Product-basis method for calculating dielectric matrices

F. Aryasetiawan and O. Gunnarsson
Max-Planck-Institut für Festkörperforschung Heisenbergstrasse 1, 70569 Stuttgart, Federal Republic of Germany
(Received 15 December 1993; revised manuscript received 24 February 1994)

We present a method for calculating dielectric matrices of periodic systems. Unlike the conventional method, which uses a plane-wave basis, the present method employs a product basis, which, in the linear-muffin-tin-orbital formalism, consists of products of orbitals. The method can be used for any system, including *sp* as well as narrow band systems. We demonstrate the applicability of our method by calculating the energy-loss spectra of Ni and Si, including local-field effects that require the full dielectric matrix. Good agreement with experiment is found. The small number of basis functions makes the method suitable for self-energy calculations within the *GW* approximation, without making the so-called plasmon-pole approximation for the dielectric matrix.

I. INTRODUCTION

The response function of an electronic system has long been recognized as a basic quantity to describe many important properties of the system, such as optical, transport, electronic, and magnetic properties. In this paper, we are concerned with the density-density response function P, i.e., the function describing a change in the electron density $\delta\rho$ due to an external perturbation δV^{ext}: $\delta\rho = \delta V^{\text{ext}} P$.

P is a basic ingredient in many-body perturbation theory, which has proven to be very successful in atomic physics.[1] Its application to electronic structures in solid-state physics, which is our main interest, has been very slow due to a prohibitively large computational size in obtaining P. Fortunately, *ab initio* many-body perturbation calculations have become feasible over the last few years, due to rapid progress in computer technology and a better understanding of theory itself. With the invention of the linear method,[2] we can now obtain single-particle band structures of complex systems with relative ease. They define the zeroth order Hamiltonian in a many-body perturbation calculation. A working approximation for the evaluation of a perturbation series was developed in the early sixties by Hedin and Lundqvist known as the *GW* approximation (*GWA*),[3,4] and it has been applied to many systems with success.[5–8]

In most *GW* calculations, the screened potential $W = \epsilon^{-1}v$, where ϵ^{-1} is the screening matrix and v is the bare Coulomb potential, is calculated in a plane-wave basis. This is a natural choice for systems that can be suitably treated with pseudopotentials[9] for two reasons: first, the single-particle wave functions, which are needed in the evaluation of P, are expanded in plane waves, and second, the Coulomb potential becomes trivial, since it is diagonalized by the plane waves. Other advantages of the plane-wave basis are good control over convergence with respect to the number of basis functions and programming ease. There are, however, serious drawbacks.

(1) It is not possible to do all-electron calculations. In many cases, it is important to include core electrons. For instance, in the case of Ni, the exchange of a valence state with the core states is $\sim -10\,\text{eV}$.

(2) The number of plane waves becomes prohibitively large for narrow band systems.

Moreover, the plane waves have no direct physical interpretation.

To overcome these drawbacks, we use the linear muffin-tin-orbital (LMTO) method,[2] which allows us to treat any system. The LMTO method uses a minimal number of basis functions and we carry over the concept of minimal basis in band structures to the dielectric matrix ϵ. Instead of a plane-wave basis, we use a "product basis," which consists of products of LMTO's. As will become clear later, the product basis constitutes a minimal basis for ϵ within the LMTO formalism. A method for inverting the dielectric matrix using localized Wannier orbitals instead of plane waves has also been used in the context of local-field and excitonic effects in the optical spectrum of covalent crystals.[10]

Our method for evaluating ϵ can be used for any system, although our interest is in transition metals and their compounds (e.g., Ni and NiO). Due to a small number of basis functions, the method allows us to obtain the dielectric matrices for these systems without making the so-called plasmon-pole approximation,[11–13] whose validity in these systems is doubtful. We demonstrate the applicability of our method by calculating the energy-loss spectra of Ni and Si, including local-field effects, which require the full calculation of the dielectric matrix. We have also performed *GW* calculations for these narrow band systems, but this will be the subject of subsequent papers. In this paper we concentrate on the numerical machinery behind the calculations.

A different method for dealing with narrow band systems has also been devised before, using a modified LAPW method for the band structure and basis functions.[8,14] This method has been applied with success for calculating the energy-loss spectra and self-energy of Ni. Due to a relatively large number of basis functions needed in the LAPW method, extension to larger systems is hampered by the computational size.

The paper is organized as follows. In Sec. II we describe the theory and the method and in Sec. III we apply the method to calculate the energy-loss spectra of Ni and Si. Finally, in Sec. IV we draw our conclusions.

II. THEORY

A. The dielectric matrix

The linear response function for a noninteracting system is given by[15]

$$P^0(\mathbf{r},\mathbf{r}';\omega) = \int_0^\infty d\omega' S^0(\mathbf{r},\mathbf{r}';\omega') \times \left\{ \frac{1}{\omega-\omega'+i\delta} - \frac{1}{\omega+\omega'-i\delta} \right\}, \quad (1)$$

where

$$S^0(\mathbf{r},\mathbf{r}';\omega) = \sum_{\mathbf{k}n}^{\text{occ}} \sum_{\mathbf{k}'n'}^{\text{unocc}} \psi^*_{\mathbf{k}n}(\mathbf{r})\psi_{\mathbf{k}'n'}(\mathbf{r})\psi^*_{\mathbf{k}'n'}(\mathbf{r}')\psi_{\mathbf{k}n}(\mathbf{r}') \times \delta(\omega-\epsilon_{\mathbf{k}'n'}+\epsilon_{\mathbf{k}n}). \quad (2)$$

$\psi_{\mathbf{k}n}$ and $\epsilon_{\mathbf{k}n}$ are the Bloch state and its eigenenergy of the corresponding noninteracting Hamiltonian H^0. We have made use of the time-reversal symmetry in the above expression, i.e., for every $\psi_{\mathbf{k}n}$ there is $\psi^*_{-\mathbf{k}n}$ with the same eigenvalue.

