Application of a general self-consistency scheme in the linear combination of atomic orbitals formalism to the electronic and structural properties of Si and W

C. T. Chan, David Vanderbilt,* and Steven G. Louie

Department of Physics, University of California, Berkeley, California 94720
and Materials and Molecular Research Division, Lawrence Berkeley Laboratory, University of California,
Berkeley, California 94720

We present a general self-consistency procedure formulated in momentum space for electronic structure and total-energy calculations of crystalline solids. It is shown that both the charge density and the change in the Hamiltonian matrix elements in each iteration can be calculated in a straightforward fashion once a set of overlap matrices is computed. The present formulation has the merit of bringing the self-consistency problem for different basis sets to the same footing. The scheme is used to extend a first-principles pseudopotential linear combination of Gaussian orbitals method to full point-by-point self-consistency, without refitting of potentials. It is shown that the set of overlap matrices can be calculated very efficiently if we exploit the translational and space-group symmetries of the system under consideration. This scheme has been applied to study the structural and electronic properties of Si and W, prototypical systems of very different bonding properties. The results agree well with experiment and other calculations. The fully self-consistent results are compared with those obtained by a variational procedure [J. R. Chelikowsky and S. G. Louie, Phys. Rev. B 29, 3470 (1984)]. We find that the structural properties for bulk Si and W (both systems have no interatomic charge transfer) can be treated accurately by the variational procedure. However, full self-consistency is needed for an accurate description of the band energies.

I. INTRODUCTION

The local-density approximation (LDA), in the density-functional formalism,[1] has been very successful in treating the electronic and structural properties of a wide variety of solid-state systems. Within the LDA, there are several popular approaches in solving the effective one-particle Kohn-Sham equations. Examples are the plane-wave (PW), the linearized augmented plane-wave (LAPW), the linear muffin-tin orbital (LMTO), the linear combination of atomic orbitals (LCAO), and the mixed-basis methods. These approaches mainly differ in the basis set chosen to expand the crystalline wave functions. The efficiency of each method is dependent on the systems under consideration. For example, the PW method, together with the *ab initio* pseudopotential approach, has been proven very successful with simple metals and semiconductors. Methods like APW, LAPW, and LMTO are more frequently applied to transition elements and systems with closed packed structures. Another commonly used approach is the LCAO method which has the feature of using a basis set that is "physical". Thus usually only a relatively small number of basis functions (per atom) are needed for the wave function. This yields a small Hamiltonian matrix, a crucial factor in electronic structure calculations for complex structure materials. The method is versatile and can be applied to a large variety of systems. Ever since Lafon *et al.*[2] and Langlinais and Callaway[3] formulated the LCAO method in a first-principles fashion, the method has become an increasingly popular computational tool.

However, the LCAO method, as with the other methods, is not without difficulties. One common difficulty is that iteration to self-consistency in the screening potential is a tedious procedure. Calculation of the electronic charge density in real space (which is the central quantity in the density-functional formalism) is usually computationally extensive as the charge density has to be constructed from the eigenfunctions in every iteration. Moreover, to make the calculations tractable, the charge densities or the potentials are usually fitted to Gaussians in every iteration. An accurate fit is in general not easy. Despite the fact that various techniques[4] have been developed, the final result is inevitably influenced by the quality of the fit. Unless great care is taken, fitting the charge densities or potentials to Gaussians limits the accuracy of the calculation because an artificial form restriction is imposed. Some of these difficulties have been circumvented by Smith, Gay, and Arlinghaus[5] in their self-consistent local orbitals (SCLO) scheme in treating the electronic structure of transition metal surfaces. There are no explicit constraints on the shape of the final self-consistent potential in their method.

Recently, Chelikowsky and Louie[6] have developed an *ab initio* pseudopotential LCAO method, which has been applied to several prototypical systems, mainly involving the *sp* electrons. The present paper is to present a scheme that will extend this method to full point-by-point self-consistency. The approach has the following desirable features.

(i) The calculation of the charge density in real or Fourier space is rather straightforward.

(ii) No fitting of the screening potential or charge is needed during the iteration process.

(iii) The crystal potential has no shape constraint.

(iv) There is no need to recompute fully the Hamiltonian matrix elements in every iteration. Only the changes in matrix elements as a result of the changing potential are computed and they can be obtained without evaluation of any three-center integrals.

The present scheme should be applicable to the self-consistency iteration for an arbitrary chosen basis set. It thus has the merit of bringing the self-consistency problem for different basis sets to the same footing. In the specific application to the LCAO basis set, the scheme in part resembles that of Smith, Gay, and Arlinghaus.[5] The key feature of both methods is the introduction of a set of overlap matrices (which have been defined in various forms before; see, for example, Ref. 3). There are, however, differences. In our method the overlap matrices are used to calculate both the charge density and the change of Hamiltonian matrix elements while they are used only for the latter by Smith *et al.* The major problem of this formulation is that the computation of these matrices are not easy in general. We will show that for many systems, both surface and bulk, the evaluation of these matrices can be greatly simplified by a factorization scheme. The method can treat both bulk and surface problems (using a supercell) and can calculate electronic structure as well as total energy.

The remainder of the paper is organized in the following way. The general formalism is presented in Sec. II.

Section III presents examples to which this method is employed. Section IV is a discussion and summary.

II. FORMALISM

A. Self-consistency scheme

Within the framework of the density-functional theory,[1] the basis problem in electronic structure calculations is to solve the set of Kohn-Sham equations self-consistently:

$$(-\tfrac{1}{2}\nabla^2 + V_{\text{ion}} + V_H + V_{\text{xc}})\psi_{n\mathbf{k}} = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}} \qquad (1)$$

and

$$\rho(\mathbf{r}) = \sum_{n,\mathbf{k}}^{\text{occ}} |\psi_{n\mathbf{k}}(\mathbf{r})|^2 , \qquad (2)$$

where V_{ion} is the external potential due to the ions, V_H is the Hartree potential, V_{xc} is the exchange-correlation potential, n and $\mathbf{k}$ are band and k-point indices, respectively. In almost all total-energy calculations, the local-density approximation is used for V_{xc}. Various functional forms of local V_{xc} have been proposed,[7,8] and they have been shown to yield similar results for electronic properties.

