

Alternative approach to separable first-principles pseudopotentials

S. Goedecker and K. Maschke

EPF-Lausanne, Institut de Physique Appliquée, CH-1015 Lausanne, Switzerland

(Received 8 May 1990)

We construct a first-principles pseudo-Hamiltonian using ideas borrowed from the linear augmented-plane-wave method. Space is subdivided into the core region around the nuclei and the valence region, which is just the remainder of space. A basis set consisting of atomic-type orbitals in the core region and plane waves in the valence region leads to Hamiltonian and overlap matrices that depend on the plane-wave expansion coefficients only. The Hamiltonian and overlap matrices have the extremely useful property that their nonlocal part is separable. This allows eigenvectors and eigenvalues to be calculated with $O(n \ln(n))$ operations, where n is the number of plane waves. The method is also well suited for first-row atoms. As an example we present results obtained for oxygen.

INTRODUCTION

The solution of the Schrödinger equation can be mapped onto a matrix eigenvalue problem by representing the wave function as a linear superposition of basis functions. The choice of the basis functions is crucial for the efficiency of the algorithm. A basis set of plane waves would be convenient in many respects. Unfortunately, an unrealistically large number of plane waves is required to describe the strong oscillations near the core. Pseudopotentials eliminate the core states and the strong oscillations of the valence functions near the core and allow therefore the use of plane waves. Even though the pseudo-wave-functions differ from the true wave functions in the core region, accurate norm-conserving pseudopotentials¹ allow us to calculate many interesting quantities such as total energies, equilibrium configurations, and mechanical properties. Quantities that require the knowledge of the charge density near the core, however, cannot be described properly.

In order to describe the exact wave function everywhere in space, atomiclike orbitals in the core region and plane waves in the region of chemical bonding are the optimal basis set. Among others, the augmented-plane-wave (APW) method is based on this observation. In the following, we will show that this basis set can also be the starting point of a separable first-principles (SFP) pseudopotential. In contrast to conventional pseudopotential theory, we will actually circumvent the intermediate step of constructing a pseudopotential in real space and instead derive directly the resulting pseudo-Hamiltonian matrix.

CONSTRUCTION OF THE PSEUDO-HAMILTONIAN

In the core region, the potential is assumed to have spherical symmetry. In this region the wave functions Φ_{core} can therefore be written as a superposition of atomic wave functions:

$$\Phi_{\text{core}} = \sum_{l=0}^{l_{\text{max}}} \sum_m a_{l,m}^1 R_l^1 Y_{l,m} + \sum_{l=0}^{l_{\text{max}}} \sum_m a_{l,m}^2 R_l^2 Y_{l,m} . \quad (1)$$

Strategies for choosing optimal functions $R_l^1(r)$ and $R_l^2(r)$ will be discussed later on. For the moment we only need their properties on the surface of the sphere with radius r_c which separates the core and the valence region:

$$R_l^1(r_c) = 1, \quad \frac{\partial R_l^1(r_c)}{\partial r} = 0 , \quad (2)$$

$$R_l^2(r_c) = 0, \quad \frac{\partial R_l^2(r_c)}{\partial r} = 1 . \quad (3)$$

In the interstitial region the wave function Φ_{out} is written as a superposition of plane waves:

$$\Phi_{\text{out}} = \sum_{\mathbf{K}} c_{\mathbf{K}} e^{i\mathbf{K}\cdot\mathbf{r}} .$$

We now require that the wave function and its gradient be continuous across the boundary of the sphere. Since the weight of the different contributions to the wave function is rapidly decreasing with increasing angular momentum l , we impose these two requirements only on the low- l components up to l_{max} . This gives

$$a_{l,m}^1 = \int Y_{l,m}^* \Phi_{\text{out}}|_{|\mathbf{r}|=r_c} d\Omega ,$$

$$a_{l,m}^2 = \int Y_{l,m}^* \hat{\mathbf{n}} \cdot \nabla \Phi_{\text{out}}|_{|\mathbf{r}|=r_c} d\Omega .$$

Defining the projectors

$$p_{l,m}^{1*}(\mathbf{K}) = \int Y_{l,m}^* e^{i\mathbf{K}\cdot\mathbf{r}}|_{|\mathbf{r}|=r_c} d\Omega ,$$

$$p_{l,m}^{2*}(\mathbf{K}) = \int Y_{l,m}^* (i\mathbf{K} \cdot \hat{\mathbf{n}}) e^{i\mathbf{K}\cdot\mathbf{r}}|_{|\mathbf{r}|=r_c} d\Omega ,$$

where $\hat{\mathbf{n}}$ is the unit outwardly directed radial vector, the expansion coefficients of the core region are then given by

$$a_{l,m}^1 = \sum_{\mathbf{K}} p_{l,m}^{1*}(\mathbf{K}) c_{\mathbf{K}} , \quad (4)$$

$$a_{l,m}^2 = \sum_{\mathbf{K}} p_{l,m}^{2*}(\mathbf{K}) c_{\mathbf{K}} . \quad (5)$$