The inverse dielectric matrix, which determines the screening properties, is given by

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r}-\mathbf{r}') + \int d^3r'' v(|\mathbf{r}-\mathbf{r}''|)P(\mathbf{r}'',\mathbf{r}';\omega), \quad (3)$$

where P is the total response function which, in the random phase approximation,[16] is related to P^0 by the integral equation

$$P = P^0 + P^0 v P. \quad (4)$$

Calculating ϵ^{-1} in real space is not feasible because $\mathbf{r}$ and $\mathbf{r}'$ are not restricted to a unit cell and ϵ^{-1} can be long range. The usual procedure is to expand it in a complete set of Bloch functions $\{B_{\mathbf{k}i}\}$ in the following form:

$$\epsilon^{-1}(\mathbf{r},\mathbf{r}';\omega) = \sum_{\mathbf{k}}^{\text{1st BZ}} \sum_{ij} B_{\mathbf{k}i}(\mathbf{r})\epsilon^{-1}_{ij}(\mathbf{k},\omega)B^*_{\mathbf{k}j}(\mathbf{r}'). \quad (5)$$

Similarly, P^0, P, and v may also be expanded in the above form.

In almost all calculations, the basis functions are chosen to be plane waves, which are appropriate for *sp* systems. Even for these systems, the computational size is still too large and one resorts to a plasmon-pole approximation[11,12] where, in essence, each matrix element of ϵ^{-1} is replaced by a δ function, whose position and weight are determined by the static and f-sum rules. A more general form has also been proposed recently.[13] In simple metals and *sp* semiconductors/insulators, the main plasmon excitation forms a well defined sharp peak, with little structure and well separated from the single-particle-like excitations. The plasmon-pole approximation is expected to work well in these systems. In a heavy alkaline metal (Cs), however, the plasmon excitations form several peaks.[17] In transition metals the situation is a lot worse because there are no well defined plasmon peaks in the first place.[18,14] The plasmon and single-particle-like excitations merge, resulting in a complicated structure with two main peaks separated by a few eV, as can be seen in the energy loss spectra of Cu or Ni. For these systems, the plasmon-pole approximation is not expected to work well. Should one nevertheless try to use the plasmon-pole approximation by having several δ functions instead of one δ function to represent these peaks, one would run out of sum rules, which are needed to determine the positions and weights of these δ functions.

We would like to be able to treat not only *sp* systems, but also *d* and *f* systems with narrow bands without resorting to uncontrolled approximations such as the plasmon-pole approximation. Already with the former systems, it is a large undertaking to calculate ϵ^{-1} without the plasmon-pole approximation, even when the systems are small. With the latter systems, we simply cannot afford to use plane waves and therefore we have taken a completely different approach, which we describe below.

B. The product basis

The method described in this section is quite general, but we have applied it within the framework of the LMTO method,[2] which uses a minimal number of basis functions. For an accurate description of the valence states, one needs typically 9 (*spd*) or 16 (*spdf*) LMTO's. To calculate P^0, however, we need both the valence and conduction states; the latter can be up to 50–100 eV. One expects that additional basis functions are needed for a proper description of the conduction bands. In a previous paper,[19] we described how the LMTO method could be extended to handle both valence and high lying conduction states by having multiple orbitals with different energies for a given $L \equiv lm$ channel.

In the multiple LMTO method, the basis functions are of the form

$$\chi_{\mathbf{R}L\nu} = \phi_{\mathbf{R}L\nu} + \sum_{\mathbf{R}'L'\nu'} \dot{\phi}_{\mathbf{R}'L'\nu'} h_{\mathbf{R}'L'\nu',\mathbf{R}L\nu} \quad (6)$$

where $L\nu$ denotes the angular momentum and principal quantum number, $\phi_{\mathbf{R}L\nu}$ is a radial solution to the Schrödinger equation inside a sphere centered on an atom at site $\mathbf{R}$ for a certain energy ϵ_ν, normally chosen at the center of the band, and $\dot{\phi}$ is the energy derivative of ϕ at ϵ_ν. We have used the atomic sphere approximation where the Wigner-Seitz cell is replaced by sphere(s) of equal volume. An important point to note is that both ϕ and $\dot{\phi}$ are *independent* of $\mathbf{k}$ and band index n. Information about the lattice structure and the scattering phase

shifts (logarithmic derivatives of the wave functions at the surface of the Wigner-Seitz spheres) is carried in h, which therefore contains the $\mathbf{k}$ dependence when a Bloch sum is performed. For more details, we refer to Ref. 19. The wave functions are expanded in these LMTO's

$$\psi_{\mathbf{k}n} = \sum_{\mathbf{R}L\nu} \chi_{\mathbf{R}L\nu}(\mathbf{k}) b_{\mathbf{R}L\nu}(\mathbf{k}n). \tag{7}$$

Unlike the pseudopotential method with a plane-wave basis, it is not clear anymore what basis functions should now be used to expand ϵ^{-1}. Let us introduce a notation:

$$\{P^0\} \equiv \text{Hilbert space spanned by } P^0 \ . \tag{8}$$

To solve the problem, we first figure out what Hilbert space is spanned by ϵ^{-1}. We expand Eq. (4)

$$P = P^0 + P^0 v P^0 + P^0 v P^0 v P^0 + \cdots \ . \tag{9}$$

Since v is always sandwiched between two P^0's, it is clear that $\{P\}$, is identical to $\{P^0\}$. Thus, from Eq. (3) we see that $\{\epsilon^{-1}\}$ is the union between $\{P^0\}$ and $\{v\}$,

$$\{\epsilon^{-1}\} = \{P^0\} \cup \{v\}. \tag{10}$$

In practice, however, we usually take matrix elements of ϵ^{-1}, or use it in conjunction with other quantities so that we may not actually need the entire space. We consider two examples that are closely related to our area of interest.