Once (1) and (2) are solved self-consistently, the total energy of the system can be determined using

$$E_{\text{total}} = \sum_{n,\mathbf{k}} \epsilon_{n\mathbf{k}} - \int V_H(\rho_{\text{in}})\rho_{\text{out}}\, d^3r + \tfrac{1}{2}\int V_H(\rho_{\text{out}})\rho_{\text{out}}\, d^3r - \int V_{\text{xc}}(\rho_{\text{in}})\rho_{\text{out}}\, d^3r + \int \epsilon_{\text{xc}}(\rho_{\text{out}})\rho_{\text{out}}\, d^3r + E_{\text{ion-ion}} , \qquad (3)$$

where ϵ_{xc} is the exchange-correlation energy density and $E_{\text{ion-ion}}$ is the electrostatic interaction energy among the bare ions.

Here we distinguish between input (ρ_{in}) and output (ρ_{out}) charge densities. They should in principle be equal if the charge density is fully self-consistent. However, in practice, a small (but controllable) difference is inevitable due to the fact that Eqs. (1) and (2) have to be solved iteratively. Formulating the total energy as in Eq. (3) guarantees that the energy is a functional of the output charge density only,[6] with no explicit dependence on the input charge density. This procedure gives better results since $E[\rho]$ is a variational minimum for the correct ρ, as emphasized in Ref. 6. It has been shown that in some cases,[6,9] self-consistency iterations can be avoided while state-of-the-art accuracy is obtained if Eq. (3) were used.

Nevertheless, there are many problems of interests, such as surfaces and chemisorption systems, in which the self-consistency of the electronic screening potential is important and demanding. It is thus desirable to have a scheme which can handle self-consistency accurately and efficiently. For the same system under consideration, the amount of labor needed in achieving self-consistency varies with the basis set used. As has been mentioned in the Introduction, achieving self-consistency in the LCAO approach is usually more tedious when compared with basis sets like plane waves. We show in this section that a

general self-consistency scheme can be formulated in which most of the computation labor (besides the diagonalization of the Hamiltonian matrices) can be transformed to the computation of a set of overlap matrices. The scheme is formally independent of the basis set chosen. It is only the evaluation of the overlap matrices themselves that depends on the basis set. It will be shown in the next section how these overlap matrices can be computed for specific different basis sets, in particular for the LCAO basis set.

For a crystalline solid with a set of reciprocal-lattice vectors $\{\mathbf{G}\}$ and a set of Bloch functions $\{\phi_i(\mathbf{k},\mathbf{r})\}$ chosen as the basis functions to expand the electronic wave function ψ:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_i c_i(n,\mathbf{k})\phi_i(\mathbf{k},\mathbf{r}) , \qquad (4)$$

where i is a general composite index, Eq. (1) becomes a generalized eigenvalue problem:

$$H\psi = ES\psi . \qquad (5)$$

The Hamiltonian matrix is given by

$$H_{ij}(\mathbf{k}) = \langle \mathbf{k},i \mid -\tfrac{1}{2}\nabla^2 + V_{\text{ion}} + V_H + V_{\text{xc}} \mid \mathbf{k},j \rangle , \qquad (5a)$$

and

$$S_{ij}(\mathbf{k}) = \langle \mathbf{k},i \mid \mathbf{k},j \rangle \qquad (5b)$$

is the usual overlap matrix. $|\mathbf{k},i\rangle$ here denotes $\phi_i(\mathbf{k},\mathbf{r})$. The solution of Eq. (5) gives the eigenvectors ψ, for each n and $\mathbf{k}$, in the basis space in the form of a column vector with the c_i's as components. (We assume that the basis set chosen is fixed and does not change with iterations.)

Most of the labor associated in achieving point-by-point self-consistency in the electronic screening potential can be transformed into the computation of a set of overlap matrices $\Pi_{ij}(\mathbf{k},\mathbf{G})$ which are defined by the equation

$$\Pi_{ij}(\mathbf{k},\mathbf{G})=\left\langle \mathbf{k},j \left| \frac{e^{-i\mathbf{G}\cdot\mathbf{r}}}{\Omega_c} \right| \mathbf{k},i \right\rangle , \qquad (6)$$

where Ω_c is the volume of the unit cell. The $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices are completely determined by the basis set and the reciprocal lattice. These matrices have the same dimension as the Hamiltonian matrix.

The total charge density $\rho(\mathbf{r})$ is

$$\rho(\mathbf{r})=\sum_{n,\mathbf{k}}\sum_{i,j}c_i(n,\mathbf{k})c_j^*(n,\mathbf{k})\phi_i(\mathbf{k},\mathbf{r})\phi_j^*(\mathbf{k},\mathbf{r}) , \qquad (7)$$

and its Fourier components, $\rho(\mathbf{G})$, defined by

$$\rho(\mathbf{r})=\sum_{\mathbf{G}}\rho(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{r}} , \qquad (8)$$

are then, using (6), given by

$$\rho(\mathbf{G})=\sum_{n,\mathbf{k}}\sum_{i,j}c_i(n,\mathbf{k})c_j^*(n,\mathbf{k})\Pi_{ij}(\mathbf{k},\mathbf{G}) , \qquad (9)$$

or formally in matrix notation as

$$\rho(\mathbf{G})=\sum_{n,\mathbf{k}}\psi_{n\mathbf{k}}^{\dagger}\Pi(\mathbf{k},\mathbf{G})\psi_{n\mathbf{k}} . \qquad (9a)$$

Note that in many electronic structure calculations, charge densities are computed in every iteration with the same amount of effort using (7). Now that if $\Pi_{ij}(\mathbf{k},\mathbf{G})$ have been computed and stored, the charge density can be simply evaluated using (9) for every iteration. For a LCAO-type basis set, the dimension of the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ is usually small and thus making the calculation of charge density particularly simple. In particular, if pseudopotentials are employed, the wave functions are nodeless. The charge density is smooth and can be represented very accurately by using a reasonably small cutoff in Fourier space.

After obtaining the charge density in Fourier space from Eq. (9), the electron screening potentials may be computed in the usual way. The Hartree potential is straightforwardly given by

$$V_H(\mathbf{G})=\frac{8\pi\rho(\mathbf{G})}{G^2} , \qquad (10)$$

and the exchange-correlation potential, in the LDA by

$$V_{xc}(\mathbf{r})=V_{xc}[\rho(\mathbf{r})] . \qquad (11)$$

The charge density in real space can be computed from $\rho(\mathbf{G})$ either directly from (8) or more effectively by using fast-Fourier-transform techniques. Likewise, V_{xc} can be transformed to $V_{xc}(\mathbf{G})$ by fast Fourier transforms.

