j}&=\langle\mathbf{k}j|\Delta^{(1)}V_{\text{eff}}|\mathbf{k}j\rangle\\&=\sum_R\left\{\mathbf{Q}_R\left\langle\mathbf{k}j\left|\frac{\delta^+V_{\text{eff}}}{\delta R}\right|\mathbf{k}j\right\rangle+\mathbf{Q}_R^*\left\langle\mathbf{k}j\left|\frac{\delta^-V_{\text{eff}}}{\delta R}\right|\mathbf{k}j\right\rangle\right\}\\&\equiv\sum\{\mathbf{Q}\langle\mathbf{k}j|\delta^+V_{\text{eff}}|\mathbf{k}j\rangle+\mathbf{Q}^*\langle\mathbf{k}j|\delta^-V_{\text{eff}}|\mathbf{k}j\rangle\},\end{aligned} \quad (17)$$

where we have used the definition like (4) for the change in the effective potential due to the displacements in the form (1). Each matrix element in this expression is equal zero since $\psi^*_{\mathbf{k}j}\psi_{\mathbf{k}j}$ is periodical at the original lattice while $\delta^{\pm}V_{\text{eff}}$ translates with wave vector $\pm\mathbf{q}$. The same arguments can be applied if we consider the expression (12) for $\Delta^{(1)}\epsilon_{\mathbf{k}j}$. In this case, matrix elements containing an incomplete-basis-set correction [second contribution in (12)] are also equal to zero since the expression under the integral there also translates with wave vector $\pm\mathbf{q}$. This is so since

$$|\Delta^{(1)}\mathbf{k}j\rangle=\sum\mathbf{Q}|\delta^+\mathbf{k}j\rangle+\sum\mathbf{Q}^*|\delta^-\mathbf{k}j\rangle \quad (18)$$

is a superposition of two traveling wave with wave vectors $\mathbf{k}\pm\mathbf{q}$, while the one-electron Hamiltonian $-\nabla^2+V_{\text{eff}}$ is periodical at the original structure and $\psi_{\mathbf{k}j}$ translates with wave vector $\mathbf{k}$.

Since we have proved that $\Delta^{(1)}\epsilon_{\mathbf{k}j}=0$ for $\mathbf{q}\neq 0$, then $\Delta^{(1)}f_{\mathbf{k}j}=\delta(\epsilon_F-\epsilon_{\mathbf{k}j})(\Delta^{(1)}\epsilon_F-\Delta^{(1)}\epsilon_{\mathbf{k}j})$ is also equal to zero. The contribution in (16) from the second-order change in the occupation numbers reads

$$\begin{aligned}\sum_{\mathbf{k}j}\Delta^{(2)}f_{\mathbf{k}j}\epsilon_{\mathbf{k}j}=&\sum_{\mathbf{k}j}\epsilon_{\mathbf{k}j}\{\Delta^{(1)}\epsilon_F-\Delta^{(1)}\epsilon_{\mathbf{k}j}\}^2\frac{d}{d\epsilon_F}\delta(\epsilon_F-\epsilon_{\mathbf{k}j})\\&+\epsilon_F\sum_{\mathbf{k}j}\delta(\epsilon_F-\epsilon_{\mathbf{k}j})\{\Delta^{(2)}\epsilon_F-\Delta^{(2)}\epsilon_{\mathbf{k}j}\}.\end{aligned} \quad (19)$$

Since $\Delta^{(1)}\epsilon_F=\Delta^{(1)}\epsilon_{\mathbf{k}j}=0$, the whole expression is equal to $\epsilon_F\Delta^{(2)}N_{\text{val}}=0$. We have thus proved that for the most interesting case $\mathbf{q}\neq 0$ we can omit the contributions from the change in the occupation numbers.

Two comments should be said for the limiting case $\mathbf{q}\rightarrow 0$. First, if the crystal is an insulator, the change in the occupation numbers is always equal to zero in all orders. However, due to the appearance of a longitudinal electric field care should be taken when calculating the LO-TO splitting. The details of such calculations can, for instance, be found in Ref. 6. Second, for metals it is much easier to find the limit $\mathbf{q}\rightarrow 0$ in the final expression for the dynamical matrix rather than to work it out from the starting expression (16). Taking this limit involves the substitution of the intraband contribution by the Fermi surface integral, i.e.,

$$\sum_{\mathbf{k}}\frac{f_{\mathbf{k}j}-f_{\mathbf{k}+\mathbf{q}j}}{\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}+\mathbf{q}j}}\rightarrow-\sum_{\mathbf{k}}\delta(\epsilon_F-\epsilon_{\mathbf{k}j}). \quad (20)$$

Since it can be done straightforwardly within our formulation we shall not return to this point later in this paper.

We now discuss the second-order change in the eigenvalues. It is obtained from Eq. (12) and reads

$$\begin{aligned}\Delta^{(2)}\epsilon_{\mathbf{k}j}=&\langle\mathbf{k}j|\Delta^{(2)}V_{\text{eff}}|\mathbf{k}j\rangle+\{\langle\Delta^{(2)}\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\mathbf{k}j\rangle+\text{c.c.}\}\\&+2\langle\Delta^{(1)}\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\Delta^{(1)}\mathbf{k}j\rangle\\&+2\langle\Delta^{(1)}\mathbf{k}j|\Delta^{(1)}V_{\text{eff}}|\mathbf{k}j\rangle+2\langle\mathbf{k}j|\Delta^{(1)}V_{\text{eff}}|\Delta^{(1)}\mathbf{k}j\rangle.\end{aligned} \quad (21)$$

Again, we keep the second contribution in this expression due to incompleteness of the basis set just like we did it in the expression (12) for $\Delta^{(1)}\epsilon_{\mathbf{k}j}$. Inserting (21) into the expression (16) for $\Delta^{(2)}E$ and performing the summation over the occupied states, we transform the first matrix element in (21) to the integral with $\rho\Delta^{(2)}V_{\text{eff}}$ and the last two matrix elements to the integral with $2\Delta^{(1)}\rho\Delta^{(1)}V_{\text{eff}}$. The final expression for the second-order change in the total energy is given by

$$\Delta^{(2)}E=\sum_{\mathbf{k}j} f_{\mathbf{k}j}\{\langle\Delta^{(2)}\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\mathbf{k}j\rangle+\text{c.c.}\}$$
$$+2\sum_{\mathbf{k}j} f_{\mathbf{k}j}\langle\Delta^{(1)}\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\Delta^{(1)}\mathbf{k}j\rangle$$
$$+\int \Delta^{(1)}\rho\Delta^{(1)}V_{\text{eff}}+\int \Delta^{(1)}\rho\Delta^{(1)}V_{\text{ext}}$$
$$+\int \rho\Delta^{(2)}V_{\text{ext}}. \tag{22}$$

In order to derive the expression for the dynamical matrix Λ we should separate out the infinitesimal polarization vectors $\mathbf{Q}_R$. We define

$$\Delta^{(2)}E=\sum_{R'\mu'R\mu}\Lambda_{R'R}^{\mu'\mu}(\mathbf{q})Q_{R'\mu'}Q^*_{R\mu}+\text{c.c.}, \tag{23}$$

where $\{\mu\}$ denote directions of the polarization. The first-order changes in the quantities like $\Delta^{(1)}V$ and $\Delta^{(1)}\psi_{\mathbf{k}j}$ are given by the formulas (4) and (18). Changes in the second order are easy to find. As follows from Eq. (2),

$$\Delta^{(2)}V_{\text{ext}}=\sum_{R,\mu\mu'}\left\{Q_{R\mu}Q_{R\mu'}\sum_t e^{2i\mathbf{q}\mathbf{t}}\nabla_\mu\nabla_{\mu'}\frac{-Z_Re^2/2}{|\mathbf{r}-\mathbf{R}-\mathbf{t}|}\right.$$
$$\left.+Q_{R\mu}Q^*_{R\mu'}\sum_t\nabla_\mu\nabla_{\mu'}\frac{-Z_Re^2/2}{|\mathbf{r}-\mathbf{R}-\mathbf{t}|}\right\}+\text{c.c.} \tag{24}$$

It is convenient to rewrite this expression in the form

$$\Delta^{(2)}V_{\text{ext}}=\sum_{R'\mu'R\mu}\left\{Q_{R\mu}Q_{R'\mu'}\frac{\delta^+\delta^+V_{\text{ext}}}{\delta R'_{\mu'}\delta R_\mu}\right.$$
$$\left.+Q_{R\mu}Q^*_{R'\mu'}\frac{\delta^+\delta^-V_{\text{ext}}}{\delta R'_{\mu'}\delta R_\mu}\right\}+\text{c.c.}, \tag{25}$$

by introducing the following notations

$$\frac{\delta^+\delta^-V_{\text{ext}}}{\delta R'_{\mu'}\delta R_\mu}\equiv\delta^+\delta^-V_{\text{ext}}=\delta_{R'R}\sum_t\nabla_{\mu'}\nabla_\mu\frac{-Z_Re^2/2}{|\mathbf{r}-\mathbf{R}-\mathbf{t}|}. \tag{26}$$

Similar definition can also be found for $\delta^+\delta^+V_{\text{ext}}$. Omitting indexes $R'\mu'R\mu$, it is directly seen that the variation $\delta^+\delta^-V_{\text{ext}}$ is a periodical function at the original lattice while the function $\delta^+\delta^+V_{\text{ext}}$ translates like a wave of wave vector $2\mathbf{q}$. The expansion like (25) can be written for any second-order change. For example,

$$|\Delta^{(2)}\mathbf{k}j\rangle=\sum\{QQ|\delta^+\delta^+\mathbf{k}j\rangle+QQ^*|\delta^+\delta^-\mathbf{k}j\rangle\}+\text{c.c.} \tag{27}$$

represents a second-order change in the wave function as a superposition of Bloch waves with wave vectors $\mathbf{k}+2\mathbf{q}$, $\mathbf{k}$, and $\mathbf{k}-2\mathbf{q}$. Using the definitions like (27) in the expression (22) and taking into account the form (23), we finally arrive at the expression for the dynamical matrix:

$$\Lambda_{R'R}^{\mu'\mu}(\mathbf{q})\equiv\Lambda=\sum_{\mathbf{k}j} f_{\mathbf{k}j}\langle\delta^+\delta^-\mathbf{k}j+\delta^-\delta^+\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\mathbf{k}j\rangle$$
$$+\sum_{\mathbf{k}j} f_{\mathbf{k}j}2\langle\delta^+\mathbf{k}j|H-\epsilon_{\mathbf{k}j}|\delta^+\mathbf{k}j\rangle+\int\delta^+\rho\delta^-V_{\text{eff}}$$
$$+\int\delta^+\rho\delta^-V_{\text{ext}}+\int\rho\delta^+\delta^-V_{\text{ext}}. \tag{28}$$

We see that this expression does not contain the second-order contributions with wave vectors $\pm 2\mathbf{q}$. $|\delta^\pm\delta^\mp\mathbf{k}j\rangle$ are the functions of wave vector $\mathbf{k}$ and only they contribute to the matrix elements with $|\mathbf{k}j\rangle$. $\delta^+\delta^-V_{\text{ext}}$ is periodical and only this gives a nonzero contribution to the integral with ρ. (Therefore, we consider the operator $\delta^\pm$ as a variation of a Bloch wave, whereby $\pm\mathbf{q}$ gets added to its wave vector.)

The first term in expression (28) is not zero if the unperturbed states are approximate solutions found from the eigenvalue problem (9). If, on the other hand, one neglects it, and performs a variation of Λ with respect to the first-order corrections, the self-consistent linear-response equations (6), (7) will be recovered. The expression (28) is variational with respect to the first-order changes in the wave functions just like the unperturbed total energy is variational with respect to the unperturbed states $|\mathbf{k}j\rangle$. This property of the density functional follows from the Hohenberg-Kohn-Sham variational principle. It is not unique and represent a particular case of the powerful ''$2n+1$'' theorem of perturbation theory: the knowledge of the perturbations in the wave functions up to (n)th order allows one to find the $(2n+1)$th correction to the eigenenergy.[30] A recent publication[21] gives a full description of generalizing density functional theory to *arbitrary* order of perturbation. There, the variational properties of even derivatives of the total energy were also discussed in detail. An excellent demonstration of those by direct minimizing the dynamical matrix in terms of the conjugate-gradient method has also appeared recently in the literature.[8]

The formulated variational principle is important for us since the calculation of the dynamical matrix can be done accurately: while the first-order changes in the wave functions and the charge densities are only variationally accurate, the error will be of second order with respect to the error in $|\delta^\pm\mathbf{k}j\rangle$. In particular, the convergence of the dynamical matrix during the iterations towards self-consistency is much faster than the convergence of the induced charge density. At its minimum, expression (28) contains no second and third terms and it may therefore be interpreted as the Hellmann-Feynman result (last two terms) plus incomplete-basis-set correction (the first contribution). The latter is the analog of the ''Pulay force'' known from force calculations.