(1) The energy loss spectra which are given by

$$\epsilon^{-1}(\mathbf{q},\omega) = 1 + v(\mathbf{q})\{\langle \mathbf{q}|P^0|\mathbf{q}\rangle + \langle \mathbf{q}|P^0 v P^0|\mathbf{q}\rangle + \cdots\}, \tag{11}$$

where $|\mathbf{q}\rangle = 1/\sqrt{V}\exp(i\mathbf{q}\cdot\mathbf{r})$ and $v(\mathbf{q}) = 4\pi/q^2$. We can read from the above equations that the required Hilbert space is

$$\{\epsilon^{-1}(\mathbf{q},\omega)\} = \{P^0\}. \tag{12}$$

(2) The self-energy of a given state $\psi_{\mathbf{k}n}$ in the GWA schematically has the form[8]

$$\begin{aligned}\Sigma(\mathbf{k}n,\omega) &= \langle \psi_{\mathbf{k}n}|GW|\psi_{\mathbf{k}n}\rangle \\ &= \langle \psi_{\mathbf{k}n}\psi|v|\psi\psi_{\mathbf{k}n}\rangle + \langle \psi_{\mathbf{k}n}\psi|vPv|\psi\psi_{\mathbf{k}n}\rangle. \end{aligned} \tag{13}$$

The first term is the bare exchange and the second term is the correlated part. The extra ψ consisting of core, valence, and conduction states comes from the Green function G,

$$\{\psi\} \equiv \{\psi_{\text{core}}\} + \{\psi_{\text{val}}\} + \{\psi_{\text{cond}}\}. \tag{14}$$

For the exchange term, the required Hilbert space is

$$\{\Sigma^x\} = \{\psi_{\mathbf{k}n}(\psi_{\text{core}} + \psi_{\text{val}})\}, \tag{15}$$

since the exchange part only involves summation over occupied states. In principle we need only the overlap between the above space and $\{v\}$, but it is unlikely that there exists a function $f \in \{\Sigma^x\}$, such that $\langle f|v|f\rangle = 0$. For the correlated part, the Hilbert space is

$$\{\Sigma^c\} = \{\psi_{\mathbf{k}n}\psi\} \cup \{P^0\}, \tag{16}$$

which also covers $\{\Sigma^x\}\cdot\{v\}$ alone is also possible, but this is probably more than is necessary.

Since it is unlikely that one can construct a small but complete basis for expanding $v = 1/|\mathbf{r}-\mathbf{r}'|$ in a separable form as in Eq. (5), we proceed by considering the space in Eq. (16). From Eq. (1) we have

$$\begin{aligned}\{P^0\} &= \{\psi^{\text{occ}}\psi^{\text{unocc}}\} \\ &= \{\psi_{\text{core}}\psi_{\text{cond}}\} + \{\psi_{\text{val}}\psi_{\text{cond}}\}. \end{aligned} \tag{17}$$

The first term is only necessary when core polarization is important. In GW calculations, one is usually interested in the valence states and a few of the low lying conduction states. Let us assume that $\{\psi_{\mathbf{k}n}\} = \{\psi_{\text{val}}\}$, keeping in mind that it might include some conduction states. Then from Eqs. (16) and (17), the total Hilbert space required is

$$\{\Sigma\} = \{(\psi_{\text{core}} + \psi_{\text{val}})(\psi_{\text{val}} + \psi_{\text{cond}})\}. \tag{18}$$

For practical purpose, the above equation is not very useful because $\{\psi\}$ depends on $\mathbf{k}$. In the LMTO method, however, the basis functions consist of ϕ and $\dot{\phi}$, which are $\mathbf{k}$ independent. Schematically, from Eqs. (6) and (7), we have

$$\psi\psi = [\phi\phi + (\phi\dot{\phi} + \dot{\phi}\phi)h + \dot{\phi}\dot{\phi}h^2]b^2. \tag{19}$$

Since h and b are just coefficients, we may rewrite Eq. (19) as

$$\{\Sigma\} = \{\phi\phi\} + \{\phi\dot{\phi}\} + \{\dot{\phi}\dot{\phi}\}. \tag{20}$$

This equation displays very clearly the Hilbert space spanned by $\{\Sigma\}$. If we include all the possible products of $\phi\phi$, $\phi\dot{\phi}$, and $\dot{\phi}\dot{\phi}$, then we have a complete basis *by construction*. The number of products in Eq. (20) can still be large. With nine *spd* orbitals, excluding core states, the total number of product functions is $3\times 9\times(9+1)/2 = 135$. With 16 *spdf* orbitals we have $3\times 16\times(16+1)/2 = 408$.

We reduce the number of product functions in three steps.

(1) We neglect terms containing $\dot{\phi}$, since they are small ($\dot{\phi}$ is typically 10% of ϕ). This reduces the number of product functions by a factor of 3.

(2) In Eq. (18), there are no products between conduction states. Therefore, in *sp* systems, products of $\phi_d\phi_d$ can be neglected and similarly in *d* systems, products of $\phi_f\phi_f$ may be neglected.

(3) The remaining product functions turn out to have a large number of linear dependencies, typically 30–50 %. These linear dependencies can be eliminated systematically, a process which is described below.

In general, a product function is defined by

$$\tilde{B}_i(\mathbf{r}) \equiv \chi_{\mathbf{R}L\nu}(\mathbf{r})\chi_{\mathbf{R}'L'\nu'}(\mathbf{r}) \ ; \ i = (\mathbf{R}L\nu, \mathbf{R}'L'\nu'). \tag{21}$$

With the simplification (1), the product function be-

comes

$$\tilde{B}_i(\mathbf{r}) \equiv \phi_{\mathbf{R}L\nu}(\mathbf{r})\phi_{\mathbf{R}L'\nu'}(\mathbf{r}) = \phi_{\mathbf{R}l\nu}(r)\phi_{\mathbf{R}l'\nu'}(r)y_L(\hat{\mathbf{r}})y_{L'}(\hat{\mathbf{r}}), \quad (22)$$

where $i = (\mathbf{R}, L\nu, L'\nu')$. This function is only nonzero inside a sphere centered on atom $\mathbf{R}$. Thus, there are no products between orbitals centered on different spheres. For a periodic system we need a Bloch basis and perform a Bloch sum

$$\tilde{B}_{\mathbf{k}i}(\mathbf{r}) = \sum_{\mathbf{T}} e^{i\mathbf{k}\cdot\mathbf{T}}\phi_{\mathbf{R}L\nu}(\mathbf{r}-\mathbf{R}-\mathbf{T})\phi_{\mathbf{R}L'\nu'}(\mathbf{r}-\mathbf{R}-\mathbf{T}). \quad (23)$$