Defining the difference between the input and output potentials as

$$\delta V=\delta V_{\text{output}}-\delta V_{\text{input}} , \qquad (12)$$

the change in Hamiltonian matrix elements,

$$\delta H_{ij}(\mathbf{k})=\langle \mathbf{k},i \,|\, \delta V \,|\, \mathbf{k},j \rangle , \qquad (13)$$

is then given by

$$\delta H_{ij}(\mathbf{k})=\sum_{\mathbf{G}}\delta V(\mathbf{G})\Pi_{ij}^*(\mathbf{k},\mathbf{G})\Omega_c . \qquad (14)$$

Hence the corrections to the Hamiltonian matrix for the next iteration can be expressed simply in terms of $\Pi_{ij}(\mathbf{k},\mathbf{G})$ and computed easily. Other than errors due to incompleteness of bases or a finite $\mathbf{G}$ space, the formulation here is exact within the LDA. There is no need to re-fit the charge density or the potential, and there is no form constraint on the screening potential, even if a LCAO-type basis set is used. As long as the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices can be calculated, the iterative procedure is greatly simplified. The iteration procedure is illustrated schematically in Fig. 1.

In practice, the δV in Eq. (12), defined as the difference between output and input potential, is damped to accelerate convergence by anticipating electron-screening effects. A popular method is to assume Thomas-Fermi–type screening; then

$$\delta V'(\mathbf{G})=\frac{\delta V(\mathbf{G})}{1+\lambda^2/G^2} , \qquad (15)$$

where λ is the Thomas-Fermi screening wave vector, and $\delta V'$ instead of δV is used in Eq. (14). More sophisticated approaches like Broyden's method[10] or the dielectric matrix screening scheme[11] can further accelerate convergence.
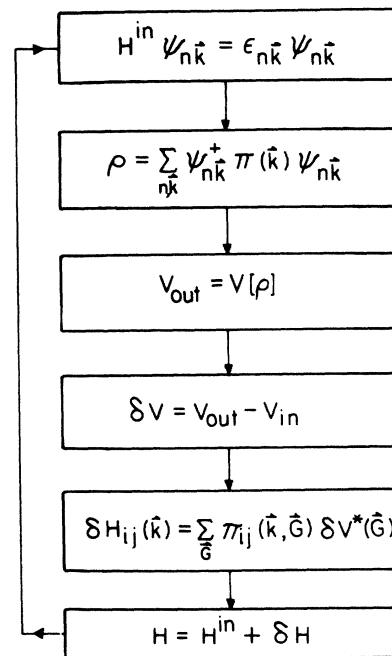

FIG. 1. A schematic chart of the self-consistency procedure using the π matrix.

B. Evaluation of Π_{ij} matrix

In this section, the general form of the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices will first be given. It will then be shown that for both plane-wave and LCAO (with Gaussian orbitals) basis sets, it reduces to simple analytic expressions and can be evaluated efficiently by exploiting the symmetry of the system.

We express an arbitrary basis set of Bloch functions in an expansion of plane waves,

$$\phi_i(\mathbf{k},\mathbf{r}) = \sum_{\mathbf{G}} \phi_i(\mathbf{k}+\mathbf{G}) \, | \, \mathbf{k}+\mathbf{G} \rangle . \tag{16}$$

Then the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices, as defined by Eq. (6), have the general form:

$$\Pi_{ij}(\mathbf{k},\mathbf{G}) = \frac{1}{\Omega_c} \sum_{\mathbf{G}'} \phi_i(\mathbf{k}+\mathbf{G}')\phi_j^*(\mathbf{k}+\mathbf{G}'-\mathbf{G}) . \tag{17}$$

For a plane-wave basis set, ϕ_i's are plane waves themselves and so the index i labels a $\mathbf{G}$ vector. The Fourier coefficients in Eq. (16) are just Kronecker deltas so that Eq. (17) reduces to

$$\Pi_{ij}(\mathbf{k},\mathbf{G}) = \Pi_{\mathbf{G}',\mathbf{G}''}(\mathbf{k},\mathbf{G}) = \delta_{\mathbf{G}'',\mathbf{G}'-\mathbf{G}} , \tag{18}$$

which is independent of $\mathbf{k}$ and trivially simple in this special case. The charge density and the change in matrix elements are just, using (18) in (9) and (14),

$$\rho(\mathbf{G}) = \sum_{n,\mathbf{k}} \sum_{\mathbf{G}'} \psi_{\mathbf{k}+\mathbf{G}'}(n,\mathbf{k})\psi_{\mathbf{k}+\mathbf{G}'-\mathbf{G}}^*(n,\mathbf{k}) , \tag{19}$$

and

$$\delta H_{\mathbf{G}\mathbf{G}'} = \delta V(\mathbf{G}-\mathbf{G}') . \tag{20}$$

The present formalism thus reduces Eqs. (9) and (14) to the usual expressions in the plane-wave representation.

For a LCAO basis set in general, we have a Bloch sum of atomiclike orbitals localized on the atomic sites of the form:

$$\phi_i(\mathbf{k},\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\tau_i)} f_i(\mathbf{r}-\mathbf{R}-\tau_i) , \tag{21}$$

where Ω is the crystal volume, $\mathbf{R}$ is a lattice vector, τ_i is a basis vector, and f_i is a localized function. Using (21), the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ in Eq. (6) becomes

$$\Pi_{ij}(\mathbf{k},\mathbf{G}) = \frac{e^{-i\mathbf{G}\cdot\tau_j}}{\Omega_c} \sum_{\mathbf{G}'} e^{-i\mathbf{G}'\cdot(\tau_i-\tau_j)} f_i(\mathbf{k}+\mathbf{G}')$$

$$\times f_j^*(\mathbf{k}+\mathbf{G}'-\mathbf{G}) , \tag{22}$$

where

$$f(\mathbf{k}+\mathbf{G}) = \frac{1}{\Omega_a} \int e^{-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} f(\mathbf{r}) d\mathbf{r}^3 \tag{23}$$

is the Fourier transform of the atomiclike orbitals. Ω_a is the atomic volume and Ω_c is the unit-cell volume.