It is worth to pointing out that the knowledge of the first-order corrections to the wave functions allows us to consider changes in the total energy up to *third* order, in the same way as the zeroth-order unperturbed states allow calculating such first-order derivatives as, for instance, forces. Consequently, third-order anharmonicity constants, Grüneisen parameters and other nonlinear coefficients are, in principle, easily accessed within the linear-response formalism. One of such applications has also appeared recently in the literature.[31]

C. First-order corrections

We now turn out to the construction of the basis functions which represent the first-order perturbations. As a first illustration, let us take an expression of the standard perturbation theory:

$$|\delta^{\pm}\mathbf{k}j\rangle=\sum_{j'}|\mathbf{k}\pm\mathbf{q}j'\rangle\frac{\langle\mathbf{k}\pm\mathbf{q}j'|\delta^{\pm}V_{\text{eff}}|\mathbf{k}j\rangle}{\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}\pm\mathbf{q}j'}}.\tag{29}$$

The sum here should, in principle, go over both $\mathbf{k}'$ and j' indexes. However, since matrix element $\langle\mathbf{k}'j'|\delta^{\pm}V_{\text{eff}}|\mathbf{k}j\rangle$ is not zero only for $\mathbf{k}'=\mathbf{k}\pm\mathbf{q}$, the summation over $\mathbf{k}'$ can be omitted. After representing the unperturbed state $|\mathbf{k}\pm\mathbf{q}j'\rangle$ as an expansion (8), the expression (29) can be rewritten as follows:

$$|\delta^{\pm}\mathbf{k}j\rangle=\sum_{\alpha}|\chi_{\alpha}^{\mathbf{k}\pm\mathbf{q}}\rangle\delta^{\pm}A_{\alpha}^{\mathbf{k}j},\tag{30}$$

where the change in the expansion coefficients, $\delta^{\pm}A_{\alpha}^{\mathbf{k}j}$ is given by

$$\delta^{\pm}A_{\alpha}^{\mathbf{k}j}=\sum_{j'}A_{\alpha}^{\mathbf{k}\pm\mathbf{q}j'}\frac{\langle\mathbf{k}\pm\mathbf{q}j'|\delta^{\pm}V_{\text{eff}}|\mathbf{k}j\rangle}{\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}\pm\mathbf{q}j'}}.\tag{31}$$

From this simple illustration we see that the expression (30) does not take into account the change in the MT-basis set with respect to the perturbation and, therefore, it can be slowly convergent. Since the unperturbed state is given by expansion (8), the first-order change $|\delta^{\pm}\mathbf{k}j\rangle$ must include both the change $\delta^{\pm}A_{\alpha}^{\mathbf{k}j}$ in the expansion coefficients as well as the change $|\delta^{\pm}\chi_{\alpha}^{\mathbf{k}}\rangle$ in the MT-basis set, i.e.,

$$|\delta^{\pm}\mathbf{k}j\rangle=\sum_{\alpha}\{|\chi_{\alpha}^{\mathbf{k}+\mathbf{q}}\rangle\delta^{\pm}A_{\alpha}^{\mathbf{k}j}+|\delta^{\pm}\chi_{\alpha}^{\mathbf{k}}\rangle A_{\alpha}^{\mathbf{k}j}\}.\tag{32}$$

Since $|\delta^{\pm}\mathbf{k}j\rangle$ is a Bloch state with wave vector $\mathbf{k}\pm\mathbf{q}$, so are $|\chi_{\alpha}^{\mathbf{k}+\mathbf{q}}\rangle$ and $|\delta^{\pm}\chi_{\alpha}^{\mathbf{k}}\rangle$. The first function is the original linear MT orbital of wave vector $\mathbf{k}\pm\mathbf{q}$ and the second one is the change in the MT orbital due to the movements of atoms. In Sec. III we will give detailed formulas for the change in the basis functions. Here we note that since the original basis $|\chi_{\alpha}^{\mathbf{k}}\rangle$ is a Bloch sum of atom-centered localized orbitals, the important contribution to the change in the Bloch sum is connected with the *rigid* movement of these orbitals due to the rigid movement of the potentials for displaced atoms. The expansion (32) is rapidly convergent because the basis $|\delta^{\pm}\chi_{\alpha}^{\mathbf{k}}\rangle$ can be tailored to the perturbation just like the basis $|\chi_{\alpha}^{\mathbf{k}}\rangle$ is tailored to the unperturbed potential. Equation (32) can be interpreted as an expansion of $|\delta^{\pm}\mathbf{k}j\rangle$ in terms of $|\chi_{\alpha}^{\mathbf{k}+\mathbf{q}}\rangle$ in the local coordinate system displaced with the atom; the convergence with respect to the number of orbitals per atom must be about the same as for the unperturbed state. This is in contrast with the expression of the standard perturbation theory where, for the expansion of $|\delta^{\pm}\mathbf{k}j\rangle$, only the change in the coefficients $A_{\alpha}^{\mathbf{k}j}$ is taken into account [first contribution to (32)].

We shall now consider the second-order changes $|\delta^{\pm}\delta^{\mp}\mathbf{k}j\rangle$ which appear in expression (28) for the incomplete-basis-set corrections to the dynamical matrix. By performing the variation of the expansion (8) to second order we obtain

$$\begin{aligned}|\delta^{\pm}\delta^{\mp}\mathbf{k}j\rangle=\sum_{\alpha}\{&|\chi_{\alpha}^{\mathbf{k}}\rangle\delta^{\pm}\delta^{\mp}A_{\alpha}^{\mathbf{k}j}+|\delta^{\mp}\chi_{\alpha}^{\mathbf{k}\pm\mathbf{q}}\rangle\delta^{\pm}A_{\alpha}^{\mathbf{k}j}\\&+|\delta^{\pm}\chi_{\alpha}^{\mathbf{k}\mp\mathbf{q}}\rangle\delta^{\mp}A_{\alpha}^{\mathbf{k}j}+|\delta^{\pm}\delta^{\mp}\chi_{\alpha}^{\mathbf{k}}\rangle A_{\alpha}^{\mathbf{k}j}\},\end{aligned}\tag{33}$$

where $\delta^{\pm}\delta^{\mp}A_{\alpha}^{\mathbf{k}j}$ and $|\delta^{\pm}\delta^{\mp}\chi_{\alpha}^{\mathbf{k}}\rangle$ are the second-order changes in the expansion coefficients and the basis functions, respectively. By inserting (33) in the first term of (28) one sees that the second-order changes $\delta^{\pm}\delta^{\mp}A_{\alpha}^{\mathbf{k}j}$ do not contribute because they enter as coefficients to the unperturbed basis functions and

$$\sum_{\alpha}\delta^{\pm}\delta^{\mp}A_{\alpha}^{\mathbf{k}j}\langle\chi_{\alpha}^{\mathbf{k}}|H-\epsilon_{\mathbf{k}j}|\mathbf{k}j\rangle\equiv 0.\tag{34}$$

The absence of the coefficients $\delta^{\pm}\delta^{\mp}A_{\alpha}^{\mathbf{k}j}$ in our formulation of the problem has an important consequence: since $|\delta^{\pm}\delta^{\mp}\mathbf{k}j\rangle$ has only the unknown contribution from the first-order changes in $A_{\alpha}^{\mathbf{k}j}$ and since the Hilbert space $\{|\chi\rangle,|\delta\chi\rangle\}$ of the basis functions is fixed, we see that the variational freedom of the functional (28) is provided only by the coefficients $\delta^{\pm}A_{\alpha}^{\mathbf{k}j}$. This is again in close analogy to that in band-structure calculations the variational freedom of the total energy is provided only by the unperturbed coefficients $A_{\alpha}^{\mathbf{k}j}$. In the total-energy calculations this has the consequence when calculating the forces: due to the stationarity condition the force formula does not contain any first-order derivatives in $A_{\alpha}^{\mathbf{k}j}$. In the dynamical-matrix calculation this will have the same consequence when calculating third-order nonlinear coefficients: the corresponding formulas will not contain any second- and third-order derivatives of $A_{\alpha}^{\mathbf{k}j}$ and, thus, can be explicitly evaluated from only the knowledge of $\delta^{\pm}A_{\alpha}^{\mathbf{k}j}$. Note however, that together with the matrix elements containing $|\chi\rangle$, $|\delta\chi\rangle$, $|\delta^{(2)}\chi\rangle$, a contribution from third-order changes in the basis sets must be taken into account.

We shall now derive the equations for the first-order changes in the expansion coefficients. This is done by minimization of Λ with respect to $\delta^{\pm}A_{\alpha}^{\mathbf{k}j}$. We obtain

$$\begin{aligned}&\sum_{\alpha}\langle\chi_{\beta}^{\mathbf{k}\pm\mathbf{q}}|H-\epsilon_{\mathbf{k}j}|\chi_{\alpha}^{\mathbf{k}\pm\mathbf{q}}\rangle\delta^{\pm}A_{\alpha}^{\mathbf{k}j}+\sum_{\alpha}\{\langle\chi_{\beta}^{\mathbf{k}\pm\mathbf{q}}|\delta^{\pm}V_{\text{eff}}|\chi_{\alpha}^{\mathbf{k}}\rangle\\&\quad+\langle\delta^{\pm}\chi_{\beta}^{\mathbf{k}\mp\mathbf{q}}|H-\epsilon_{\mathbf{k}j}|\chi_{\alpha}^{\mathbf{k}}\rangle+\langle\chi_{\beta}^{\mathbf{k}\pm\mathbf{q}}|H-\epsilon_{\mathbf{k}j}|\delta^{\pm}\chi_{\alpha}^{\mathbf{k}}\rangle\}A_{\alpha}^{\mathbf{k}j}=0.\end{aligned}\tag{35}$$

This linear system of equations is, in fact, a variation of the original eigenvalue problem (9). It determines the position of the minimum of Λ in the space of the coefficients $\delta^{\pm}A_{\alpha}^{\mathbf{k}j}$, and none of the second-order changes, such as $|\delta^{\pm}\delta^{\mp}\chi_{\alpha}^{\mathbf{k}}\rangle$, affect it. The functions $|\delta^{\pm}\delta^{\mp}\chi_{\alpha}^{\mathbf{k}}\rangle$, on the other hand, define the value of Λ itself in its minimum and must be taken into account in the evaluation of the dynamical matrix.

We must now solve Eq. (35). This equation involves only the *occupied* states of the unperturbed system, which are necessary for constructing the induced charge density according to (5). It may be solved using an iterative algorithm with the number of operations proportional to

$N_{\text{band}} \times N^2_{\text{basis}}$, where N_{band} is a number of filled bands and N_{basis} is a number of the basis functions used for representing the unperturbed wave functions and their first-order corrections. This scheme is advantageous when using the LAPW or plane-wave pseudopotential methods where the conventional matrix-diagonalization algorithms represent the most time-consuming step which scales as the cube of the size of the basis. The LMTO method, on the other hand, has a small basis and the inversion of the matrix $\langle \chi_\beta^{\mathbf{k}\pm\mathbf{q}} | H - \epsilon_{\mathbf{k}j} | \chi_\alpha^{\mathbf{k}\pm\mathbf{q}} \rangle$ required for solving Eq. (35), can easily be performed because its eigenvalues are $\epsilon_{\mathbf{k}\pm\mathbf{q}j'} - \epsilon_{\mathbf{k}j}$ and eigenvectors are $A_\alpha^{\mathbf{k}\pm\mathbf{q}j'}$, where $j' = 1, N_{\text{basis}}$. Because of the minimal size of the basis in the LMTO method, it is not a time-consuming step to find eigenvalues $\epsilon_{\mathbf{k}j}$ and eigenvectors $A_\alpha^{\mathbf{k}j}$ for all energy bands $(= N_{\text{basis}})$ at some grid of wave vectors $\mathbf{k}$. This is independent of the phonon mode and therefore needs to be done only once. We therefore use the original eigenstates for the matrix inversion. The result for $\delta^\pm A_\alpha^{\mathbf{k}j}$ is then substituted into (32) which gives the final expression for $|\delta^\pm \mathbf{k}j\rangle$ in the form

$$
\begin{aligned}
|\delta^\pm \mathbf{k}j\rangle = {} & \sum_\alpha |\delta^\pm \chi_\alpha^{\mathbf{k}}\rangle A_\alpha^{\mathbf{k}j} + \sum_{j'} \frac{|\mathbf{k}\pm\mathbf{q}j'\rangle}{\epsilon_{\mathbf{k}j} - \epsilon_{\mathbf{k}\pm\mathbf{q}j'}} \\
& \times \left\{ \left\langle \mathbf{k}\pm\mathbf{q}j' | H - \epsilon_{\mathbf{k}j} \middle| \sum_\alpha \delta^\pm \chi_\alpha^{\mathbf{k}} A_\alpha^{\mathbf{k}j} \right\rangle \right. \\
& + \left\langle \sum_\alpha \delta^\pm \chi_\alpha^{\mathbf{k}\mp\mathbf{q}} A_\alpha^{\mathbf{k}\pm\mathbf{q}j'} \middle| H - \epsilon_{\mathbf{k}j} | \mathbf{k}j \right\rangle \\
& \left. + \langle \mathbf{k}\pm\mathbf{q}j' | \delta^\pm V_{\text{eff}} | \mathbf{k}j \rangle \right\}. \qquad (36)
\end{aligned}
$$