The $\mathbf{k}$ dependence is in some sense artificial because the function has no overlap with neighboring spheres, similar to core states. After leaving out unnecessary products [step (2)], we optimize the basis by eliminating linear dependencies [step (3)]. This is done by orthogonalizing the overlap matrix

$$O_{ij} = \langle \tilde{B}_i | \tilde{B}_j \rangle, \quad (24)$$

$$Oz = ez, \quad (25)$$

and neglecting eigenvectors z_i with eigenvalues $\epsilon_i <$ tolerance. The resulting orthonormal basis is a linear combination of the product functions:

$$B_i = \sum_j \tilde{B}_j z_{ji}, \quad (26)$$

and typically we have ~70–100 functions per atom with *spdf* orbitals. The above procedure ensures that we have the smallest number of basis functions. Further approximations may be introduced to reduce the basis.

In Table I we show a completeness test for the basis. The slight discrepancy for high lying states is due to the neglect of $\dot{\phi}$'s in the product basis, which become more important for the broad high lying conduction states, and also because the optimization procedure puts less weight on those products that have smaller overlap. This is not crucial for two reasons: the matrix elements become smaller for the higher states and, in relation to *GW* calculations, there is a factor of $1/\omega$ that makes the higher states less important.

TABLE I. The completeness test of the optimal product basis for nickel. A product of two wave functions is expanded in the basis: $\psi^*_{\mathbf{k}n}\psi_{\mathbf{k}'n'} = \sum_i B_i c_i$ with $\mathbf{k} = (0\ 0\ 0)$, $\epsilon_{\mathbf{k}n} = -1.22$ eV (the highest valence state at the Γ point) and $\mathbf{k}' = (0.5\ 0.5\ 0.5)$. The basis is complete if $\sum_i |c_i|^2 = |\langle\psi_{\mathbf{k}n}|\psi_{\mathbf{k}'n'}\rangle|^2$. The number of optimal product basis functions is 101 and 82 with and without $3s, 3p$ core states, respectively.

Core	$\lvert\langle\psi_{\mathbf{k}n}\rvert\psi_{\text{core}}\rangle\rvert^2$	$\sum_i \lvert c_i\rvert^2$	Error
3s	0.158 114	0.158 113	0.000 001
3p	0.066 833	0.066 832	0.000 001
3p	0.209 174	0.209 172	0.000 002
3p	0.168 184	0.168 183	0.000 001

$\epsilon_{\mathbf{k}'n'}$(eV)	$\lvert\langle\psi_{\mathbf{k}n}\rvert\psi_{\mathbf{k}'n'}\rangle\rvert^2$	$\sum_i \lvert c_i\rvert^2$	Error
-9.09	0.052 222	0.052 220	0.000 002
-2.23	0.181 724	0.181 722	0.000 002
-2.23	0.101 124	0.101 124	0.000 000
-2.23	0.167 753	0.167 743	0.000 010
-1.22	0.119 586	0.119 583	0.000 003
-1.22	0.014 565	0.014 537	0.000 028
24.39	0.060 638	0.060 634	0.000 004
28.19	0.013 188	0.013 074	0.000 114
28.19	0.008 852	0.008 736	0.000 116
28.19	0.001 883	0.001 738	0.000 145
42.29	0.015 563	0.015 400	0.000 163
42.29	0.018 353	0.018 167	0.000 186
42.29	0.006 292	0.006 040	0.000 252
73.89	0.017 300	0.016 660	0.000 640
73.89	0.011 847	0.011 409	0.000 438
73.89	0.013 877	0.012 911	0.000 966

III. APPLICATIONS: ELECTRON ENERGY LOSS SPECTRA OF NI AND SI

The electron energy-loss spectra measure the energy loss suffered by fast electrons when going through a thin film by creating single-particle-like excitations and collective excitations, mainly plasmons. It is given by the imaginary part of the inverse dielectric function[16]

$$\text{Im}\,\epsilon^{-1}(\mathbf{q},\omega) = \text{Im}\,\frac{1}{V}\int d^3r d^3r' e^{-i\mathbf{q}\cdot\mathbf{r}}\epsilon^{-1}(\mathbf{r},\mathbf{r}';\omega)e^{i\mathbf{q}\cdot\mathbf{r}'}. \quad (27)$$

In most calculations, $\epsilon^{-1}(\mathbf{q},\omega)$ is approximated by $1/\epsilon(\mathbf{q},\omega)$, which corresponds to neglecting the local-field arising from the nonhomogeneity of the induced charge

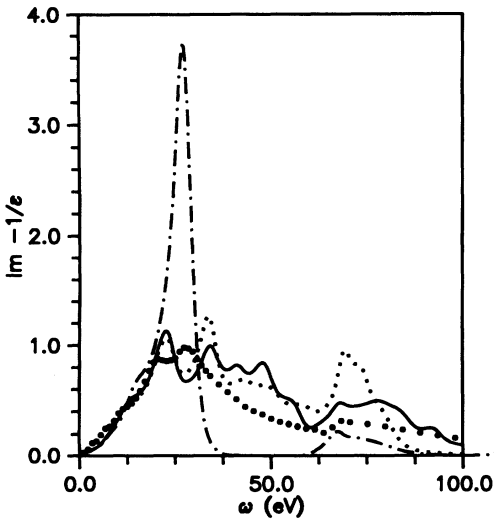

FIG. 1. Loss spectra of ferromagnetic Ni obtained with increasing number of LMTO basis: with $4s$, $4p$, $3d$ orbitals (dash-dotted line); $4s$, $4p$, $3d$, $4f$ orbitals (dotted line); $4s$, $4p$, $3d$, $4f$, $5g$ orbitals (solid line). All cases include core excitations. The solid circles are the experimental results.

density. A number of calculations have included the local field effects in semiconductors[20,21,13] and one calculation for a transition metal (Ni).[14] Most recently, local-field effects have also been considered in the calculations of plasmon dispersions in light alkali metals[22] and heavy alkali metals.[17]