If Gaussian orbitals are used for the radial part of the atomiclike functions, as is frequently the case, analytic expressions exist for $f_i(\mathbf{k}+\mathbf{G})$. For example, we can take $f(\mathbf{r})$ to be[6]

$$f_{alm}(\mathbf{r}) = A_{alm} e^{-\alpha r^2} K_{lm}(r,\theta,\phi) , \tag{24}$$

where A_{alm} are normalization constants. Suppose that the K_{lm}'s (up to $l=2$) have the form $\{1, x, y, z, x^2, y^2, z^2, xy, yz, zx\}$ and the function $f(\mathbf{r})$ is normalized according to

$$\int f(\mathbf{r})^2 d^3 r = 1 , \tag{24a}$$

then A_{alm} is given by

$$A_{alm} = \left[\left[\frac{\pi}{2\alpha} \right]^{-3/2} \prod_{p=1}^{3} \frac{(4\alpha)^{n_{x_p}}}{(2n_{x_p}-1)!!} \right]^{1/2} , \tag{24b}$$

where $n_{x_1}, n_{x_2}, n_{x_3}$ are the powers of x, y, and z, respectively, in the Kubic harmonics.

Writing out the composite indices i and j explicitly, $\Pi_{ij}(\mathbf{k},\mathbf{G})$ then has the form

$$\Pi_{ialm,j\beta l'm'}(\mathbf{k},\mathbf{G}) = A_{alm} A_{\beta l'm'}^* \frac{e^{-i\mathbf{G}\cdot\tau_j}}{\Omega_c} \sum_{\mathbf{G}'} e^{-i\mathbf{G}'\cdot(\tau_i-\tau_j)} \prod_{p=1}^{3} [I_{an_{x_p}}(k_{x_p}+G'_{x_p})I_{\beta n_{x_p}}^*(k_{x_p}+G'_{x_p}-G_{x_p})] , \tag{25}$$

where

$$I_{an}(g) = (2i\sqrt{\alpha})^{-n} H_n(g/2\sqrt{\alpha}) \left[\left[\frac{\pi}{\alpha} \right]^{1/2} e^{-g^2/4\alpha} \right] , \tag{25a}$$

with H_n being a Hermite polynomial of order n. We note that the i and j indices for $\Pi_{ij}(\mathbf{k},\mathbf{G})$ have been used as general indices, while in (25), i and j are specifically site indices, α and β are indices for Gaussian decay constants, and lm and $l'm'$ are angular momentum indices for the localized functions.

Great savings can be achieved if the system has a unit cell (*not* necessarily primitive) that is orthorhombic. Examples are sc, fcc, bcc, diamond, hcp structures and so forth. In that case, the summations over $\mathbf{G}'$ in Eq. (25) are separable into individual x,y,z components and thus reduce the computation efforts significantly.

We first consider the simplest case in which the primitive unit cell is orthorhombic. In this case, the reciprocal-lattice vectors $\mathbf{G}$ have the form $\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3$, where (i) $\mathbf{b}_1, \mathbf{b}_2$ and $\mathbf{b}_3$ are mutually orthogonal and (ii) m_1, m_2, and m_3 are arbitrary integers. It is easy to see then the sum over $\{\mathbf{G}'\}$ on the right-hand side of Eq. (25) factorizes into the form

$$\sum_{G} [\cdots] = \left[\sum_{G_x}(\cdots)\right]\left[\sum_{G_y}(\cdots)\right]\left[\sum_{G_z}(\cdots)\right] . \qquad (26)$$

We basically reduce a N^3 summation problem to a N^1 summation problem, a huge saving if the number of G vector is large.

For a crystal structure that has an unit cell orthorhombic but its primitive unit cell not (e.g., bcc and fcc), it is always possible to partition the $\{G\}$ into a finite (and usually small) number of disjoint subsets $\{G_i\}$, each of which satisfies conditions (1) and (2) above. Then, we have

$$\sum_{G} = \sum_{\{G_1\}} + \cdots + \sum_{\{G_n\}} , \qquad (27)$$

and the individual subset summations are factorizable. For example, for a crystal with fcc lattice in real space, the G's can be partitioned into two subsets of G's which form two interpenetrating simple cubic sublattices. This is illustrated schematically in Fig. 2 for a two-dimensional reciprocal lattice.

Both the computational effort and storage for the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices can be further reduced by applying group theory. Let $\{R \mid \tau_R\}$ be the space group of the system, where R is a point-group operation and τ_R is the associated nonprimitive translation vector. It is not difficult to show that Eq. (9) can be rewritten as

$$\rho(\mathbf{G}) = \sum_{n} \sum_{\mathbf{k}\in IZ} \sum_{i,j} c_i(n,\mathbf{k}) c_j^*(n,\mathbf{k})$$

$$\times \left[\sum_{R} \Pi_{ij}(\mathbf{k},R^{-1}\mathbf{G})e^{-i\mathbf{G}\cdot\tau_R}\right] , \quad (28)$$

so that we just need to consider the $\mathbf{k}$ points in the irreducible wedge of the Brillouin zone (IZ). Furthermore, if we define a set of symmetrized $\Pi_{ij}^{sym}(\mathbf{k},\mathbf{G})$ matrices by

$$\Pi_{ij}^{sym}(\mathbf{k},\mathbf{G}) = \sum_{R} \Pi_{ij}(\mathbf{k},R^{-1}\mathbf{G})e^{-i\mathbf{G}\cdot\tau_R} , \qquad (29)$$

then, Eq. (28) becomes

$$\rho(\mathbf{G}) = \sum_{\mathbf{k}\in IZ} \sum_{n} \psi_{n\mathbf{k}}^{\dagger}\Pi^{sym}(\mathbf{k},\mathbf{G})\psi_{n\mathbf{k}} . \qquad (30)$$

Also, we need only to calculate $\rho(\mathbf{G})$ for a prototypical G in every star of the set of G's. The charge density for other G's in the same star can be generated by

$$\rho(R^{-1}\mathbf{G}) = e^{-i\mathbf{G}\cdot\tau_R}\rho(\mathbf{G}) . \qquad (31)$$
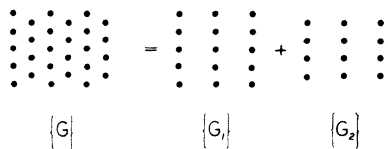
Similarly, Eq. (14) is simplified to