This formula has a simple physical meaning. The first three terms containing $|\delta\chi\rangle$ appear because of the use of *variational* solutions. They can be interpreted as incomplete-basis-set corrections to the last term (the one with $\delta^\pm V_{\text{eff}}$), which has the form of the standard perturbation theory (29). If all unperturbed states are exact and they represent *mathematically* a complete basis set, then the first and second terms in (36) cancel, the third term vanishes, and the standard perturbative formula (29) is recovered. However, if this is not the case, the use of the functions $|\delta\chi\rangle$ in the basis greatly reduces the number of states $|\mathbf{k}\pm\mathbf{q}j'\rangle$ needed to reach the convergence in (36). Namely, following the above derivation, the summation in the last three terms is over N_{basis} energy states, i.e., over the size of the basis for the unperturbed system. To illustrate the advantage of this formula we consider the so-called acoustic sum rule (ASR): suppose all atoms are displaced in the same direction by a small amount. The change in the charge density induced by the rigid movement ∇V_{eff} of the potential will be equal to $\nabla \rho$. Within the standard perturbation theory one obtains

$$
\begin{aligned}
|\delta \mathbf{k}j\rangle &= \sum_{j'} |\mathbf{k}j'\rangle \frac{\langle \mathbf{k}j' | \nabla V_{\text{eff}} | \mathbf{k}j\rangle}{\epsilon_{\mathbf{k}j} - \epsilon_{\mathbf{k}j'}} \\
&= \sum_{j'} |\mathbf{k}j'\rangle \langle \mathbf{k}j' | \nabla | \mathbf{k}j\rangle = \nabla |\mathbf{k}j\rangle. \qquad (37)
\end{aligned}
$$

The latter equality can only be obtained if the states $|\mathbf{k}j'\rangle$ represent a *mathematically complete* basis set. This is not the case in the LMTO method which employs a minimal basis set to reproduce the energy bands and wave functions within a certain energy window. On the other hand, within the minimal basis set the ASR can be trivially satisfied if one uses the expression (36) for the first-order corrections: here, by construction $|\delta\chi_\alpha^{\mathbf{k}}\rangle \equiv \nabla |\chi_\alpha^{\mathbf{k}}\rangle$, while the last three contributions vanish. [This is so because they are combined into the integral from a gradient of the periodic function: $\nabla\{\psi_{\mathbf{k}j'}(H - \epsilon_{\mathbf{k}j})\psi_{\mathbf{k}j}\}$, which is, by definition, equal to zero.]

The unoccupied states in the expression (36) should not be considered as real excitation energies and wave functions. Let us consider the induced charge density as a ground-state property of both perturbed and unperturbed systems. In both cases only the occupied states must be well reproduced, the excited states can, in principle, be arbitrary. The LMTO and LAPW methods are very suitable for this purpose: they are fast and accurate within a certain energy window, which is achieved by expanding the basis functions of the original Korringa-Kohn-Rostoker (KKR) and augmented-plane-wave (APW) methods by Taylor series around some energies ϵ_ν at the centers of interest. *The states* $|\mathbf{k}\pm\mathbf{q}j'\rangle$ *in* (36) *are the eigenstates of the Hamiltonian matrix* $\langle \chi_\beta^{\mathbf{k}\pm\mathbf{q}} | H | \chi_\alpha^{\mathbf{k}\pm\mathbf{q}} \rangle$ *which is itself constructed to reproduce the occupied energy bands well.* This is the energy window of interest and all centers of linearization ϵ_ν are in this window. In the KKR and APW methods the states $|\mathbf{k}\pm\mathbf{q}j'\rangle$ have the following meaning: since the KKR (APW) energy bands and eigenvectors are the eigenstates of the LMTO (LAPW) Hamiltonian $\langle \chi_\beta^{\mathbf{k}}(\epsilon_\nu) | H | \chi_\alpha^{\mathbf{k}}(\epsilon_\nu) \rangle$ with $\epsilon_\nu \equiv \epsilon_{\mathbf{k}j}$, the states $|\mathbf{k}\pm\mathbf{q}j'\rangle$ in (36) will be the eigenstates of the Hamiltonian $\langle \chi_\beta^{\mathbf{k}\pm\mathbf{q}}(\epsilon_{\mathbf{k}j}) | H | \chi_\alpha^{\mathbf{k}\pm\mathbf{q}}(\epsilon_{\mathbf{k}j}) \rangle$ and only those bands $\epsilon_{\mathbf{k}\pm\mathbf{q}j'}$ with energy near $\epsilon_{\mathbf{k}j}$ will be described correctly. In this case, finding $|\delta^\pm \mathbf{k}j\rangle$ requires the knowledge of this auxiliary spectrum for every occupied energy $\epsilon_{\mathbf{k}j}$. We thus finally conclude that the excited states are not to be interpreted as the exact ones, *only the knowledge of occupied energy bands is necessary* in our linear-response formulation.

III. IMPLEMENTATION

In this section, an extension of the linear muffin-tin orbital method for linear-response calculations is described. We shall first review the full-potential LMTO method, which is used as the framework in this implementation. Then, the problem of constructing the changes in the MT orbitals due to the atomic movements is considered. Other problems considered are the Brillouin-zone integration for metallic systems and the self-consistency at long wavelengths where the Coulomb singularity $4\pi/q^2$ makes the standard mixing schemes computationally inefficient.

A. Full-potential LMTO method

We first review the LMTO method, which solves the original Schrödinger equation. The space is partitioned into the nonoverlapping (or slightly overlapping) muffin-tin spheres s_R surrounding every atom and the remaining interstitial region Ω_{int}. Within the spheres, the basis functions are represented in terms of numerical solutions of the radial

Schrödinger equation for the spherical part of the potential multiplied by spherical harmonics as well as their energy derivatives taken at some set of energies ϵ_ν at the centers of interest. In the interstitial region, where the potential is essentially flat, the basis functions are spherical waves taken as the solutions of Helmholtz's equation: $(-\nabla^2-\epsilon)f(r,\epsilon)=0$ with some fixed value of the average kinetic energy $\epsilon=\kappa_\nu^2$. In particular, in the standard LMTO method using the atomic-sphere approximation (ASA),[27] the approximation $\kappa_\nu^2=0$ is chosen. In the extensions of the LMTO method for a potential of arbitrary shape (full potential), a multiple-κ basis set[32] is normally used in order to increase the variational freedom of the basis functions while recent developments of a different LMTO technique[33] promise to avoid this problem.

The general strategy for including the full-potential terms in the calculation is the use of the variational principle. A few different techniques have been developed for taking the nonspherical corrections into account in the framework of the LMTO method. They include Fourier transforms of the LMTO's in the interstitial region,[34,35] one-center spherical-harmonics expansions within atomic cells,[29] interpolations in terms of the Hankel functions[36] as well as direct calculations of the charge density in the tight-binding representation.[37] In two of these schemes[29,36] the treatment of open structures such as, e.g., the diamond structure is complicated and interstitial spheres are usually placed between the atomic spheres. In the dynamical-matrix calculation it is inconvenient to use interstitial spheres because they lead to artificial degrees of freedom for the lattice dynamics. Therefore we will develop the linear-response LMTO technique using the plane-wave Fourier representation. This allows us to apply the method for such materials as Si and NbC without interstitial spheres. Note, however, that in our previous applications[12,15,16] for bcc and fcc metals, atomic-cell spherical-harmonic expansions[29] were used.

Consider the so-called envelope function, which is a singular Hankel function,

$$K_{\kappa RL}(\mathbf{r}_R-\mathbf{t})=K_{\kappa Rl}(|\mathbf{r}_R-\mathbf{t}|)i^lY_{lm}(\mathbf{r}_R-\mathbf{t}), \tag{38}$$

centered at site $\mathbf{R}+\mathbf{t}$ and with an energy $\epsilon=\kappa_\nu^2$. Y_{lm} denotes a complex spherical harmonic with the phase convention after Ref. 38. Inside any other site $\mathbf{R}'+\mathbf{t}'$ the Hankel function can be represented as an expansion in terms of the Bessel functions, $J_{\kappa R'L'}(\mathbf{r}_{R'}-\mathbf{t}')$, i.e.,

$$K_{\kappa RL}(\mathbf{r}_R-\mathbf{t})=-\sum_{L'}J_{\kappa R'L'}(\mathbf{r}_{R'}-\mathbf{t}')\gamma_{R'l'}S_{R'L'RL}(\mathbf{t}'-\mathbf{t},\kappa), \tag{39}$$

where $\gamma_{Rl}=1/s_R(2l+1)$ and $S_{R'L'RL}(\mathbf{t},\kappa)$ are the structure constants in real space. Note that, while the index L enumerating the basis functions usually runs only over s, p, and d states, the sum over L' in this expression must include higher angular momenta. Normally l' goes up to $6-8$. For convenience, we use the following prefactors in the definitions of the spherical functions:

$$K_{\kappa Rl}(|\mathbf{r}_R|)=-\frac{(\kappa s_R)^{l+1}}{(2l-1)!!}h_l(\kappa|\mathbf{r}_R|), \tag{40}$$

$$J_{\kappa Rl}(|\mathbf{r}_R|)=\frac{(2l+1)!!}{(\kappa s_R)^l}j_l(\kappa|\mathbf{r}_R|), \tag{41}$$

where $h_l=j_l-in_l$, j_l, and n_l are the spherical Hankel, Bessel, and Neumann functions, respectively. The expression for the structure constants is then

$$\begin{aligned}S_{R'L'RL}(\mathbf{t},\kappa)={}&\left(\frac{s_{R'}}{w}\right)^{l'+1}\left(\frac{s_R}{w}\right)^{l+1}\\&\times\sum_{L''}\frac{-4\pi w(2l''-1)!!}{(2l'-1)!!(2l-1)!!}C_{LL'}^{L''}(\kappa w)^{l+l'-l''}\\&\times K_{\kappa wl''}(|\mathbf{t}-\mathbf{R}'+\mathbf{R}|)(-i)^{l''}Y_{L''}^*(\mathbf{t}-\mathbf{R}'+\mathbf{R}),\end{aligned} \tag{42}$$

where w is the average Wigner-Seitz radius and the Hankel function $K_{\kappa wl}$ is defined with w instead of s_R in expression (40). The Gaunt coefficients $C_{LL'}^{L''}$ are defined by the integral

$$C_{LL'}^{L''}=\int Y_LY_{L''}Y_{L'}^*. \tag{43}$$

We now consider a Bloch sum of the Hankel functions (38), centered at different sites, which, inside the MT sphere at $\mathbf{R}'$, is represented by the expansions in the Bessel functions:

$$\begin{aligned}\sum_t e^{i\mathbf{kt}}K_{\kappa RL}(\mathbf{r}_R-\mathbf{t})={}&K_{\kappa RL}(\mathbf{r}_R)\delta_{R'R}\\&-\sum_{L'}J_{\kappa R'L'}(\mathbf{r}_{R'})\gamma_{R'l'}S_{R'L'RL}^{\mathbf{k}}(\kappa),\end{aligned} \tag{44}$$

where $S_{R'L'RL}^{\mathbf{k}}(\kappa)$ denotes the lattice sum of the structure constants (42). The linear MT orbitals $|\chi_{\kappa RL}^{\mathbf{k}}\rangle$ are now obtained by augmenting the spherical functions $K_{\kappa RL}$, $J_{\kappa RL}$ in all MT spheres by numerical radial functions Φ_{kRL}^K, Φ_{kRL}^J:

$$\begin{aligned}\chi_{\kappa RL}^{\mathbf{k}}(\mathbf{r}_{R'})={}&\Phi_{\kappa RL}^K(\mathbf{r}_R)\delta_{R'R}\\&-\sum_{L'}\Phi_{\kappa R'L'}^J(\mathbf{r}_{R'})\gamma_{R'l'}S_{R'L'RL}^{\mathbf{k}}(\kappa).\end{aligned} \tag{45}$$

The functions Φ_{kRL}^K, Φ_{kRL}^J are the linear combinations of the solutions $\phi_{RL}(\mathbf{r}_R,\epsilon_{\nu\kappa Rl})\equiv\phi_{\kappa RL}$ to the radial Schrödinger equation as well as their energy derivatives $\dot{\phi}_{RL}(\mathbf{r}_R,\epsilon_{\nu\kappa Rl})\equiv\dot{\phi}_{\kappa RL}$ taken at the energies $\epsilon_{\nu\kappa Rl}$. In the interstitial region, the linear MT orbitals are represented as multicenter expansions [left-hand side of Eq. (44)]. In order to calculate the interstitial-potential matrix elements and represent the charge density, we use the Fourier transform of the LMTO's in the interstitial region. It is impossible to consider the Fourier transform of the expression (44) directly because of the singularities in the Hankel functions. On the other hand, since this representation will be used for the description of the basis functions only within Ω_{int}, we can substitute the divergent part of the Hankel function by a smooth function for $r_R<s_R$. This regular function is defined in the

Appendix and it is denoted $\widetilde{K}_{\kappa RL}$. We thus introduce a pseudo-LMTO $|\widetilde{\chi}^{\mathbf{k}}_{\kappa RL}\rangle$ defined in all space as follows:

$$\widetilde{\chi}^{\mathbf{k}}_{\kappa RL}(\mathbf{r})=\sum_{t} e^{i\mathbf{kt}}\widetilde{K}_{\kappa L}(\mathbf{r}_R-\mathbf{t})=\sum_{G} \widetilde{\chi}_{\kappa RL}(\mathbf{k}+\mathbf{G})e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}, \tag{46}$$

which is identical with the true sum (44) in the interstitial region.