To calculate ϵ^{-1} we first calculate P^0 in the product basis:

$$P^0_{ij}(\mathbf{q},\omega) = \sum_{\mathbf{k}} \sum_{n}^{\text{occ}} \sum_{n'}^{\text{unocc}} \langle B_{\mathbf{q}i}\psi_{\mathbf{k}n}|\psi_{\mathbf{k}+\mathbf{q}n'}\rangle\langle\psi_{\mathbf{k}+\mathbf{q}n'}|\psi_{\mathbf{k}n}B_{\mathbf{q}j}\rangle \left\{ \frac{1}{\omega - \epsilon_{\mathbf{k}+\mathbf{q}n'} + \epsilon_{\mathbf{k}n} + i\delta} - \frac{1}{\omega + \epsilon_{\mathbf{k}+\mathbf{q}n'} - \epsilon_{\mathbf{k}n} - i\delta} \right\}. \quad (28)$$

We must also calculate the Coulomb potential v in the product basis, which is described in the Appendix. We then use Eqs. (3) and (4) to obtain P and ϵ^{-1}.

To calculate the energy-loss spectra, however, it is more efficient to calculate S^0 in Eq. (2) than to calculate P^0 directly. In the actual calculations, the δ function is replaced by a Gaussian

$$\delta(\omega) \to \frac{1}{\sigma\sqrt{\pi}} e^{-\omega^2/\sigma^2}, \quad (29)$$

with $\sigma = 0.2$ Ry. The energy-loss spectra including local field is given by

$$\text{Im}\ \epsilon^{-1}(\mathbf{q},\omega) = \text{Im} \sum_{ij} \langle \mathbf{q}|B_{\mathbf{q}i}\rangle, \epsilon^{-1}_{ij}(\mathbf{q},\omega)\langle B_{\mathbf{q}j}|\mathbf{q}\rangle \quad (30)$$

where $\langle \mathbf{q}|B_{\mathbf{q}i}\rangle$ is the Fourier transform of the product basis and $\epsilon^{-1}_{ij}(\mathbf{q},\omega)$ is the matrix element of ϵ^{-1} in the product basis.

We have made a systematic study of Ni energy-loss spectra concerning the importance of high lying states and the sensitivity with respect to the quality of the transition matrix elements. Our starting Hamiltonian for the evaluation of P^0 is the local density Hamiltonian[23]

$$H^0 = T + V_H + V_{xc}. \quad (31)$$

We have calculated all spectra with and without local field for $\mathbf{q} = (0.25,\ 0,\ 0)2\pi/a$, $a = 6.654\ a_0$.

The loss spectra obtained with different sets of LMTO orbitals are shown in Fig. 1. The double peak at around 25 eV corresponds most likely to plasmons. The sharp edge at 67 eV is due to the onset of $3p$ core excitations. Inclusion of the f orbitals is essential because the transition probability from the d to f is large, which is reflected in the absence of intensity between 40 and 60 eV when the f orbitals are neglected. Inclusion of the g orbitals improves the spectra above 60 eV but there are still some remaining discrepancies that suggest that the band structure and the transition matrix elements are not accurate enough. Indeed, we have shown in a previous paper[19] that *spdf* or even *spdfg* orbitals are not sufficient for a proper description of the high lying conduction states. Inclusion of the second d orbitals, i.e., $4d$ orbitals, in the basis functions is crucial and they were found to change the band structure above 30 eV dramatically.

In Fig. 2 we show the effects of adding the $4d$ orbitals to the LMTO basis. The main improvements are the shifting of the second main peak by 3 eV to lower energy and the reduction in intensity in the region between 40

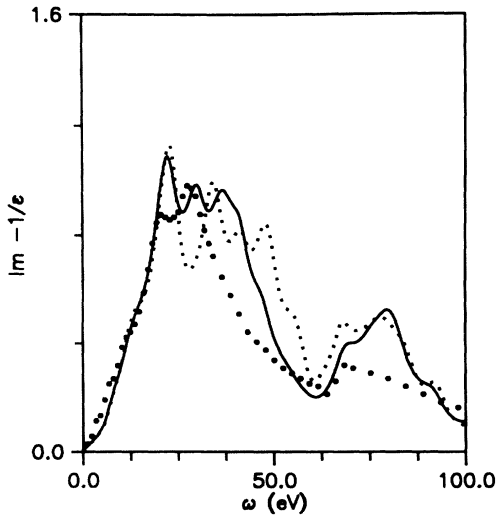

FIG. 2. The effect of including the $4d$ orbitals on the loss spectra of Ni: with $4s$, $4p$, $3d$, $4f$, $5g$ orbitals (dotted line) and with $4s$, $4p$, $3d$, $4d$, $4f$, $5g$ orbitals (solid line). Both cases include core excitations. The solid circles are the experimental result.

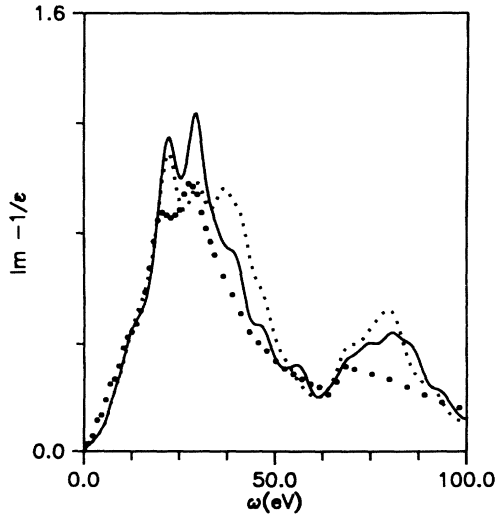

FIG. 3. The effect of including an empty sphere on the loss spectra of Ni. The empty sphere is located at $(0.5\ 0.5\ 0.5)a$. The solid and dotted lines are the loss spectra obtained with and without an empty spheres, respectively. Both spectra are calculated with $4s$, $4p$, $3d$, $4d$, $4f$, $5g$ orbitals including core excitations. The solid circles are the experimental result.

and 60 eV. The large peak at the onset of the $3p$ core excitations is shifted to higher energy, but the magnitude is still too large.