{G}   {G_i}   {G_2}

FIG. 2. Example of partitioning a set of two-dimensional G vectors into two disjoint subsets, each of which is primitive orthorhombic.

$$\delta H_{ij}(\mathbf{k}) = \sum_{\text{stars}} \delta V(\mathbf{G}_s)\Pi_{ij}^{sym}(\mathbf{k},\mathbf{G}_s)^* \frac{N_R}{N}\Omega_c , \qquad (32)$$

where the sum is over the stars of G vectors (not the individual G) and G_s on the right-hand side of Eq. (32) is a prototypical G vector (one for each star) for a given star. N_R is the number of G vectors in a star of G vectors and N is the order of the space group. Hence, we see that one needs only to calculate and store the symmetrized $\Pi_{ij}^{sym}(\mathbf{k},\mathbf{G}_s)$ matrix for $\mathbf{k}$ points in the IZ and for one G_s in each star.

We have here illustrated the evaluation of the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices for LCAO's and plane waves. It is clear that a mixed basis,[12] which contains both plane waves and Gaussian orbitals can similarly be treated by the above $\Pi_{ij}(\mathbf{k},\mathbf{G})$ approach. The present approach will also improve the efficiency of obtaining self-consistency.

III. APPLICATION TO Si AND W

The present self-consistent scheme is applied to extend the pseudopotential LCAO method in Ref. 6 to full point-by-point self-consistency. This approach, which utilizes linear combination of Gaussian orbitals and the variational property of the total energy functional, has been applied to treat the cohesive and structural properties of several insulators, semiconductors, and transition metals. We demonstrate in this section the effects of self-consistency on the band structure and cohesive properties of two prototypical systems: Si and W. These two systems were chosen because of the availability of self-consistent calculations that we can compare results with [e.g., for Si (Refs. 13 and 14), and for W (Refs. 15—17] and also because of their different bonding characters (Si being a covalent semiconductor and W a transition metal), which should test the versatility of the present scheme.

For Si, the calculation procedure can be summarized as follows.

(i) The norm-conserving ionic pseudopotential is generated by the method of Hamann et al.[18]

(ii) The atomic pseudo-charge-density (corresponding to an sp^3 configuration) is superimposed to form a crystalline charge density from which the input Hartree and exchange-correlation potentials are generated. The Hedin-Lundqvist[7] form of local exchange-correlation potential is used.

(iii) The ionic pseudopotential for $l=2$ (V_d) is used as the local potential. It is screened by the Hartree and exchange-correlation potentials. The screened local potential is then fitted to a set of Gaussians centered on the atomic sites. The nonlocal potentials ($V_s - V_d$ and $V_p - V_d$) are also fitted to on-site Gaussians. The fittings are done by a Monte Carlo simulated annealing minimization method.[19]

(iv) A LCAO basis set of the form (21) is employed to expand the wave function with radial functions of the form (24) for the atomic orbitals. Kubic harmonics up to $l=2$ (i.e., s,p,d orbitals) are used for the angular part and three Gaussian functions with exponential coefficients $\alpha = 0.202$, 0.758, and 2.837 are used for the radial part. The α's are determined by minimizing the total energy.

(v) Ten special **k** points[20] in the irreducible Brillouin zone are used in calculating the charge density. The total energy is computed using Eq. (3). Lattice constants, cohesive energy, and bulk modulus are then obtained by fitting the total energy versus volume curve to the Murnaghan equation of state.[21] (Eight points are used). These procedures give us the cohesive properties of Si as reported in Ref. 9. The calculation at this stage is not fully self-consistent and the accuracy of this method depends on the fact that the total energy is variational with respect to the charge density. This procedure can be considered as a variational approach.[6] This terminology will be used in this paper for the purpose of contrasting to the fully self-consistent results of the scheme described in Sec. II. The variational approach may be regarded as a one iteration calculation done in a very careful way.

(vi) The problem is done fully self-consistently using the present self-consistency scheme as depicted in Fig. 1. We stop iterating when the Fourier components of the output and input potentials have a maximum difference of less than 10^{-4} Ry. (This is achieved typically by 3—5 iterations).

We can now compare the variational (or one step) and the fully-self-consistent results. In Table I, the cohesive properties are given. The band structures are plotted in Fig. 3. The solid lines in Fig. 3 correspond to the fully-self-consistent results and the dashed lines correspond to the variational results. Band energies at high-symmetry **k** points calculated by the pseudopotential plane-wave (PW) method[22] are also marked (as dots) in Fig. 3 for comparison. These PW band energies are calculated with a very high plane-wave energy cutoff of 21.5 Ry. The band energies at high-symmetry **k** points from the present calculation are listed in Table II. They are also compared with the results from a highly converged (energy cutoff of 21.5 Ry) and a not-so-well-converged (energy cutoff of 5 Ry) PW calculations.[22]

From Table I, we see that the cohesive properties are almost unchanged by self-consistency except for a small increase in cohesive energy as expected. However, Fig. 3 shows that the band structure does change as full self-consistency is included. This is not surprising since there is no variational principle for the individual Kohn-Sham eigenvalues. Comparing with the PW band energies in Table II and Fig. 3, we note that the self-consistent band structure agrees very well with the high-cutoff PW result. The Hamiltonian matrix of the LCAO calculation (60×60) is approximately the same size as the low-cutoff PW calculation but is substantially smaller than that of the converged high-cutoff PW calculation (~450×450)

TABLE I. Structural and cohesive properties of Si obtained by variational and fully self-consistent LCAO methods. Experimental values are quoted from Ref. 13.