The charge density and the potential have a dual representation: spherical-harmonic expansions inside the MT spheres and plane-wave expansions in the interstitial region. This is usually done by introducing a smooth pseudocharge density $\widetilde{\rho}$ in all space defined in terms of the pseudo-LMTO's (46). The pseudodensity coincides with the true density when $\mathbf{r}\in\Omega_{\text{int}}$. In this way, the solution of the Poisson equation is straightforward and can be done along the lines developed in Ref. 39. In practical applications we have also used the technique described in the Appendix for the Fourier transform of the Coulomb interactions and for the construction of auxiliary densities. The exchange-correlation potential is found using the fast-Fourier transform and the interstitial-potential matrix elements are explicitly evaluated.

B. Changes in the linear muffin-tin orbitals

We shall now discuss the linear-response calculation. Small displacements of atoms from their equilibrium positions defined by expression (1) lead to the change in the Bloch sum of the atom-centered (pseudo) Hankel functions (46). Because of the explicit dependence of the basis functions $|\chi^{\mathbf{k}}_{\kappa RL}\rangle$ on the atomic positions $\mathbf{R}$, here and in the following the displaced atoms will be denoted by the index $\overline{\mathbf{R}}$. The change in the Bloch sum of the MT orbitals can be found analogously to the change in the external potential in Eqs. (2)–(4). As a result, we consider two traveling waves with wave vectors $\mathbf{k}+\mathbf{q}$ and $\mathbf{k}-\mathbf{q}$, i.e.,

$$\begin{aligned}\frac{\delta^{\pm}\widetilde{\chi}^{\mathbf{k}}_{\kappa RL}(\mathbf{r})}{\delta \overline{R}_{\mu}}&=-\delta_{R\overline{R}}\sum_{t} e^{i(\mathbf{k}\pm\mathbf{q})\mathbf{t}}\nabla_{\mu}\widetilde{K}_{\kappa RL}(\mathbf{r}_R-\mathbf{t})\\&=-\delta_{R\overline{R}}\sum_{G} i(k\pm q+G)_{\mu}\\&\quad\times\widetilde{\chi}_{\kappa RL}(\mathbf{k}\pm\mathbf{q}+\mathbf{G})e^{i(\mathbf{k}\pm\mathbf{q}+\mathbf{G})\mathbf{r}},\end{aligned} \tag{47}$$

which represent the change of the basis functions in the interstitial region or the change of the pseudo-LMTO's in the whole space. Here we have restored the original notations: $\delta^{\pm}\widetilde{\chi}^{\mathbf{k}}_{\kappa RL}(\mathbf{r})/\delta \overline{R}_{\mu}\equiv\delta^{\pm}\widetilde{\chi}^{\mathbf{k}}_{\kappa RL}(\mathbf{r})$. We also introduce a spherical coordinate system,[38]

$$\overline{\mathbf{R}}=\sum_{\mu}\overline{R}_{\mu}e^{\mu}=\overline{R}_{-1}e^{-1}+\overline{R}_{0}e^{0}+\overline{R}_{+1}e^{+1}, \tag{48}$$

which is connected to the Cartesian system as follows: $R_{-1}=+(R_x-iR_y)/\sqrt{2}$, $R_0=R_z, R_{+1}=-(R_x+iR_y)/\sqrt{2}$. The reason is that, in the spherical coordinates, the operation ∇_{μ} on a product of a radial function $f(r)$ multiplied by the spherical harmonic takes the simple form

$$\begin{aligned}&\nabla_{\mu}f(r)Y_{lm}(\mathbf{r})\\&\quad=\sqrt{\frac{4\pi}{3}}C^{1\mu}_{lml+1m+\mu}\left(\frac{df}{dr}-\frac{l}{r}f(r)\right)Y_{l+1m+\mu}(\mathbf{r})\\&\qquad+\sqrt{\frac{4\pi}{3}}C^{1\mu}_{lml-1m+\mu}\left(\frac{df}{dr}+\frac{l+1}{r}f(r)\right)Y_{l-1m+\mu}(\mathbf{r}).\end{aligned} \tag{49}$$

We shall now find a variation of the basis functions inside the MT spheres. In the sphere $\mathbf{R}'$, the original LMTO is defined in the expression (45). Its change must include both the changes in the numerical radial functions and the change in the structure constants:

$$\begin{aligned}\frac{\delta^{\pm}\chi^{\mathbf{k}}_{\kappa RL}(\mathbf{r}_{R'})}{\delta\overline{R}_{\mu}}&=\frac{\delta^{\pm}\Phi^{K}_{\kappa RL}(\mathbf{r}_R)}{\delta\overline{R}_{\mu}}\delta_{R'R}\\&\quad-\sum_{L'}\frac{\delta^{\pm}\Phi^{J}_{\kappa R'L'}(\mathbf{r}_{R'})}{\delta\overline{R}_{\mu}}\gamma_{R'l'}S^{\mathbf{k}}_{R'L'RL}(\kappa)\\&\quad-\sum_{L'}\Phi^{J}_{\kappa R'L'}(\mathbf{r}_{R'})\gamma_{R'l'}\frac{\delta^{\pm}S^{\mathbf{k}}_{R'L'RL}(\kappa)}{\delta\overline{R}_{\mu}}.\end{aligned} \tag{50}$$

The change in the numerical functions contains two contributions. Since $\Phi^{K}_{\kappa RL}$, $\Phi^{J}_{\kappa RL}$ are constructed from the solutions of the radial Schrödinger equation and their energy derivatives, $\phi_{\kappa RL}$ and $\dot{\phi}_{\kappa RL}$, the change in $\phi_{\kappa RL}$ and $\dot{\phi}_{\kappa RL}$ is a result of both the rigid movement of the spherical component of the potential and the change in the shape of the spherical component. In the following, it is convenient to treat the rigid movements of the potential within the MT sphere centered at $\mathbf{R}$ separately, i.e., represent the total change in the form:

$$\frac{\delta^{\pm}V_{\text{eff}}(\mathbf{r}_R)}{\delta\overline{R}_{\mu}}=-\delta_{R\overline{R}}\nabla_{\mu}V_{\text{eff}}(\mathbf{r}_R)+\frac{\delta^{\pm}_{(s)}V_{\text{eff}}(\mathbf{r}_R)}{\delta\overline{R}_{\mu}}, \tag{51}$$

where the notation $\delta^{\pm}_{(s)}$ stands for the "soft" change, i.e., the variation connected with the change in the shape of the function. The functions $\delta^{\pm}\phi_{\kappa RL}/\delta\overline{R}_{\mu}$ are represented in a form similar to (51), i.e., $\delta^{\pm}\phi_{\kappa RL}/\delta\overline{R}_{\mu}=-\delta_{R\overline{R}}\nabla_{\mu}\phi_{\kappa RL}+\delta^{\pm}_{(s)}\phi_{\kappa RL}/\delta\overline{R}_{\mu}$, where the last (soft) contribution is found by solving the radial Sternheimer equation:

$$\begin{aligned}&\left(-\nabla^{2}_{r}+\frac{l(l+1)}{r^{2}}+V^{\text{SPH}}_{\text{eff}}-\epsilon_{\nu\kappa Rl}\right)\frac{\delta^{\pm}_{(s)}\phi_{\kappa RL}}{\delta\overline{R}_{\mu}}\\&\quad+\left(\frac{\delta^{\pm}_{(s)}V^{\text{SPH}}_{\text{eff}}}{\delta\overline{R}_{\mu}}-\frac{\delta^{\pm}_{(s)}\epsilon_{\nu\kappa Rl}}{\delta\overline{R}_{\mu}}\right)\phi_{\kappa RL}=0.\end{aligned} \tag{52}$$

The superscript "*SPH*" here denotes the spherical component of the potential and the perturbation. It is, in principle, not a problem to take all nonspherical terms of the perturbation into account. If this is done, the first-order changes in the radial functions are no longer given by a single spherical harmonic but as an expansion in Y_{lm}. One obtains an *uncoupled* system of radial equations, which can easily be solved.[11] However, in the problem of lattice dynamics the change in $\phi_{\kappa RL}$ and $\dot{\phi}_{\kappa RL}$ due to the change in the shape of the spherical component of the potential is small. This is because the motions of atoms mainly distort the dipole part of the potential. If the change in the shape of the spherical component can be described by some constant shift of the energy, it may be canceled by appropriate choice of the change $\delta^{\pm}_{(s)}\epsilon_{\nu\kappa Rl}/\delta R_{\mu}$ in the energies $\epsilon_{\nu\kappa Rl}$. This cancelation can, for instance, be obtained by finding $\delta^{\pm}_{(s)}\epsilon_{\nu\kappa Rl}/\delta R_{\mu}$ with fixed logarithmic derivatives $D_{\nu\kappa Rl}$. (The derivatives $D_{\nu\kappa Rl}$ are evaluated at the occupied centers of gravities of the bands for the unperturbed crystal.) We thus see that the influence of the constant shifts to the change in the basis set can be eliminated and, therefore, one can neglect the contribution $\delta^{\pm}_{(s)}\phi_{\kappa RL}/\delta R_{\mu}$ in practical calculations. The accuracy of this approximation is quite good which has already been confirmed by good agreement between total energy and force calculations with the original LMTO method[29] where the same approximation was used for deriving the force formula.

We now give the formula for the change in the structure constants which enters Eq. (50). It is expressed as the difference between the gradients of the structure constants for wave vectors $\mathbf{k}$ and $\mathbf{k}\pm\mathbf{q}$, i.e.,

$$\frac{\delta^{\pm}S^{\mathbf{k}}_{R'L'RL}(\kappa)}{\delta R_{\mu}}=\delta_{RR'}\nabla_{\mu}S^{\mathbf{k}}_{R'L'RL}(\kappa)-\delta_{RR}\nabla_{\mu}S^{\mathbf{k}\pm\mathbf{q}}_{R'L'RL}(\kappa). \tag{53}$$

The gradient is with respect to $\mathbf{R}'-\mathbf{R}$. From (49), using the recursion relations for the Hankel functions, it follows that the change in the structure constants can be expressed in terms of the structure constants:

$$\begin{aligned}\nabla_{\mu}S^{\mathbf{k}}_{R'l'm'Rlm}(\kappa)\\ &=i\sqrt{\frac{4\pi}{3}}C^{1\mu}_{l'-1m'-\mu l'm'}\frac{\kappa^{2}s_{R'}}{(2l'-1)}S^{\mathbf{k}}_{R'l'-1m'-\mu Rlm}(\kappa)\\ &\quad+i\sqrt{\frac{4\pi}{3}}C^{1\mu}_{l'+1m'-\mu l'm'}\frac{(2l'+1)}{s_{R'}}S^{\mathbf{k}}_{R'l'+1m'-\mu Rlm}(\kappa).\end{aligned} \tag{54}$$

Here, the left index of the structure constants has changed to $l'\pm 1,m'-\mu$, but the right index lm remains the same. An analogous formula exist in which the right index change is $l\pm 1,m+\mu$ in and the left index is unchanged.