To further improve the band structure and the transition matrix elements, we add empty spheres. Although they have little effect on the valence states, we found significant effects on the high lying conduction states. The result is shown in Fig. 3. A good agreement with experiment is obtained; in particular, the positions of the two main peaks are very close to the experimental peaks, with a 1 eV shift to higher energy. The magnitude is somewhat too large, because it is sensitive to the choice of the broadening used in the calculations, as may be seen in Fig. 4 where the spectra are calculated with $\sigma = 0.1$ Ry and $\sigma = 0.2$ Ry. Increasing the broadening will reduce the heights of the peaks further, but the structure in the rest of the spectra is lost.

It is interesting to note that although the f-sum rule is off by almost a factor of two when the f orbitals are neglected, the static sum rule is given correctly. This is due to the fact that the static sum rule is a ground-state property. It has an important implication in the GW calculations because the self-energy is given by an integral of the form $\int d\omega' W(\omega + \omega')/\omega'$.

In our calculations, the local-field effects are found to be small, except for energies above 50 eV, as may be seen in Fig. 5. This result is surprising because the charge density in Ni is very far from homogeneous and one expects that the corresponding screened potential would be highly inhomogeneous, resulting in a large local-field correction. Since P^0 itself must have high Fourier components, the result suggests that the screened potential does not have significantly high Fourier components. Since the objective of this paper is to describe the method of calculating dielectric matrices, we present a detailed investigation of the local-field effects in a forthcoming paper.[24]

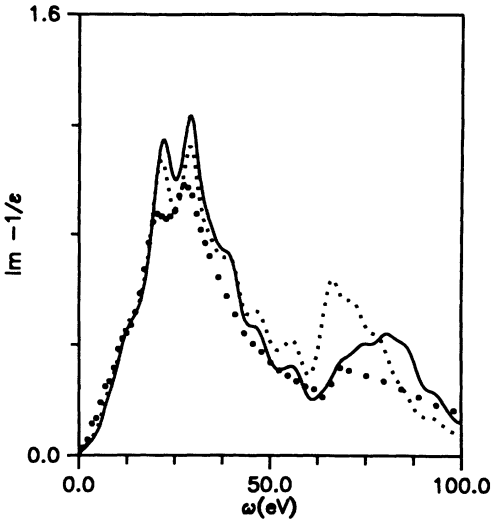

FIG. 5. The effect of local field on the loss spectra of Ni. The solid and dotted lines are the loss spectra with and without local field corrections, respectively. Both spectra are calculated with $4s$, $4p$, $3d$, $4d$, $4f$, $5g$ orbitals, including an empty sphere at $(0.5\ 0.5\ 0.5)a$ and core excitations.

We compare in Fig. 6 our present results with a previous work which used a modified LAPW method for the band structure and the basis functions. The second main peak in the previous calculation is somewhat larger than the present one but, on the other hand, the intensity at high energies is better described than in the present calculations. As shown in Fig. 4, the magnitude of the peaks can be sensitive to the choice of broadening. The

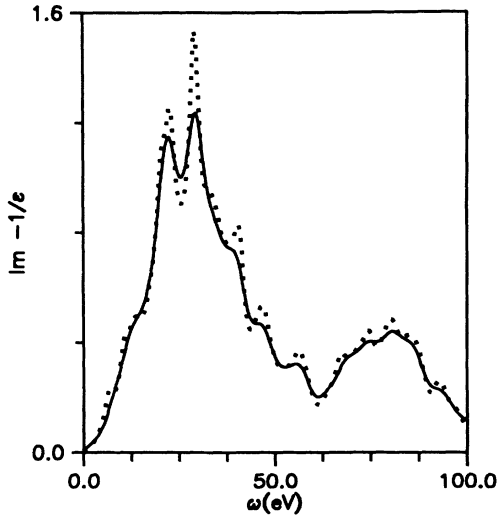

FIG. 4. The effect of Gaussian broadening on the loss spectra of Ni. The solid and dotted lines are the loss spectra obtained with $\sigma = 0.2$ Ry and $\sigma = 0.1$ Ry, respectively. Both spectra are calculated with $4s$, $4p$, $3d$, $4d$, $4f$, $5g$ orbitals, including an empty sphere at $(0.5\ 0.5\ 0.5)a$ and core excitations. The Gaussian broadening $\sigma = 0.2$ Ry.

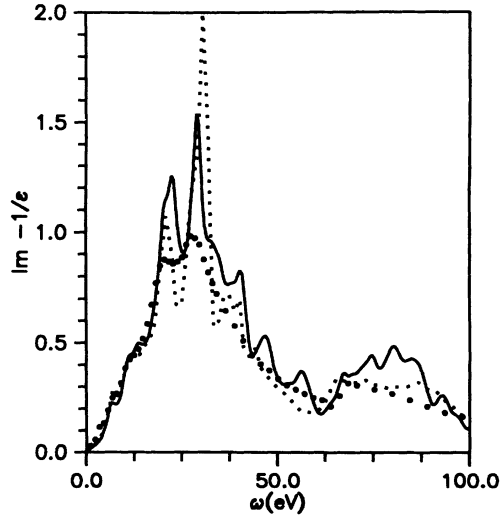

FIG. 6. Comparison between the loss spectra of Ni obtained with a modified LAPW method[14] (dotted line) and the present method (solid line). Both spectra are calculated with $\sigma = 0.1$ Ry and the solid curve is calculated with $4s$, $4p$, $3d$, $4d$, $4f$, $5g$ orbitals, including an empty sphere at $(0.5\ 0.5\ 0.5)a$. Both cases include core excitations.