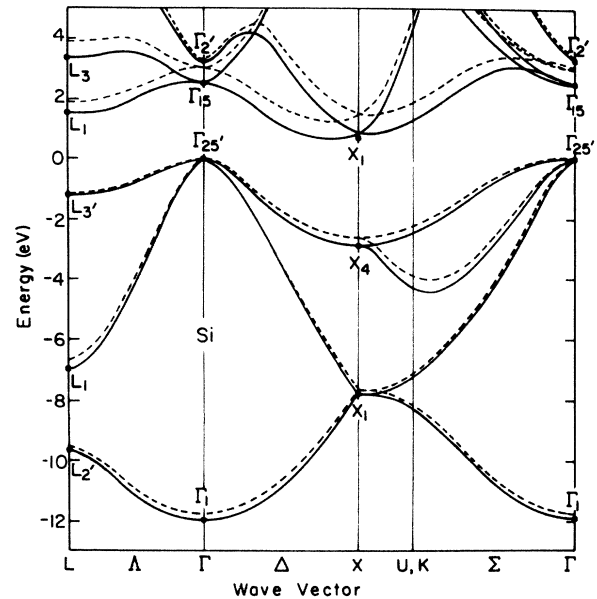|	Variational	Fully Self-consistent	Experiment
Lattice constant (Å)	5.35	5.35	5.43
Bulk modulus (Mbar)	1.15	1.16	0.99
Cohesive energy (eV)	4.9	5.0	4.63

FIG. 3. Band structure of Si calculated by the full self-consistent LCAO approach (solid lines) and the variational LCAO approach (dashed lines). The band energies from a 21.5-Ry cutoff plane-wave calculation (Ref. 22) are marked as dots at points Γ, X and L.

despite the fact that Si is one of elements that should be very efficiently treated by plane waves. For more localized systems such as those involving C, the self-consistent LCAO method is even more practical. The valence-band energies of the variational approach differ only slightly from the fully self-consistent results. The changes are more pronounced for the conduction bands. Comparing with the band energies of the low-cutoff (5 Ry) plane-wave calculation in Table II, it is fair to say that even the band structure of the non-self-consistent variational calculation is better than the low-cutoff self-consistent plane-wave results.

The sensitivity of the result with respect to the choice of the radial exponential coefficients are also tested. In Tables III and IV, we compare the structural properties and band structure of Si calculated using three decay constants $\alpha = (0.202, 0.758, 2.837)$ and four decay constants $\alpha = (0.2, 0.41, 0.85, 1.75)$. The change is indeed rather small except for a slight increase in the cohesive energy for the more complete basis set.

In Table I, we note that the lattice constant is about 1.5% too small compared with the experimental value whereas PW calculations[13] have shown an agreement of better than 1%. This is partially due to the fact the Hedin-Lundqvist exchange-correlation potential is used in the present calculation while the Wigner[8] exchange-correlation potential is used by Yin and Cohen.[13] Using the Wigner form consistently gives a larger lattice constant and a smaller cohesive energy (both of which would improve the present results compared with experiment.) Another difference is that we have used the d ionic pseudopotential as the local potential whereas Yin and Cohen

TABLE II. Comparison of band energies (in eV) of Si obtained by LCAO and plane-wave (PW) (Ref. 22) methods.

	Variational LCAO	Fully self-consistent LCAO	PW $E_{cut}=21.5$ Ry	PW $E_{cut}=5$ Ry
Γ_{1v}	-11.73	-11.99	-11.91	-11.21
Γ'_{25v}	0.0	0.0	0.0	0.0
Γ_{15c}	3.00	2.53	2.55	2.72
Γ'_{2c}	3.35	3.21	3.28	5.62
X_{1v}	-7.69	-7.84	-7.76	-6.88
X_{4v}	-2.64	-2.86	-2.86	-2.80
X_{1c}	1.44	0.84	0.66	1.29
L'_{2v}	-9.50	-9.65	-9.56	-8.82
L_{1v}	-6.70	-7.00	-6.96	-6.50
L'_{3v}	-1.11	-1.21	-1.20	-1.17
L_{1c}	1.89	1.54	1.50	3.36
L_{3c}	3.89	3.38	3.33	3.79

used the s ionic pseudopotential as the local potential. We found that the use of the higher angular momentum pseudopotential as the local potential gives a slightly smaller lattice constant. This also contributes to the slight difference in the results of the two calculations.

We also note that the cohesive properties of Si can be described very accurately without d states in the LCAO basis set. (In other words by just treating a 24×24 matrix, using three decay constants in the radial part of the atomic orbitals.) This makes the method even more appealing in treating structural properties. The d states are however needed for an accurate description of the band structure, especially the conduction bands.

Tungsten is studied is essentially the same way as Si, except for the following differences.

(i) The scalar relativistic ionic pseudopotential[23] is used. This is because W is much heavier than Si and the band structure is changed substantially when relativistic effects are included.

(ii) The Wigner[8] form of local exchange-correlation potential is used. The nonlinearity of the exchange-correlation interaction between core and valence electron is treated by the partial-core method,[24] owing to the relatively large extension of the $5p$ core-state wave function.

(iii) A uniform grid of 47 k points is used for the

charge density. Four exponential decay constants $\alpha = (0.2, 0.49, 1.22, 3.0)$ are used for the atomic orbitals. These values of α are again determined by minimizing the total energy of the system.

(iv) For the non-self-consistent part of the calculation, whereas the "variational" result of Si is done by a single iteration process, W is done in a two-step process. This is the way that Mo has been treated, as has been reported briefly previously.[9] The physical reason for this is that the s and p atomic pseudowave-functions for Si are essentially occupying the same radial region in real space, while the d pseudo-wave-functions in W (and in other transition elements) being more localized, occupy different regions as compared with the s or p states. Consequently, $\rho(\mathbf{r})$ from the superposition of atomic charge densities for Si is only weakly dependent on the atomic configuration

TABLE III. Comparison of structural properties of Si obtained by the fully self-consistent LCAO scheme with three and four radial Gaussian functions. B'_0 is the derivative of the bulk modulus at the equilibrium volume as determined by the Murnaghan equation of state (Ref. 21).

	LCAO Three radial Gaussians	LCAO Four radial Gaussians
Lattice constant (Å)	5.35	5.35
Bulk modulus (Mbar)	1.16	1.13
Cohesive energy (eV)	5.0	5.14
B'_0	3.5	3.6

TABLE IV. Comparison of band energies (in eV) of Si obtained by the fully self-consistent LCAO scheme with three and four radial Gaussian functions. Band energies from a high-cutoff (21.5 Ry) plane-wave calculation (Ref. 22) are also quoted for comparison.