It is seen that the change in the MT orbital (50) can be represented as a rigid part, a small soft part and a contribution from the change in the structure constants:

$$\begin{aligned}\frac{\delta^{\pm}\chi^{\mathbf{k}}_{\kappa RL}(\mathbf{r}_{R'})}{\delta R_{\mu}}&=-\delta_{RR'}\nabla_{\mu}\chi^{\mathbf{k}}_{\kappa RL}(\mathbf{r}_{R'})+\frac{\delta^{\pm}_{(s)}\chi^{\mathbf{k}}_{\kappa RL}(\mathbf{r}_{R'})}{\delta R_{\mu}}\\ &\quad-\sum_{L'}\Phi^{J}_{\kappa R'L'}(\mathbf{r}_{R'})\gamma_{R'l'}\frac{\delta^{\pm}S^{\mathbf{k}}_{R'L'RL}(\kappa)}{\delta R_{\mu}}.\end{aligned} \tag{55}$$

It is convenient to separate the rigid part since it gives rise to a rigid contribution in the electronic response:

$$\frac{\delta^{\pm}\rho(\mathbf{r}_{R})}{\delta R_{\mu}}=-\delta_{RR}\nabla_{\mu}\rho(\mathbf{r}_{R})+\frac{\delta^{\pm}_{(s)}\rho(\mathbf{r}_{R})}{\delta R_{\mu}}. \tag{56}$$

Since the induced charge density (56) has the same form as the change in the potential (51), we need not calculate the gradients of the charge density and the potential. This is important since these gradients are huge in the core region, which could result in large numerical errors. The second term in (55) is $\delta^{\pm}_{(s)}\chi^{\mathbf{k}}_{\kappa RL}(\mathbf{r}_{R'})/\delta R_{\mu}$. It is constructed from the changes $\delta^{\pm}_{(s)}\phi_{\kappa RL}/\delta R_{\mu}$ and their energy derivatives which are numerically small. This function is exactly equal to zero together with its first-order radial derivative at the sphere $s_{R'}$. It translates like a Bloch wave with vector $\mathbf{k}\pm\mathbf{q}$ because the original form of one-center expansion (45) translates with wave vector $\mathbf{k}$ while the first-order changes $\delta^{\pm}_{(s)}\phi_{\kappa RL}/\delta R_{\mu}$ translate, like the perturbation, with wave vector $\pm\mathbf{q}$. The whole expression (55) also translates with wave vector $\mathbf{k}\pm\mathbf{q}$ and fits into the multicenter expansion of the change in the basis set in the interstitial region [formula (47)]. However, since the original LMTO's are continuous and only differentiable to the first order at the boundaries of the MT spheres, the matching of the change in the basis set is only continuous but not differentiable. This, in principle, leads to a kink in the change of the charge density. However, it does not have any effect in the calculation of the dynamical matrix if the latter is compared with the second-order derivative of the total energy derived from the frozen-phonon supercell calculation. This is so because the extension of the LMTO method described here is just an analytical version of the finite-difference approach employed in the supercell technique. When applied to the same problem, the results of both approaches have to be the same except for the errors introduced by taking finite differences. This concerns the comparison of not only the dynamical matrix and the phonon frequencies, but also the changes in the basis set, the expansion coefficients, the charge densities, and in all other quantities which can be obtained by the frozen-phonon LMTO technique.

We now turn to the problem of calculating the change in the expansion coefficients $A^{\mathbf{k}j}_{\kappa RL}$, which are necessary to compute the first-order corrections according to (32). From expression (36), the change $\delta^{\pm}A^{\mathbf{k}j}_{\kappa RL}/\delta R_{\mu}$ is given by

$$\begin{aligned}\frac{\delta^{\pm}A^{\mathbf{k}j}_{\kappa RL}}{\delta R_{\mu}}&=\sum_{j'}\frac{A^{\mathbf{k}\pm\mathbf{q}j'}_{\kappa RL}}{\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}\pm\mathbf{q}j'}}\\ &\quad\times\left(\frac{\delta^{\pm}H^{\mathbf{k}\pm\mathbf{q}j'\mathbf{k}j}}{\delta R_{\mu}}-\epsilon_{\mathbf{k}j}\frac{\delta^{\pm}O^{\mathbf{k}\pm\mathbf{q}j'\mathbf{k}j}}{\delta R_{\mu}}\right)\end{aligned} \tag{57}$$

and it is expressed in terms of the change in the Hamiltonian and the overlap matrices. Here, the change in the matrix elements of the Hamiltonian is given by the band representation:

$$\frac{\delta^{\pm} H^{\mathbf{k}\pm\mathbf{q}j'\mathbf{k}j}}{\delta \overline{R}_{\mu}} = \sum_{\kappa' R' L'} \sum_{\kappa R L} A^{\mathbf{k}\pm\mathbf{q}j'*}_{\kappa' R' L'} \frac{\delta^{\pm} H_{\kappa' R' L' \kappa R L}}{\delta \overline{R}_{\mu}} A^{\mathbf{k}j}_{\kappa R L} \tag{58}$$

and a similar formula holds for the matrix elements of the overlap integral. In the original, $\{\kappa RL\}$, representation the changes $\delta^{\pm} H_{\kappa' R' L' \kappa RL}/\delta \overline{R}_{\mu}$ and $\delta^{\pm} O_{\kappa' R' L' \kappa RL}/\delta \overline{R}_{\mu}$ are readily computed using the formulas (47) and (55) for the first-order changes in the basis set. It is indeed even more advantageous to find the corresponding formulas by directly varying the expressions for the Hamiltonian and the overlap matrices. This avoids the problem of combining the contributions with the gradients of numerical radial functions to the surface integrals. One point about calculating the change in the interstitial kinetic-energy matrix elements and the interstitial overlap integrals is worth noticing. Since these matrix elements contain energy derivative of the structure constants, the change in these matrix elements will contain the change in this energy derivative. The corresponding formula can be found by taking the derivative with respect to κ^2 in the expressions (53) and (54).

Another problem is to find second-order changes in the LMTO basis functions as well as second-order variations in the Hamiltonian and the overlap matrices. They are necessary for computing the incomplete-basis-set corrections in (28) for the dynamical matrix. In the interstitial region the second-order change in the pseudo-LMTO's is simply given by

$$\frac{\delta^{+}\delta^{-}\chi^{\mathbf{k}}_{\kappa RL}(\mathbf{r})}{\delta \overline{R}'_{\mu'}\delta \overline{R}_{\mu}} = \frac{1}{2}\delta_{\overline{R'}R}\delta_{\overline{R}R}\sum_{G} i(k+G)_{\mu'} i(k+G)_{\mu} \times \chi_{\kappa RL}(\mathbf{k}+\mathbf{G}) e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}}. \tag{59}$$

Inside the MT spheres the expression is more complicated, but can be found straightforwardly by performing one more variation $\delta^{-}/\delta R_{\mu'}$ of expression (50) for the first-order change. It will contain second-order changes in the numerical radial functions and second-order changes in the structure constants as well as different products of the first-order changes in these quantities. The second-order changes in the structure constants are given by

$$\frac{\delta^{+}\delta^{-} S^{\mathbf{k}}_{R'L'RL}(\kappa)}{\delta \overline{R}'_{\mu'}\delta \overline{R}_{\mu}} = \frac{1}{2}\left\{ \delta_{\overline{R}R'} \nabla_{\mu} \frac{\delta^{-} S^{\mathbf{k}}_{R'L'RL}(\kappa)}{\delta \overline{R}'_{\mu'}} - \delta_{\overline{R}R} \nabla_{\mu} \frac{\delta^{-} S^{\mathbf{k}+\mathbf{q}}_{R'L'RL}(\kappa)}{\delta \overline{R}'_{\mu'}} \right\}. \tag{60}$$

This is obtained from the expression (53) and $\delta^{-}S^{\mathbf{k}}$ is expressed via the difference between the gradients of the structure constants for the wave vectors $\mathbf{k}$ and $\mathbf{k}-\mathbf{q}$, while $\delta^{-}S^{\mathbf{k}+\mathbf{q}}$ is the difference between ∇S for the wave vectors $\mathbf{k}+\mathbf{q}$ and $\mathbf{k}$. Alternatively, the expression (60) can be found by first considering the expression for the structure constants in the supercell and then, assuming the form (1) for the atomic displacements, transferring the supercell expression to the original structure. The second-order gradients $\nabla_{\mu'}\nabla_{\mu} S$ are calculated using (54) and they are again the structure constants with the left index changed to $l' \pm 2, m' - \mu' - \mu$ and the unchanged right index. Analogously, they can be expressed in terms of the structure constants of the same left index $l'm'$ and the right index: $l \pm 2, m + \mu' + \mu$.

The second-order changes in the numerical radial functions must also be calculated. They contain contributions $\nabla_{\mu'}\nabla_{\mu}\Phi^{K,J}_{\kappa RL}$ due to the rigid movement of the spherical part of the potential to second order, changes due to the rigid movements of the first-order variations in the shape of the spherical part (rigid movement of the soft part), as well as the contributions arising from the change in the shape of the spherical part to the second order (second-order soft part). As we discussed above, one can neglect by the influence of the change in the shape of $V^{\mathrm{SPH}}_{\mathrm{eff}}$ to the change in the basis. Therefore, we must only keep the rigid contributions described by $\nabla_{\mu'}\nabla_{\mu}\Phi^{K,J}_{\kappa RL}$.

C. Brillouin-zone integrals

After computing the first-order corrections to the wave functions, we have to perform the $\mathbf{k}$-space integration over the first Brillouin zone (BZ) in order to find the change in the charge density from Eq. (6). The BZ integration is also required for calculating the incomplete-basis-set corrections to the dynamical matrix. It is in general a full-zone integration while for the high-symmetry wave vectors the integrals are reduced to that portion of the BZ which is irreducible with respect to the symmetry of the perturbation vector.

Two kinds of the integrals have to be performed in the linear-response calculation. The first one has the following form:

$$I_1(\mathbf{q}) = \sum_{\mathbf{k}j} 2 f_{\mathbf{k}j} A_{kj}(\mathbf{q}), \tag{61}$$

and the second one is given by

$$I_2(\mathbf{q}) = \sum_{\mathbf{k}jj'} \frac{2 f_{\mathbf{k}j}(1 - f_{\mathbf{k}\pm\mathbf{q}j'})}{\epsilon_{\mathbf{k}j} - \epsilon_{\mathbf{k}\pm\mathbf{q}j'}} M^{\mathbf{k}\pm\mathbf{q}j'\mathbf{k}j}, \tag{62}$$

where $A_{kj}(\mathbf{q})$ and $M^{\mathbf{k}\pm\mathbf{q}j'\mathbf{k}j}$ are the matrix elements which presumably are smooth functions of wave vectors. In order to calculate these integrals we use the tetrahedron method in Ref. 40. In this method, the BZ is set up by the reciprocal-lattice translational vectors and it is divided into small primitive cells exactly as in standard fast-Fourier-transform analysis. The calculation becomes simpler if the $\mathbf{q}$ vector coincides with a mesh point because $\mathbf{k}\pm\mathbf{q}$ vectors are also mesh points. In this way the energy bands, the expansion coefficients, and the structure constants have to be calculated only once at the mesh of the irreducible wave vectors $\mathbf{k}$ for the unperturbed crystal. Applying symmetry operations, these quantities can be found for any general $\mathbf{k}$.

When applied to a semiconductor, the tetrahedron method is identical to the special-point method of Monkhorst and Pack,[41] which means that the occupation numbers $f_{\mathbf{k}j}$ in (61) and (62) can be regarded as the geometrical weights of the **k** points. Both integrals (61) and (62) converge rapidly with respect to the number of **k** points. The integral $I_2(\mathbf{q})$ reduces to the integral $I_1(\mathbf{q})$ by performing the summation over the unoccupied bands j'.

For metallic systems a significantly larger number of **k** points (N_k) is necessary when the matrix elements as well as the energy denominator $\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}\pm\mathbf{q}j'}$ are interpolated linearly within the tetrahedron. For these systems there are two sources of errors: the first is connected with the interpolation of the matrix elements and the second is connected with the interpolation of the Fermi surface. The latter can easily be circumvented in the linear-response calculation, since the Fermi surface can be determined accurately from the band structure of the unperturbed crystal. For the integrals $I_1(\mathbf{q})$ this can be done using the method described in Ref. 42 which is based on considering two, coarse and dense, meshes. In the tetrahedron method the integration weight of a particular **k** point is calculated by integrating over the occupied parts of those tetrahedra that contain this point. The occupied part of the tetrahedron is found by linear interpolation between the energies at the corners of this tetrahedron. Suppose we introduce a much denser mesh that also contains the original coarse mesh. We will need only the energies $\epsilon_{\mathbf{k}j}$ at this dense mesh, which will define the accurate Fermi surface (for example, also by linear interpolation). Then, the occupied part of the tetrahedron at the coarse grid can be found by not interpolating linearly the energies known at its corners but as a piece of the accurate Fermi surface found with help of the dense grid. The same is applicable to the integrals $I_2(\mathbf{q})$: we consider the dense and the coarse grids. The band energies are known at the dense grid. To find the integration weights we must find a region in the tetrahedron where the state $|\mathbf{k}j\rangle$ is occupied and the state $|\mathbf{k}\pm\mathbf{q}j\rangle$ is unoccupied. This can be done using the linear interpolation but on the dense grid. We must also include the energy denominator $\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}\pm\mathbf{q}j'}$. This is also interpolated linearly but again using the dense grid. Consequently, all the effects from the energy bands and the Fermi surface are treated exactly in such a scheme which allows us to avoid this source of errors in the integration.