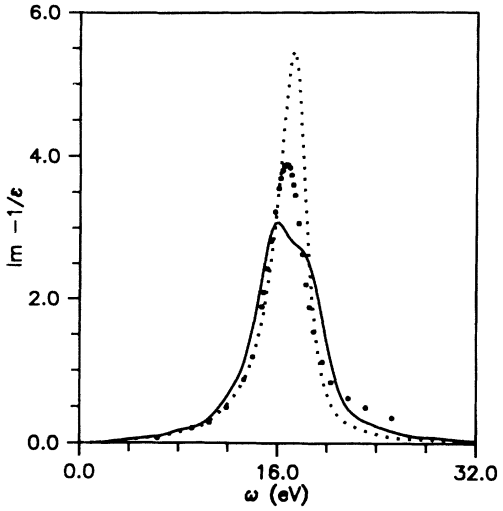

FIG. 7. The loss spectra of Si with local field (solid line) and without local field (dotted line), calculated with $3s$, $3p$, $3d$, $4f$ orbitals and with a Gaussian broadening $\sigma = 0.1$ Ry. Core excitations are included. The solid circles are the experimental results.

agreement is nevertheless satisfactory considering that the two methods are very different.

To show that our method also works for sp systems, we have calculated the energy-loss spectra of Si which have been the subject of several investigations.[13,20,21] The result is shown in Fig. 7. A large reduction in the magnitude of the plasmon peak due to local-field effects is observed, which has also been reported elsewhere. The calculated spectra correspond to $\mathbf{k} = (0.25\ 0\ 0)2\pi/a = 0.3$ Å^{-1}, which is larger than the wave vector in the experiment (0.067 Å^{-1}). This explains the smaller magnitude of the calculated plasmon peak, which results in a larger width. Another interesting feature is the splitting of the plasmon peak, as also found by Engel and Farid,[13] whose origin may be assigned to local-field effects. However, the separation between the two peaks is quite large (2.5 eV), significantly larger than the experimental resolution ($\sim$ 0.8 eV), and therefore it is not likely that the experiment has missed this splitting, unless there are other broadening effects such as those due to impurities. On the other hand, the peak structure can be sensitive to the quality of the wave functions.

IV. SUMMARY AND CONCLUSIONS

We have developed a method for calculating the dielectric matrix of a periodic system using a product basis. The method is applicable to any system, sp as well as narrow band systems, as the examples on energy-loss spectra have shown. Our method has several advantages over methods based on plane waves. The main advantages are that there is a small number of basis functions, core polarization can be easily incorporated, and there is no restriction to any particular system. In the Ni case, for instance, there is no difficuly in including all the core states and the computational effort is about the same as for Si. This is not the case in methods based on a plane-wave basis. A very large number of plane waves would be needed to describe the rather localized d states, not to mention the core states, which makes the computation very expensive, if at all possible.

The relatively small number of basis functions allows us to perform GW calculations without making a plasmon-pole approximation. Nevertheless, it is still very demanding computationally and an efficient way of calculating P^0 is needed without sacrificing too much accuracy. It is one of the objectives of the present work to find such an efficient scheme, but in order to have a controlled scheme, one must be able to do the full calculations for the purpose of comparison.

As applications of the method, we have calculated the energy-loss spectra of Ni and Si. A good agreement with experimental results is found in both cases. We have done a systematic study regarding the importance of high lying states and the quality of the transition matrix elements in Ni. It was found that inclusion of the $4d$ orbitals is important for a good description of the energy-loss spectra. This is consistent with the band-structure calculations described in a previous work,[19] where it was found that the $4d$ orbitals dramatically changed the band structure above 30 eV.

Local-field effects are found to be small in Ni but large in Si. The origin of local-field effects is an interesting subject, but has not been studied in detail so far. In a forthcoming paper,[24] we investigate in detail the origin of the local-field effects and offer an explanation of why the local-field effects in Ni are small whereas in Si they are large.

ACKNOWLEDGEMENTS

We would like to thank O. K Andersen for stimulating discussions and K. Karlsson for a critical reading of the manuscript. This work has been supported in part by the European Community program Human Capital and Mobility through Contract No. CHR-CT93-0337.

APPENDIX: EVALUATION OF THE COULOMB MATRIX

We consider one atom per unit cell for simplicity. Extension to several atoms is straightforward. The Coulomb matrix is given by

$$v_{ij}(\mathbf{q}) = \frac{1}{N}\int d^3r \int d^3r' \frac{B^*_{\mathbf{q}i}(\mathbf{r})B_{\mathbf{q}j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \qquad \text{(A1)}$$

where $B_{\mathbf{q}i}$ is normalized to unity in the unit cell. The integrations over the whole space may be reduced to integrations over a unit cell Ω by using the property

$$B_{\mathbf{q}i}(\mathbf{r}+\mathbf{T}) = e^{i\mathbf{q}\cdot\mathbf{T}}B_{\mathbf{q}i}(\mathbf{r}) \qquad \text{(A2)}$$

and noting that the integration over $\mathbf{r}'$ is independent of the origin of $\mathbf{r}$. This gives

$$v_{ij}(\mathbf{q}) = \int_\Omega d^3r \int_\Omega d^3s B^*_{\mathbf{q}i}(\mathbf{s}) E_{\mathbf{q}}(\mathbf{s},\mathbf{r}) B_{\mathbf{q}j}(\mathbf{r}), \quad \text{(A3)}$$

where

$$E_{\mathbf{q}}(\mathbf{s},\mathbf{r}) = \sum_{\mathbf{T}} \frac{e^{i\mathbf{q}\cdot\mathbf{T}}}{|\mathbf{s}-\mathbf{r}-\mathbf{T}|} = \frac{4\pi}{\Omega}\sum_{\mathbf{G}} \frac{e^{-(\mathbf{q}+\mathbf{G})^2/4\alpha^2}}{(\mathbf{q}+\mathbf{G})^2} e^{i(\mathbf{q}+\mathbf{G})\cdot(\mathbf{s}-\mathbf{r})} + \alpha \sum_{\mathbf{T}} e^{i\mathbf{q}\cdot\mathbf{T}} \frac{\mathrm{erfc}(\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|)}{\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|}. \quad \text{(A4)}$$