	LCAO Three radial Gaussians	LCAO Four radial Gaussians	PW $E_{cut}=21.5$ Ry
Γ_{1v}	-11.99	-11.92	-11.91
Γ'_{25v}	0.0	0.0	0.0
Γ_{15c}	2.53	2.59	2.55
Γ'_{2c}	3.21	3.29	3.28
X_{1v}	-7.84	-7.78	-7.76
X_{4v}	-2.86	-2.81	-2.86
X_{1c}	0.84	0.75	0.66
L'_{2v}	-9.65	-9.58	-9.56
L_{1v}	-7.00	-6.97	-6.96
L'_{3v}	-1.21	-1.20	-1.20
L_{1c}	1.54	1.58	1.50
L_{3c}	3.38	3.38	3.33

chosen, and is reasonably close to the final self-consistent charge density. On the other hand, the superposition of atomic charge densities, and hence the results of the first iteration, do depend on the particular atomic configuration chosen in the case of W. To circumvent this difficulty, the output crystalline charge density in the first iteration is refitted to a superposition of effective atomic charges centered on the atomic sites.

$$\rho(\mathbf{r}) = \sum_{\mathbf{R},\tau} \rho_{\text{eff}}(\mathbf{r} - \mathbf{R} - \tau) \ , \tag{33}$$

with the constraints that the total number of electrons is conserved and that the charge density is positive definite. The complete procedure is then repeated once more. The resultant charge density at the second iteration is used to compute the total energy by Eq. (3). We found that there is no point to go beyond the second step. This is because the purpose of the second step is to avoid the dependence of the result on the input atomic configuration and a second iteration already served the purpose. Moreover, the fit as described by Eq. (33) is a difficult constrained nonlinear minimization problem and the fitting is usually not of sufficiently high quality to justify further iterations. The self-consistent procedure developed in Sec. II of course alleviates all these difficulties. For the present discussion, we refer the variational results as the results of this two iteration process as contrast to the fully self-consistent results which were obtained by the procedure described by Fig. 1.

The structural properties are compared in Table V. We see that the two-step variational procedure is adequate for a fairly good description of structural and cohesive properties for W. We have arrived at the same conclusion for Mo.[9] However, the self-consistency scheme does improve the result. There is some difference between Si and W as far as the effects of self-consistency on the structural properties are concerned. For Si, self-consistency lowers the energy (as a function of volume) almost uniformly, inducing practically no change in the calculated lattice constant and bulk modulus. Even the derivative of the bulk modulus does not change significantly. For W, the change in energy as a function of volume due to self-consistency is small but not uniform, giving a small change in lattice constant and bulk modulus. This gives a small improvement as compared with the experimental results. The agreement with experiment is very good. We note that Bylander and Kleinman[15] and Zunger and Cohen[16] and Jansen and Freeman[17] have also obtained

structural results that agree very favorably with experimental data.

The band-energy results for W are shown in Table VI. Results by Bylander and Kleinman[15] are also quoted for comparison. Comparing the band energies in Table VI, we found that the self-consistent results and the variational results of the present formalism differ by an average of about 0.1 eV. An exception is the position of Γ_1 relative to the d-bands. The difference Γ'_{25}-Γ_1 changes by 0.3 eV, indicating that the s-d separation is quite sensitive to self-consistency. Comparing our results with that of Bylander and Kleinman,[15] we find that our band energies agree rather well with theirs. The only discrepancy is the position of Γ_1 relative to the d bands. The present self-consistent value differ from that of Bylander and Kleinman by almost 0.5 eV. Such a discrepancy is surprising as the two calculations are very similar and only differ in the details of the starting screening potential, the basis set employed, and the way in which self-consistency is treated. In view of this, we compare in Table VII the Fourier components of the self-consistent pseudo-charge-density with those of Bylander and Kleinman[15] and Zunger and Cohen.[16] Though not exactly equal, our form factors agree fairly well with those of Bylander and Kleinman. Those of Zunger and Cohen are rather different probably because a nonrelativistic hard-core pseudopotential was used. As a further test, we have done some preliminary calculation on Pd and Ni and for both cases, we obtain band energies very similar to the self-consistent result of Moruzzi et al.[25]

VI. DISCUSSION AND SUMMARY

The purpose of the present work is to extend an *ab initio* pseudopotential LCAO scheme[6] to full point-by-point self-consistency. Making use of the matrices $\Pi_{ij}(\mathbf{k},\mathbf{G})$, we arrive at a scheme that is formally independent of the basis set chosen. When applied to a LCAO basis set, it avoids some well-known problems like the fitting of the charge density and potential. The method is exact within the basis subspace and the LDA.

TABLE V. Structural and cohesive properties of W calculated by the variational and fully self-consistent LCAO method. Experimental values are quoted from C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976).

	Variational	Fully self-consistent	Experiment
Lattice constant (Å)	3.12	3.13	3.16
Bulk modulus (Mbar)	3.438	3.335	3.232
Cohesive energy (eV)	8.43	8.46	8.90

TABLE VI. Band energies of W (in eV) calculated by the variational LCAO and the fully self-consistent LCAO methods. Results of Bylander and Kleinman (BK) (Ref. 15) are quoted for comparison. The energy scale is set to 0 at Γ'_{25}.

	Variational	Fully self-consistent LCAO	BK
Γ_1	−8.946	−8.645	−8.186
Γ'_{25}	0.0	0.0	0.0
Γ_{12}	3.507	3.491	3.425
N_1	−5.251	−5.071	−4.934
N_2	−2.402	−2.341	−2.315
N'_1	1.844	2.161	2.311
N_1	3.420	3.388	3.333
N_4	4.437	4.364	4.249
H_{12}	−4.756	−4.593	−4.557
P_4	−1.883	−1.749	−1.641
P_3	4.599	4.518	4.403

TABLE VII. Comparison of the Fourier components of the fully self-consistent valence charge density in electrons/cell between the present self-consistent LCAO calculation, the Bylander and Kleinman (BK) (Ref. 15) results, and the Zunger and Cohen (ZC) (Ref. 16) results.

G	Present work	BK $\rho(G)$	ZC
(000)	6.000	6.000	6.000
(110)	0.412	0.388	0.691
(200)	−0.157	−0.152	−0.024
(211)	−0.091	−0.107	0.034
(220)	−0.099	−0.123	−0.076
(310)	−0.174	−0.180	−0.176
(222)	−0.052	−0.052	−0.043
(321)	−0.063	−0.055	−0.075

We have formulated the $\Pi_{ij}(\mathbf{k},\mathbf{G})$ matrices in Fourier space. This is because the lower Fourier components of the screening potential are more sensitive to changing environment and contribute importantly to the total energy. Another key reason is that the periodic electron screening potential in a crystal can be expressed arbitrarily accurately by a Fourier expansion. This almost dictates a formulation in Fourier space as can be seen from Eqs. (13) and (14) if we want the potential to be free of form constraints. The use of pseudopotentials in the present calculations further favors Fourier-space expansions.