Another source of errors is connected with the linear interpolation of the matrix elements. We have already mentioned that the matrix elements are normally smooth functions of wave vectors and one can expect that after eliminating the errors connected with the approximate treatment of the Fermi surface, the number of **k** points need not be too large. However, in practical calculations a large cancellation occurs between the two kinds of the integrals, (61) and (62). If one uses different integration weights, it will lead to a large numerical errors connected with the different convergency with respect to N_k in these integrals. Our task is thus to extract a large contribution from the integral of the type $I_2(\mathbf{q})$ and reduce it to the form $I_1(\mathbf{q})$. This is achieved by rewriting the energy denominator $\Delta=\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}\pm\mathbf{q}j'}$ in the expression (62) as follows:

$$\frac{1}{\Delta}=\frac{\Delta}{\Delta^2+\delta^2}+\frac{1}{\Delta}\left(1-\frac{\Delta^2}{\Delta^2+\delta^2}\right), \tag{63}$$

where the broadening δ is usually chosen ~0.1 Ry. Then, the sum over unoccupied bands j' in the integral containing $\Delta/(\Delta^2+\delta^2)$ is readily performed because this expression remains regular when $\Delta\rightarrow 0$. Consequently, this integral is reduced to the integral of the type $I_1(\mathbf{q})$. The second integral in (63) contains $1/\Delta$ and must be treated as the integral of the type $I_2(\mathbf{q})$ where the original matrix element $M^{\mathbf{k}\pm\mathbf{q}j'\mathbf{k}j}$ is now multiplied by the expression in brackets in (63). However, because the latter rapidly goes to zero for $\Delta\gg\delta$, the whole integral remains small and it is nonzero only for the band transitions $j\rightarrow j'$ between the states near the Fermi level. In practical calculations of the dynamical matrix, this procedure allows us to avoid the errors connected with the large cancellations.

We finally mention that a simple correction formula which significantly improves the convergency of the integrals $I_1(\mathbf{q})$ by taking into account the curvature of the matrix elements beyond the linear interpolation was derived by Blöchl.[37,42] Unfortunately, it is hard to derive such a correction for the integrals $I_2(\mathbf{q})$ because of the appearance of the energy denominator but we always use the Blöchl correction for the integrals (61).

D. Self-consistency at long wavelengths

The change in the charge density (5) induced by the displacements of nuclei screens the external perturbation (4), and the linear-response equations (6)–(8) must, therefore, be solved self-consistently. Let us assume that we have found the response of the electrons, $\delta\rho^{(0)}$, to the external perturbation δV_{ext} or the perturbation screened by some guessed $\delta\rho^{\text{guess}}$ (here we omit ''±'' for simplicity). The latter could, for instance, be the rigid shifts of the charge density around the displaced nuclei and in practical calculations the external perturbation is always considered as the change in the bare Coulomb potential (4) plus the term $\nabla\rho$ within the MT sphere. The response $\delta\rho^{(0)}$ is found along the lines described above and, consequently, it can be considered as some polarization operator $\hat{\Pi}$ that acts on δV_{ext}, i.e.,

$$\delta\rho^{(0)}=\hat{\Pi}\delta V_{\text{ext}}. \tag{64}$$

If we omit the terms containing the change in the basis functions and forget about the completeness problem of the unperturbed states, the operator $\hat{\Pi}$ is given by the independent-particle polarizability function $\hat{\pi}$:

$$\pi_{\mathbf{q}}(\mathbf{r},\mathbf{r}')=\sum_{\mathbf{k}jj'}\frac{f_{\mathbf{k}j}-f_{\mathbf{k}+\mathbf{q}j'}}{\epsilon_{\mathbf{k}j}-\epsilon_{\mathbf{k}+\mathbf{q}j'}}\times\psi_{\mathbf{k}+\mathbf{q}j'}(\mathbf{r})\psi^*_{\mathbf{k}j}(\mathbf{r})\psi^*_{\mathbf{k}+\mathbf{q}j'}(\mathbf{r}')\psi_{\mathbf{k}j}(\mathbf{r}'). \tag{65}$$

The operator π is an integral operator while $\hat{\Pi}$ is not necessarily one. It denotes the procedure of how to construct the change $\delta\rho$ from δV_{ext}. In particular, $\hat{\Pi}$ contains those part of the operator π in which the sum over conduction states runs only over the number which is equal to the number of the basis functions, N_{basis}. It also contains the contribution from the change in the basis functions according to (47), (55).

After the initial response $\delta\rho^{(0)}$ has been found, we have to calculate the screened perturbation (7). Let us call the

Coulomb interaction, $e^2/|\mathbf{r}-\mathbf{r}'|$, for v_C and the exchange-correlation interaction in the LDA, $dV_{\mathrm{xc}}/d\rho\times\delta(\mathbf{r}-\mathbf{r}')$, for v_{xc}. Then, the change δV_{eff} can be written as follows:

$$\delta V_{\mathrm{eff}}=\delta V_{\mathrm{ext}}+(v_C+v_{\mathrm{xc}})\delta\rho, \tag{66}$$

and the new electronic response $\delta\rho=\hat{\Pi}\delta V_{\mathrm{eff}}$. We thus see that the self-consistency of the induced charge density means solving the Dyson-like self-consistency equation:

$$\delta\rho=\delta\rho^{(0)}+\hat{\Pi}(v_C+v_{\mathrm{xc}})\delta\rho. \tag{67}$$

When $q\to 0$, the integral $v_C\delta\rho$ diverges as $1/q^2$ which immediately means that searching for the solution of Eq. (67) by iterations, i.e., $\delta\rho=\delta\rho^{(0)}+\hat{\Pi}(v_C+v_{\mathrm{xc}})\delta\rho^{(0)}+\dots$, is impossible. However, it is possible when the input to the next, $(i+1)$th, iteration is prepared by mixing the output and input densities from the previous, (i)th, iteration, i.e., $\delta\rho_{i+1}^{\mathrm{inp}}=\lambda_{\mathrm{mix}}\delta\rho_i^{\mathrm{out}}+(1-\lambda_{\mathrm{mix}})\delta\rho_i^{\mathrm{inp}}$, but the mixing parameter λ_{mix} must be chosen to be proportional to q^2. This makes the standard mixing schemes in the long-wavelength limit extremely time consuming.

This divergency problem is well known, and in the dielectric-matrix approach it is avoided by writing the solution (67) in the form

$$\delta\rho=\epsilon^{-1}\delta\rho^{(0)}, \tag{68}$$

where $\epsilon^{-1}=(1-\Pi v_C-\Pi v_{\mathrm{xc}})^{-1}$ is an inverse dielectric matrix of the crystal. [The relation (68) is usually written for the potentials δV_{eff} and δV_{ext} but in the present context it is more convenient to remain within the density language.] If for a metal Πv_C is proportional to $N(\epsilon_F)/q^2$, where $N(\epsilon_F)$ is the density of states at the Fermi energy ϵ_F, then ϵ^{-1} behaves as q^2 when $q\to 0$: this is the well-known long-wavelength behavior of the metallic dielectric function. What we actually do when solving (67) by iterations is trying to sum up $1-x+x^2-\dots=1/(1+x)$ for $x>1$.

In order to avoid this problem we use a Thomas-Fermi-like screening theory. To explain the idea we assume that the change in charge density and the potential are expanded in plane waves,

$$\delta\rho(\mathbf{r})=\sum_G \delta\rho(\mathbf{G})e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}. \tag{69}$$

We divide the Coulomb interaction v_C into long-range and short-range parts, i.e., $v_C=v_C^{\mathrm{long}}+v_C^{\mathrm{short}}$, where $v_C^{\mathrm{long}}=4\pi e^2/q^2\times\exp[i\mathbf{q}(\mathbf{r}-\mathbf{r}')]$. The exchange correlation in the LDA is always short ranged and can be treated together with v_C^{short}, i.e., $v_C^{\mathrm{short}}+v_{\mathrm{xc}}=w^{\mathrm{short}}$. The self-consistency equation can then be written as follows:

$$\delta\rho(\mathbf{r})=\delta\rho^{(0)}(\mathbf{r})+\frac{4\pi e^2}{q^2}\delta\rho(0)\Pi_{\mathbf{q}}(\mathbf{r})+\{\hat{\Pi}w^{\mathrm{short}}\delta\rho\}(\mathbf{r}), \tag{70}$$

where we have separated out the divergent contribution, $\delta\rho(0)\equiv\delta\rho(\mathbf{G}=0)$, and where we have called the response of electrons to the perturbation given by a single plane wave $\exp(i\mathbf{q}\mathbf{r})$ for $\Pi_{\mathbf{q}}(\mathbf{r})$. It can be written as an integral over the unit cell Ω_c:

$$\Pi_{\mathbf{q}}(\mathbf{r})=\int_{\Omega_c}\pi_{\mathbf{q}}(\mathbf{r},\mathbf{r}')e^{i\mathbf{q}\mathbf{r}'}d\mathbf{r}'. \tag{71}$$

The $\mathbf{G}=\mathbf{0}$ part of Eq. (70) can be written as follows:

$$\delta\rho(0)=\epsilon_{\mathrm{long}}^{-1}[\delta\rho^{(0)}(0)+\{\hat{\Pi}w^{\mathrm{short}}\delta\rho\}(\mathbf{G}=\mathbf{0})], \tag{72}$$

where we have introduced an effective dielectric constant:

$$\epsilon_{\mathrm{long}}=1-\frac{4\pi e^2}{q^2}\Pi_{\mathbf{q}}(\mathbf{G}=\mathbf{0}). \tag{73}$$

Inserting Eq. (72) in to the self-consistency equation (70), we obtain

$$\delta\rho(\mathbf{r})=\delta\rho^{(0)}(\mathbf{r})+\frac{4\pi e^2}{q^2+\kappa_D^2}[\delta\rho^{(0)}(0)+\{\hat{\Pi}w^{\mathrm{short}}\delta\rho\}\times(\mathbf{G}=\mathbf{0})]\Pi_{\mathbf{q}}(\mathbf{r})+\{\hat{\Pi}w^{\mathrm{short}}\delta\rho\}(\mathbf{r}), \tag{74}$$

where $\kappa_D^2=-4\pi e^2\Pi_{\mathbf{q}}(\mathbf{G}=\mathbf{0})$ is the Debye screening radius. The screened equation (74) is free of the difficulties discussed above and can be solved iteratively. First, one has to find the function $\Pi_{\mathbf{q}}(\mathbf{r})$ as the response of electrons to a single plane wave $\exp(i\mathbf{q}\mathbf{r})$, and from that obtain κ_D^2. Then the initial distribution $\delta\rho^{(0)}(\mathbf{r})$ is calculated. During the iterations one first finds the response to the short-range part of the perturbation, i.e., $\{\hat{\Pi}w^{\mathrm{short}}\delta\rho\}(\mathbf{r})$, and, second, the long-wavelength contribution is added as given by the second term in the right-hand side of Eq. (74). The output change in the charge density is usually mixed with the input $\delta\rho$ to obtain an input for the new iteration. This makes the self-consistent cycle stable, but the mixing parameter λ_{mix} in this case does not have to go to zero for $\mathbf{q}\to\mathbf{0}$ and it is usually chosen to be $0.2-0.5$. In practical applications we have found that the number of iterations required to solve (74) is about 10 while for solving the original equation (67) the number of iterations varies from 50 to 200 depending on the length $|\mathbf{q}|$ of the wave vector. The latter is, of course, not true for those phonon modes where $\delta\rho(0)\equiv 0$ by symmetry.

One can obviously consider the screening of not only the component $\delta V(\mathbf{G})\sim\delta\rho(\mathbf{G})/|\mathbf{q}+\mathbf{G}|^2$ with $\mathbf{G}=\mathbf{0}$ but all the components within a certain sphere $|\mathbf{q}+\mathbf{G}|\leqslant E_{\mathrm{cutoff}}$. This, for instance, is necessary for those zone-boundary wave vectors where $|\mathbf{q}|=|\mathbf{q}+\mathbf{G}|$. In this case the function $\Pi_{\mathbf{q}}(\mathbf{r})$ is replaced by the functions $\Pi_{\mathbf{q}+\mathbf{G}}(\mathbf{r})$, i.e., at the beginning it is necessary to calculate the response of the electrons to the perturbation $\exp[i(\mathbf{q}+\mathbf{G})\mathbf{r}]$. The corresponding self-consistency equation should be written again to account for the fact that ϵ_{long} is now the matrix $\epsilon_{\mathrm{long}}(\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}')$. This will reduce the number of iterations even more.