The Ewald method has been used to obtain the above decomposition into summations in the reciprocal and real space. erfc is the complementary error function equal to (1−erf), and α is an arbitrary constant whose value is chosen to give a fast convergence in the number of reciprocal lattice vectors and the number of neighbors. The essence of the Ewald method is to add and substract a Gaussian charge distribution which breaks the Coulomb potential from a point charge into a short and long range part. The short range part is done in real space and the long range part is done in reciprocal space. The main task is to calculate the potential

$$P_{\mathbf{q}j}(\mathbf{s}) = \sum_i \tilde{P}_{\mathbf{q}i}(\mathbf{s}) z_{ij}, \quad \text{(A5)}$$

with z given by Eq. (26) and

$$\tilde{P}_{\mathbf{q}i}(\mathbf{s}) = \int_\Omega d^3r \tilde{B}_{\mathbf{q}i}(\mathbf{r}) E_{\mathbf{q}}(\mathbf{s},\mathbf{r}) = \sum_{\mathbf{G}} \tilde{P}_{\mathbf{q}i}(\mathbf{s},\mathbf{G}) + \sum_{\mathbf{T}} \tilde{P}_{\mathbf{q}i}(\mathbf{s},\mathbf{T}), \quad \text{(A6)}$$

where

$$\tilde{P}_{\mathbf{q}i}(\mathbf{s},\mathbf{G}) = \frac{4\pi}{\Omega} \frac{e^{-(\mathbf{q}+\mathbf{G})^2/4\alpha^2}}{(\mathbf{q}+\mathbf{G})^2} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{s}} \times \int_\Omega d^3r e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} \tilde{B}_{\mathbf{q}i}(\mathbf{r}) \quad \text{(A7)}$$

and

$$\tilde{P}_{\mathbf{q}i}(\mathbf{s},\mathbf{T}) = \alpha e^{i\mathbf{q}\cdot\mathbf{T}} \int_\Omega d^3r \frac{\mathrm{erfc}(\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|)}{\alpha|\mathbf{s}-\mathbf{r}-\mathbf{T}|} \tilde{B}_{\mathbf{q}i}(\mathbf{r}). \quad \text{(A8)}$$

It is straightforward to calculate $\tilde{P}_{\mathbf{q}j}(\mathbf{s},\mathbf{G})$, since it is a Fourier transform of $\tilde{B}_{\mathbf{q}j}(\mathbf{r})$. To calculate $\tilde{P}_{\mathbf{q}i}(\mathbf{s},\mathbf{T})$, we use the following expansion formulas

$$\frac{1}{|\mathbf{s}-\mathbf{r}|} = \sum_L \frac{4\pi}{2l+1} \frac{r_<^l}{r_>^{l+1}} y_L(\hat{\mathbf{s}}) y_L(\hat{\mathbf{r}}) \quad \text{(A9)}$$

and

$$\frac{\mathrm{erf}(\alpha|\mathbf{s}-\mathbf{r}|)}{\alpha|\mathbf{s}-\mathbf{r}|} = \sum_L \frac{4\pi}{2l+1} g_l(r,s) y_L(\hat{\mathbf{s}}) y_L(\hat{\mathbf{r}}). \quad \text{(A10)}$$

The coefficients $g_l(r,s)$ are determined by numerical integrations using special directions.[14]

At the central sphere, $\tilde{B}_{\mathbf{q}i}(\mathbf{r})$ has no $\mathbf{q}$ dependence and is given by Eq. (22), so that

$$\tilde{P}_{\mathbf{q}i}(\mathbf{s},\mathbf{T}) = \alpha e^{i\mathbf{q}\cdot\mathbf{T}} \sum_L w_{li}(s_{\mathbf{T}}) y_L(\hat{\mathbf{s}_{\mathbf{T}}}) \int d\Omega y_{L_1} y_{L_2} y_L \quad \text{(A11)}$$

where $\mathbf{s_T} = \mathbf{s} - \mathbf{T}$ and

$$w_{li}(s_{\mathbf{T}}) = \frac{4\pi}{2l+1} \int_0^R dr\, r^2 \left\{ \frac{r_<^l}{r_>^{l+1}} - g_l(r, s_{\mathbf{T}}) \right\} \tilde{B}_i(r). \quad \text{(A12)}$$

We note that the sum over L is cut off by L_1 and L_2 in the product function $\tilde{B}$. The final integration over $\mathbf{s}$ is easily done with special directions.[14]

[1] I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer-Verlag, Berlin, 1986).
[2] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).
[3] L. Hedin, Phys. Rev. **139**, A796 (1965).
[4] L. Hedin and S. Lundqvist, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1969), Vol. 23.
[5] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).
[6] M. S. Hybertsen and S. G. Louie, Comments Condens. Matter Phys. **13**, 223 (1987).
[7] R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. B **37**, 10 159 (1988).
[8] F. Aryasetiawan, Phys. Rev. B **46**, 13 051 (1992)
[9] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959); **116**, 880 (1959); see also V. Heine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970), Vol. 24.
[10] W. Hanke and L. J. Sham, Phys. Rev B **12**, 4501 (1975).
[11] W. von der Linden and P. Horsch, Phys. Rev. B **37**, 8351 (1988).
[12] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **35**, 5585 (1987).
[13] G. E. Engel and B. Farid, Phys. Rev. B **47**, 15 931 (1993).
[14] F. Aryasetiawan, U. von Barth, P. Blaha, and K. Schwarz (unpublished).
[15] See, e.g., A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
[16] D. Pines, *Elementary Excitations in Solids* (Benjamin, New York, 1963).
[17] F. Aryasetiawan and K. Karlsson (unpublished).
[18] L. A. Feldcamp, M. B. Stearns, and S. S. Shinozaki, Phys. Rev. B **20**, 1310 (1979).
[19] F. Aryasetiawan and O. Gunnarsson, Phys. Rev. B **49**,

7219 (1994).
[20] S. G. Louie, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. Lett. **34**, 155 (1975).
[21] R. Daling, W. van Haeringen, and B. Farid, Phys. Rev. B **45**, 8970 (1992).
[22] A. A. Quong and A. G. Eguiluz, Phys. Rev. Lett. **70**, 3955 (1993).
[23] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, *ibid.* **140**, A1133 (1965).
[24] F. Aryasetiawan and O. Gunnarsson (unpublished).