The method is tested on Si and W. The calculated structural and electronic properties agree well with experiment and previous calculations. The fact that there are only very small changes in structural properties in going from the variational results to full self-consistent results justify the previous claim[6,9] that self-consistent iterations can be avoided in some applications if the total energy is formulated carefully as an explicit functional of the output charge density only. The self-consistency scheme is however an important improvement. For example, treating W fully self-consistently using the present scheme is actually easier than the two-step calculation which needs the fitting of the output crystalline charge density to a superposition of effective atomic charge density. The self-consistent result is also more satisfactory and reliable from a theoretical point of view. The fully self-consistent scheme would be indispensable in treating systems like ionic compounds and surface problems where charge rear-

rangements are expected to be substantial. We also note that despite Si and W having rather different bonding properties, they can be treated in the same way by the present scheme. In fact, their properties were calculated by the same computer code taking approximately the same amount of time. This demonstrates the versatility of the method.

Lastly, we make a comment about the storage of the $\Pi_{ij}^{\text{sym}}(\mathbf{k},\mathbf{G}_s)$ matrices, which are the key quantities in the present approach. These matrices have four indices and the total number of elements is proportional to the number of $\mathbf{k}$ points, the square of the number of basis functions used, and the number of stars of $\mathbf{G}$ vectors used. For W, we used 47 $\mathbf{k}$ points, 40 basis functions, and about 40 stars of $\mathbf{G}$ vectors (corresponding to about 1000 $\mathbf{G}$ vectors with a $\mathbf{G}$ space cutoff of 8 a.u.$^{-1}$). The $\Pi_{ij}^{\text{sym}}(\mathbf{k},\mathbf{G}_s)$ matrices have altogether about 3×10^6 numbers. Si has two atoms per unit cell but less $\mathbf{k}$ points are needed for a semiconductor, so the $\Pi_{ij}^{\text{sym}}(\mathbf{k},\mathbf{G}_s)$ matrices are of approximately the same size as W. The matrix sizes quoted here are for high convergent results and they can be reduced without introducing significant errors.[26] The use of pseudopotentials is important because the electron wave functions are then nodeless and the (pseudo)charge density can be represented by a reasonably small number of stars of $\mathbf{G}$ vectors. For surface calculations, the $\pi_{ij}^{\text{sym}}(\mathbf{k},\mathbf{G}_s)$ matrices are large. However, it is usually the low Fourier components of the potential that are most sensitive to self-consistency in surface problems. One can reduce storage and input and/or output load during the iteration cycles by employing a lower cutoff in $\mathbf{G}$ space and use a higher cutoff (i.e., more stars of $\mathbf{G}$ vectors) in the final stage of calculating the total energy when the potential is self-consistent.

ACKNOWLEDGMENTS

*Present address: Lyman Laboratory of Physics, Harvard University, Cambridge, MA 02138.

[1]P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964); W. Kohn and L. J. Sham, ibid. 140, A1133 (1965); Theory of the Inhomogeneous Electron Gas, edited by N. H. March and S. Lundqvist (Plenum, New York, 1983).

[2]E. Lafon, R. C. Chaney, and C. C. Lin, in Computational Methods in Band Theory, edited by P. M. Marcus, D. F. Janak, and A. R. Williams (Plenum, New York, 1971), p. 284.

[3]J. Langlinais and J. Callaway, Phys. Rev. B 5, 124 (1972).

[4]See for example, B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. 71, 3396 (1979); H. Sambe and R. H. Felton, ibid. 62, 1122 (1975); C. S. Wang and A. J. Freeman, Phys. Rev. B 18, 1714 (1978); J. Bernholc and N. A. W. Holzwarth, J. Chem. Phys. 81, 3987 (1984).

[5]J. R. Smith, J. G. Gay, and F. J. Arlinghaus, Phys. Rev. B 21, 2201 (1980).

[6]J. R. Chelikowsky and S. G. Louie, Phys. Rev. B 29, 3470 (1984).

[7]L. Hedin and B. I. Lundqvist, J. Phys. C 4, 2064 (1971).

[8]E. Wigner, Phys. Rev. **46**, 1002 (1934).

[9]J. R. Chelikowsky, S. G. Louie, D. Vanderbilt, and C. T. Chan, Int. J. Quantum Chem.: Quant. Chem. Symp. **18**, 105 (1984).

[10]See, for example, P. Bendt and Z. Zunger, Phys. Rev. B **26**, 3114 (1982).

[11]K. M. Ho, J. Ihm, and J. D. Joannopoulous, Phys. Rev. B **25**, 4260 (1982).

[12]S. G. Louie, K. M. Ho, and M. L. Cohen, Phys. Rev. B **19**, 1774 (1979).

[13]M. T. Yin and M. L. Cohen, Phys. Rev. Lett. **45**, 1004 (1980).

[14]B. Harmon, W. Weber, and D. R. Hamann, Phys. Rev. B **25**, 1109 (1982).

[15]D. M. Bylander and L. Kleinman, Phys. Rev. B **27**, 3152 (1983).

[16]A. Zunger, and M. L. Cohen, Phys. Rev. B **20**, 4082 (1979).

[17]H. J. F. Jansen and A. J. Freeman, Phys. Rev. B **30**, 561 (1984).

[18]D. H. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[19]D. Vanderbilt and S. G. Louie, J. Comput. Phys. **56**, 259 (1984).

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

[21]F. D. Murnaghan, Proc. Natl. Acad. Sci. USA **3**, 244 (1944).

[22]K. J. Chang, private communication; M. Hybertsen, private communication.

[23]L. Kleinman, Phys. Rev. B **21**, 2630 (1980); G. B. Bachelet and M. Schulter, *ibid.* **25**, 2103 (1982).

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

[25]V. L. Moruzzi, J. F. Janak, and A. R. Williams, *Calculated Electronic Properties of Metals* (Pergamon, New York, 1978).

[26]For example, we have tested a G-space cutoff of 5 a.u.$^{-1}$ for W and the results are practically the same as those calculated with a G-space cutoff of 8 a.u.$^{-1}$