Finally, we would like to point out that it should be possible to apply the same idea to the self-consistency problem in the standard band-structure calculation. In the crystal, due to electroneutrality of the charge density, the contribution to the Coulomb potential from $\rho(\mathbf{G}=\mathbf{0})$ drops out. However, for those reciprocal-lattice vectors which are small, the components of the potential $V(\mathbf{G})\sim\rho(\mathbf{G})/|\mathbf{G}|^2$ might be large. This is especially the case for large many-atomic unit cells. As a consequence, the mixing parameter λ_{mix} has to be chosen very small. The procedure described above will require the calculation of the polarizability (65) with $\mathbf{q}=\mathbf{0}$ at each

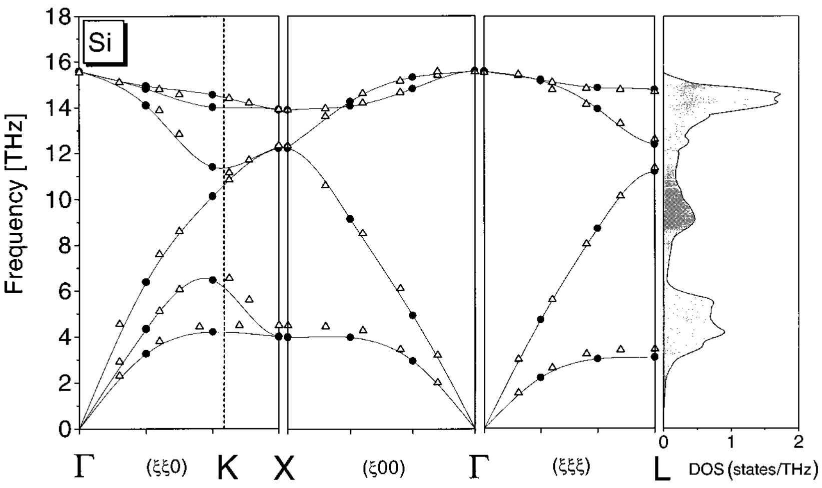

FIG. 1. Calculated phonon dispersion for Si (circles) along the high-symmetry directions in comparison with the experiment (Ref. 44) (triangles). The lines are the result of interpolation between theoretical points. Also shown is the calculated phonon density of states (DOS).

self-consistent iteration, i.e., the response of electrons to the plane waves $\exp[i\mathbf{Gr}]$ according to the expression (71) for all small vectors $|\mathbf{G}|\leqslant G_{\text{cutoff}}$. The cutoff can be chosen as the radius of the smallest first coordination sphere in G space. The computational time for finding the $\Pi_G(\mathbf{r})$ functions should presumably not exceed the time of one self-consistent iteration while the total number of iterations needed to reach the convergency is expected to decrease by approximately one order of magnitude, which is the case in linear-response calculation. Note that the idea just outlined is different from the idea of finding the self-consistent charge transfer in terms of the linear-response theory.[43] For large cells we are screening the small G components of the potential which result from some average density distribution. On the other hand, such details as the charge transfer between nearest atoms is described by large G components of $\rho(\mathbf{G})$.

IV. APPLICATIONS

In recent publications[12,15,16] we have applied our linear-response method to compute whole phonon dispersions and electron-phonon interactions in transition metals Nb and Mo. In the present paper we will describe the application of the method for calculating phonon dispersions in the materials with a few atoms per unit cell and with a relatively open crystalline structures. Two systems have been chosen for the applications. The first one is Si which is an excellent test case because of its open diamond structure. The second one is a transition-metal carbide NbC. This is a well-known classic superconductor with $T_c=11.5$ K and its phonon dispersions show many anomalies that are not present in other simple-metallic and insulating systems. Studying these anomalies as well as their influence on superconductivity and transport is interesting in itself and also represents a hard test for our method. Here we will only describe the calculations for the phonon dispersion curves in NbC and compare the results with experiments. The calculated electron-phonon interaction and transport properties will be published elsewhere.

A. Si

Si is a well-studied elemental semiconductor from both experimental and theoretical sides and its phonon dispersions have been measured a long time ago.[44] Recent linear-response[6,9] and supercell[20] calculations have allowed us to determine its lattice dynamics for the wave vectors in the entire Brillouin zone and the results show a good agreement with the experiment. These calculations were based on the linear-augmented-plane-wave and plane-wave pseudopotential methods. Within the localized-orbital representation employed in the LMTO method it is generally difficult to treat the materials with the diamond structure and, to reach close packing, interstitial spheres are usually placed into the empty sites of the lattice. This complicates the determination of the dynamical matrix. However, this problem is avoided by the use of the Fourier transform for the LMTO's in the interstitial region.

We calculate the dynamical matrix of Si as a function of wave vector for a set of irreducible $\mathbf{q}$ points in a (6,6,6)-reciprocal lattice grid (16 points per 1/48th part of the BZ). The (I,J,K) reciprocal-lattice (or Monkhorst-Pack) grid is defined in the usual manner: $\mathbf{q}_{ijk}=(i/I)\mathbf{G}_1+(j/J)\mathbf{G}_2+(k/K)\mathbf{G}_3$, where $\mathbf{G}_1$, $\mathbf{G}_2$, $\mathbf{G}_3$ are the primitive translations in reciprocal space. The details of the calculations for every $\mathbf{q}$ point are the following: We use a $3\kappa-spd$ LMTO basis set (27 orbitals per atom) with the one-center expansions inside the MT spheres performed up to $l_{\max}=6$. In the interstitial region, the s-, p-, and d-basis functions are expanded in plane waves up to 15.1, 22.3, 31.7 Ry (282, 530, 868 plane waves), respectively. The induced charge densities and screened potentials are represented inside the MT spheres by spherical harmonics up to $l_{\max}=6$ and by plane waves with the 110.2 Ry energy cutoff (5208 plane waves) in the interstitial region. The $\mathbf{k}$ integration over the BZ is performed over the (6,6,6) grid (the same grid as for the phonon wave vectors $\mathbf{q}$) by means of the improved tetrahedron method,[42] which is identical in the case of Si to the special-point method of Monkhorst and Pack. The MT-sphere radius was taken to be 2.214 a.u. and the von Barth–Hedin-like exchange-correlation formulas after Ref. 45 are employed. We use theoretically determined lattice parameter in the calculation (the volume ratio $V/V_{exp}=0.991$).

Figure 1 shows a comparison between calculated and experimental phonon dispersion curves along the major high-symmetry directions. The calculated phonon density of states is plotted at the right part in the figure. The theoretical fre-

TABLE I. Comparison between calculated and experimental phonon frequencies at the high-symmetry points Γ, X, and L for Si (THz).

	Γ_{LTO}	X_{TA}	X_{LAO}	X_{TO}	L_{TA}	L_{LA}	L_{TO}	L_{LO}
Theory	15.56	4.00	12.27	13.90	3.09	11.20	14.78	12.38
Exp.[a]	15.53	4.49	12.32	13.90	3.43	11.35	14.68	12.60

[a] Reference 44.

quencies are denoted by circles and the experimental ones are denoted by triangles. The lines result from the interpolation between the theoretical points. The calculated and experimental phonon frequencies at the high-symmetry points Γ, X, and L are also listed in Table I. We see that the agreement between theory and experiment is very good. Especially, in the optical region the discrepancy is about 1–1.5 %, which is surprising because the accuracy of the measured phonon modes is of the same order of magnitude. We also reproduce the extended flat regions of the transverse acoustic modes indicating the accurate description of long-range interactions between Si atoms as well as the correct long-wavelength behavior showing the good accuracy of calculated elastic properties of this crystal. Larger discrepancy is found for the frequencies of the TA modes, where the theoretical branches are approximately 10% softer than the experimental ones. For instance, the calculated frequency of the X_{TA} mode is 4.00 THz, while $\omega_{\mathrm{exp}}(X_{\mathrm{TA}})=4.49\pm0.06$ THz.[44] The same kind of discrepancy has also been recently reported in Refs. 9 and 20. The agreement is slightly improved when we recalculate the dynamical matrix at the X point using the experimental lattice constant. We have found that the frequency of the X_{TA} mode is increased from 4.00 to 4.27 THz. This shows that the mode has a large negative Gruneizen parameter and it is thus very sensitive to the unit-cell volume used in the calculation. Because of the large LMTO basis sets, large l_{max}, and plane-wave energy cutoffs, this discrepancy is hard to relate to the internal parameters in the calculation. Nevertheless, a very good agreement found for this mode by the previous pseudopotential calculation[6] could point out that this problem needs more careful examination.

B. NbC

The lattice-dynamical properties of transition-metal carbides and, especially, NbC have attracted much attention in the past because of the existence of pronounced anomalies in its acoustic branches and their influence to superconductivity. While some model calculations of the phonon dispersions exist in the literature and various mechanisms explaining these anomalies have been proposed,[46] no *ab initio* investigation of the lattice dynamics for NbC have so far been performed. Here we apply the linear-response approach to the phonon spectrum of NbC in order to check the accuracy of our method.

The dynamical matrix of NbC is calculated at the 29 irreducible **q** points of the (8,8,8) reciprocal-lattice grid. The self-consistent calculations performed for every wave vector involve the following parameters: $3\kappa-spd$ LMTO basis per Nb atom (27 orbitals) and $3\kappa-sp$ LMTO basis per carbon atom (12 orbitals). The one-center expansions inside the MT spheres are performed up to $l_{\mathrm{max}}=6$. In the interstitial region the basis functions are expanded in plane waves up to 13.4, 19.6, 26.9 Ry (136, 228, 338 plane waves) for, respectively, s, p, and d orbitals of Nb, and up to 24.1, 35.8 Ry (306, 536 plane waves) for s, p orbitals of C. The changes in the charge densities and the potentials are represented inside the MT spheres by spherical harmonics up to $l_{\mathrm{max}}=6$ and by plane waves with an 121 Ry energy cutoff (3382 plane waves) in the interstitial region. The **k**-space integration for the matrix elements is performed over a (8,8,8) grid (the same grid as for the phonon wave vectors **q**) by means of the improved tetrahedron method.[41] However, the integration weights for the **k** points of this grid have been found to take into account the effects arising from the Fermi surface and the energy bands precisely. This is done with the help of a (32, 32, 32) grid (897 **k** points per 1/48 BZ) as we explained in Sec. III C of this paper. The MT-sphere radius of Nb is taken to be 2.411 a.u. and the radius of the carbon sphere is 1.786 a.u. The von Barth–Hedin-like exchange-correlation formulas after Ref. 45 are employed. As in the case of Si, we also use the theoretically determined lattice parameter in this calculation (the volume ratio $V/V_{\mathrm{exp}}=0.982$).

The results of our calculations are presented in Fig. 2, where we compare theoretically determined phonon disper-

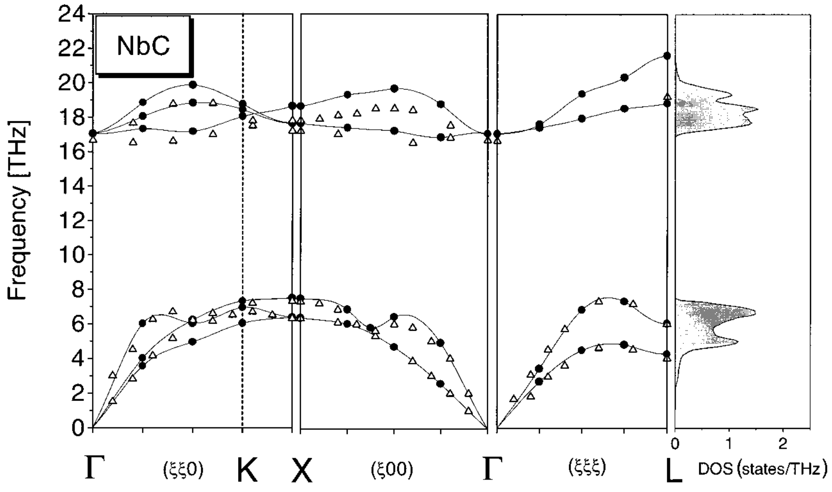

FIG. 2. Calculated phonon dispersion for NbC (circles) along the high-symmetry directions in comparison with the experiment (Ref. 47) (triangles). The lines are the result of interpolation between theoretical points. Also shown is the calculated phonon density of states (DOS).

TABLE II. Comparison between calculated and experimental phonon frequencies at the high-symmetry points Γ, X, and L for NbC (THz).

	Γ_{LTO}	X_{TA}	X_{LA}	X_{TO}	X_{LO}	L_{TA}	L_{LA}	L_{TO}	L_{LO}
Theory	17.05	6.37	7.51	17.64	18.65	4.26	6.02	18.82	21.60
Exp.[a]	16.70	6.35	7.30	17.20	17.80	4.00	6.00		19.20

[a] Reference 47.

sions (circles) with those measured by inelastic-neutron-scattering technique[47] (triangles). The calculated phonon density of states is plotted at the right part of the figure. The lines are simply the result of interpolation between the theoretical points. Since the **q** grid (8,8,8) considered here is still too coarse to resolve the anomaly of the longitudinal acoustic branch near the wave vector (0.6,0,0) $2\pi/a$, we have performed a separate calculation for the **q** point (0.625,0,0) which fits to the (16,16,16) grid in **k** space. We see that the agreement between theory and experiment is good. Most of the calculated frequencies agree within a few percent with those measured despite the fact that we have used only 29 **k** points for the BZ integration. (We list for comparison our calculated and experimental phonon frequencies at the high-symmetry points Γ, X, and L in Table II.) The theory reproduces the major anomalies presented in the acoustic branches: the well-known anomaly near the wave vector (0.6,0,0) $2\pi/a$ which is also present and well described[12] within our linear-response method in pure Nb crystal; the anomaly of the longitudinal mode near the wave vector (0.5,0.5,0) $2\pi/a$ as well as large softening of both TA and LA modes near the L point. Moreover, we also predict an anomalous behavior of the lowest transverse acoustic mode along the $(\xi\xi 0)$ direction near the wave vector (0.5,0.5,0) $2\pi/a$. Here the frequencies are not known experimentally. The anomaly found by us is, however, less pronounced compared to the results of double-shell model calculations of Weber,[48] while we have certainly not too many points along this direction to judge about its exact dispersion.

V. CONCLUSION

In conclusion, we have described in detail an all-electron linear-response approach based on the density functional theory and the LMTO technique. The method is developed to calculate lattice dynamical properties of crystalline solids and is uniquely applicable for the systems with broad and narrow energy bands. For test purposes, we have applied the method to compute phonon dispersions for Si and NbC, which have open structures and two atoms per unit cell. The results of our applications are in a good agreement with the experiment. We have thus shown that accurate calculations of lattice dynamics are now possible even for such complicated systems as transition-metal compounds. In the following paper[18] we give a description of our method for calculating electron-phonon interactions and apply the method to compute lattice-dynamical, superconducting, and transport properties for a large number of elemental metals (a brief report of this work has appeared already[15]). In another publication[49] we describe an application of the method for computing electron-phonon-coupling strength in a Ca-Sr-Cu-O high-T_c superconductor.

ACKNOWLEDGMENTS

The author is indebted to Professor O. K. Andersen for many helpful discussions and to Dr. O. Jepsen for careful reading of the manuscript prior to the publication. The work was partially supported by Grant Nos. INTAS (93-2154), ISF(MF-8300), and RFFI.

APPENDIX: FOURIER TRANSFORM OF PSEUDO-LMTO'S

Consider a Hankel function $K_{\kappa L}(\mathbf{r}) = K_{\kappa l}(r) i^l Y_{lm}(\mathbf{r})$ of energy κ^2 which is singular at the origin. The three-dimensional Fourier transform of this function $K_{\kappa L}(\mathbf{k})$ is known to behave as k^{l-2} for large k. The task is to substitute the divergent part of $K_{\kappa l}(r)$ inside some sphere s by a smooth regular but otherwise arbitrary function. This function is chosen so that the Fourier transform is convergent fast. In the full-potential LMTO method of Ref. 34, the augmenting function is the linear combination of the Bessel function $J_{\kappa L}$ and its energy derivative $\dot{J}_{\kappa L}$ matched together with its first-order radial derivative with the Hankel function at the sphere boundary. The Fourier transform becomes convergent as k^{-4}. One can obviously include higher-order energy derivatives $J_{\kappa L}^{(n)}$ in order to have a smooth matching at the sphere up to the order n. This was done in connection with the problem of solving the Poisson equation in Ref. 39. The Fourier transform here converges as $k^{-(3+n)}$ but the prefactor increases as $(2l+2n+3)!!$ and this prohibits the use of large values of n. A similar procedure has been also used in the LMTO method of Ref. 35. In the present work we will use a different approach based on the Ewald method. Instead of substituting the divergent part only for $r<s$ we consider the solution of the equation

$$(-\nabla^2-\kappa^2)\widetilde{K}_{\kappa L}(\mathbf{r}) = a_l\left(\frac{r}{s}\right)^l e^{-r^2\eta^2+\kappa^2/\eta^2} i^l Y_{lm}(\mathbf{r}). \tag{A1}$$

The function on the right-hand side of the Helmholtz equation is a decaying Gaussian. The parameter a_l is a normalization constant: $a_l = \sqrt{2/\pi}(2\eta^2)^{l+3/2} s^{2l+1}/(2l-1)!!$. The most important parameter is η. It is chosen such that the Gaussian is approximately zero when $r>s$ and η must depend on l as well as the sphere radius s. The solution $\widetilde{K}_{\kappa L}(\mathbf{r})$ is thus the Hankel function for large r, it is a regular function for small r and it is smooth together with its radial derivatives at any r. The function $\widetilde{K}_{\kappa l}(r)$ can be calculated in terms of the following error-function-like contour integral:

$$\widetilde{K}_{\kappa l}(r) = \frac{(2s)^{l+1}}{\sqrt{\pi}(2l-1)!!} r^l \int_{0+}^{\eta} \xi^{2l} e^{-r^2\xi^2+\kappa^2/4\xi^2} d\xi. \tag{A2}$$

When $\eta \to \infty$ this integral is known as the Hankel integral. The most important result is that the Fourier transform of $\widetilde{K}_{\kappa l}(r)$ decays exponentially. It is given by

$$\widetilde{K}_{\kappa l}(r) = \frac{2}{\pi}\frac{s^{l+1}}{(2l-1)!!}\int_0^\infty k^2 dk\, j_l(kr)\frac{k^l e^{(\kappa^2-k^2)/4\eta^2}}{k^2-\kappa^2}. \tag{A3}$$

Restoring the original notations, the pseudo-LMTO's $\widetilde{\chi}^{\mathbf{k}}_{\kappa RL}(\mathbf{r})$ are the Bloch waves of wave vector $\mathbf{k}$ as defined in Eq. (46). The Fourier coefficients $\widetilde{\chi}_{\kappa RL}(\mathbf{k}+\mathbf{G})$ are given by

$$\widetilde{\chi}_{\kappa RL}(\mathbf{k}+\mathbf{G}) = \frac{4\pi}{\Omega_c} \frac{s_R^{l+1}}{(2l-1)!!} \frac{|\mathbf{k}+\mathbf{G}|^l}{|\mathbf{k}+\mathbf{G}|^2-\kappa^2} \times e^{(\kappa^2-|\mathbf{k}+\mathbf{G}|^2)/4\eta_{Rl}^2} Y_L(\mathbf{k}+\mathbf{G}) e^{-i(\mathbf{k}+\mathbf{G})\mathbf{R}}, \tag{A4}$$

where Ω_c is the volume of the unit cell and where we have subscripted η with the indexes Rl and s with R.

In practical calculations the parameter η_{Rl} can be chosen from the ratio between the Hankel function at the sphere and the solution of Eq. (A1), i.e., $\widetilde{K}_{\kappa l}(s_R)/K_{\kappa l}(s_R) = 1+\delta$. The error $|\delta|$ is usually taken not larger than 0.03 which leads to the number of plane waves per atom needed for the convergency in (46) varying from 150 to 250 when $l=2$. For the s,p orbitals this number is smaller by a factor of $2-3$.

---

*On leave from P. N. Lebedev Physical Institute of the Russian Academy of Sciences, Leninski pr. 53, 117924 Moscow, Russia.

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[3] For a review, see, also, *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and S. H. March (Plenum, New York, 1983).

[4] M. S. Hybersten and S. G. Louie, Phys. Rev. B **35**, 5585 (1987); **35**, 5606 (1987).

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

[6] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).

[7] D. Kh. Blat, N. E. Zein, and V. I. Zinenko, J. Phys. Condens. Matter **3**, 5515 (1991).

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

[9] R. Yu and H. Krakauer, Phys. Rev. B **49**, 4467 (1994).

[10] R. Bauer, O. Schütt, P. Pavone, W. Windl, and D. Strauch, Phys. Rev. B **51**, 210 (1995).

[11] S. Y. Savrasov, Solid State Commun. **74**, 69 (1990).

[12] S. Y. Savrasov, Phys. Rev. Lett. **69**, 2819 (1992).

[13] N. E. Zein, Phys. Lett. A **161**, 526 (1992).

[14] A. A. Quong, B. M. Klein, Phys. Rev. B **46**, 10 734 (1992).

[15] S. Y. Savrasov, D. Y. Savrasov, and O. K. Andersen, Phys. Rev. Lett. **72**, 372 (1994).

[16] S. Y. Savrasov and E. G. Maksimov, Usp. Fiz. Nauk **165**, 773 (1995) [Sov. Phys. Usp. **38**, 737 (1995)].

[17] S. de Gironcoli, Phys. Rev. B **51**, 6773 (1995).

[18] S. Y. Savrasov and D. Y. Savrasov, following paper, Phys. Rev. B **54**, 16 487 (1996).

[19] H. Wendel and R. M. Martin, Phys. Rev. B **19**, 5251 (1979); K. Kunc and E. Tosatti, *ibid.* **29**, 7045 (1984).

[20] S. Wei and M. Y. Chou, Phys. Rev. Lett. **69**, 2799 (1992).

[21] X. Gonze, Phys. Rev. A **52**, 1096 (1995).

[22] An overview is given by A. R. Williams and U. von Barth, in Ref. 3, p. 189.

[23] See, e.g., *Ab Initio Calculations of Phonon Spectra* , edited by J. T. Devreese, V. E. Doren, and P. E. Van Camp (Plenum, New York, 1983).

[24] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).

[25] N. E. Zein, Fiz. Tverd. Tela (Leningrad) **26**, 3024 (1984) [Sov. Phys. Solid State **26**, 1825 (1984)].

[26] R. M. Sternheimer, Phys. Rev. **96**, 951 (1954); **107**, 1565 (1957); **115**, 1198 (1959).

[27] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[28] P. Pulay, Mol. Phys. **17**, 197 (1969).

[29] S. Y. Savrasov and D. Y. Savrasov, Phys. Rev. B **46**, 12 181 (1992).

[30] J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, *Advances in Quantum Chemistry* (Academic, New York, 1964), Vol. 1, pp. 267, 288, and 300.

[31] A. Debernardi, S. Baroni, and E. Molinari, Phys. Rev. Lett. **75**, 1819 (1995).

[32] M. Springborg and O. K. Andersen, J. Chem. Phys. **87**, 7125 (1987).

[33] O. K. Andersen, O. Jepsen, and G. Krier, in *Lectures on Methods of Electronic Structure Calculations*, edited by V. Kumar, O. K. Andersen, and A. Mookerjee (World Scientific, Singapore, 1994).

[34] K. H. Weyrich, Phys. Rev. B **37**, 10 269 (1988).

[35] J. Wills (unpublished).

[36] M. Methfessel, Phys. Rev. B **38**, 1537 (1988).

[37] P. Blochl, Ph.D. thesis, Max-Planck Institut für Festkörperforschung, 1989.

[38] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988).

[39] M. Weinert, J. Math. Phys. **22**, 2433 (1981).

[40] O. Jepsen and O. K. Andersen, Phys. Rev. B **29**, 5965 (1984).

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

[42] P. Blochl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16 223 (1994).

[43] H. L. Skriver and N. M. Rosengaard, Phys. Rev. B **46**, 7157 (1992).

[44] G. Dolling, in *Inelastic Scattering of Neutrons in Solids and Liquids* (IAEA, Vienna, 1963), Vol. II, p. 37; G. Nillsson and G. Nelin, Phys. Rev. B **8**, 5711 (1973).

[45] U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972); V. L. Moruzzi, J. F. Janak, and A. R. Williams, *Calculated Electronic Properties of Metals* (Pergamon, New York, 1978).

[46] For a review, see, e.g., S. K. Sinha in *Dynamical Properties of Solids*, edited by A. A. Maradudin and G. K. Horton (North-Holland, Amsterdam, 1980), Vol. 3, p. 1.

[47] H. G. Smith and W. Gläser, in *Proceedings of the International Conference on Phonons*, edited by M. A. Nusimovici (Flammarion, Paris, 1971).

[48] W. Weber, Phys. Rev. B **8**, 5082 (1973).

[49] S. Y. Savrasov and O. K. Andersen, Phys. Rev. Lett. (to be published).