The expectation value for the energy (in hartrees) can now be written in the following way:

$$E = \int_{\text{outer space}} \left(\frac{1}{2} \nabla \Phi_{\text{out}}^* \nabla \Phi_{\text{out}} + \Phi_{\text{out}}^* V \Phi_{\text{out}} \right) + \int_{\text{core}} \left(\frac{1}{2} \nabla \Phi_{\text{core}}^* \nabla \Phi_{\text{core}} + \Phi_{\text{core}}^* V \Phi_{\text{core}} \right)$$

or after rearranging the integration,

$$E = \int_{\text{all space}} \left(\frac{1}{2} \nabla \Phi_{\text{out}}^* \nabla \Phi_{\text{out}} + \Phi_{\text{out}}^* V \Phi_{\text{out}} \right) + \int_{\text{core}} \left(\frac{1}{2} \nabla \Phi_{\text{core}}^* \nabla \Phi_{\text{core}} + \Phi_{\text{core}}^* V \Phi_{\text{core}} \right) - \left(\frac{1}{2} \nabla \Phi_{\text{out}}^* \nabla \Phi_{\text{out}} + \Phi_{\text{out}}^* V \Phi_{\text{out}} \right) .$$

The second integral can be interpreted as the pseudopotential correction. Expressing everything as a function of the plane-wave expansion coefficients $c_{\mathbf{K}}$, we obtain after some algebraic manipulations using (1), (4), and (5),

$$E = \sum_{\mathbf{K}, \mathbf{K}'} c_{\mathbf{K}}^* H_{\mathbf{K}, \mathbf{K}'} c_{\mathbf{K}'} + \sum_{\mathbf{K}, \mathbf{K}'} c_{\mathbf{K}}^* H_{\mathbf{K}, \mathbf{K}'}^{\text{ps}} c_{\mathbf{K}'} .$$

$H_{\mathbf{K}, \mathbf{K}'}$ and the pseudopart of the Hamiltonian $H_{\mathbf{K}, \mathbf{K}'}^{\text{ps}}$ are given by

$$H_{\mathbf{K}, \mathbf{K}'} = \Omega \left(\frac{1}{2} \delta_{\mathbf{K}, \mathbf{K}'} K^2 + V_{\mathbf{K}-\mathbf{K}'} \right) ,$$

where Ω is the volume and

$$H_{\mathbf{K}, \mathbf{K}'}^{\text{ps}} = -\mathbf{K} F(|\mathbf{K}-\mathbf{K}'|r_c) \mathbf{K}' - \int_{\text{core}} V e^{-i(\mathbf{K}-\mathbf{K}')\mathbf{r}} + \sum_{l=0}^{l_{\text{max}}} \sum_m p_{l,m}^1(\mathbf{K}) H_l^1 p_{l,m}^{1*}(\mathbf{K}') + [p_{l,m}^1(\mathbf{K}) H_l^{12} p_{l,m}^{2*}(\mathbf{K}') + \text{H.c.}] + p_{l,m}^2(\mathbf{K}) H_l^2 p_{l,m}^{2*}(\mathbf{K}') ,$$

where

$$F(x) = 4\pi r_c^3 \frac{\sin(x) - x \cos(x)}{x^3} ,$$

$$H_l^1 = \frac{1}{2} \left[\int_0^{r_c} \left[r \frac{\partial R_l^1}{\partial r} \right]^2 dr + l(l+1) \int_0^{r_c} (R_l^1)^2 dr \right] + \int_0^{r_c} (R_l^1)^2 V(r) r^2 dr ,$$

$$H_l^{12} = \frac{1}{2} \left[\int_0^{r_c} \left[r \frac{\partial R_l^1}{\partial r} \right] \left[r \frac{\partial R_l^2}{\partial r} \right] dr + l(l+1) \int_0^{r_c} R_l^1 R_l^2 dr \right] + \int_0^{r_c} R_l^1 R_l^2 V(r) r^2 dr ,$$

$$H_l^2 = \frac{1}{2} \left[\int_0^{r_c} \left[r \frac{\partial R_l^2}{\partial r} \right]^2 dr + l(l+1) \int_0^{r_c} (R_l^2)^2 dr \right] + \int_0^{r_c} (R_l^2)^2 V(r) r^2 dr .$$

Even though the two basis sets have no overlap in real space, orthonormality of the coefficient vectors $c_{\mathbf{K}}$ does not imply orthonormality of the wave functions. We therefore have to introduce an overlap matrix. It can be calculated along the same lines as H . The result is

$$S_{\mathbf{K}, \mathbf{K}'} = \Omega \delta_{\mathbf{K}, \mathbf{K}'} - F(|\mathbf{K}-\mathbf{K}'|r_c) + \sum_{l=0}^{l_{\text{max}}} \sum_m p_{l,m}^1(\mathbf{K}) S_l^1 p_{l,m}^{1*}(\mathbf{K}') + [p_{l,m}^1(\mathbf{K}) S_l^{12} p_{l,m}^{2*}(\mathbf{K}') + \text{H.c.}] + p_{l,m}^2(\mathbf{K}) S_l^2 p_{l,m}^{2*}(\mathbf{K}') ,$$

where

$$S_l^1 = \int_0^{r_c} (R_l^1)^2 r^2 dr ,$$

$$S_l^{12} = \int_0^{r_c} (R_l^1 R_l^2) r^2 dr ,$$

$$S_l^2 = \int_0^{r_c} (R_l^2)^2 r^2 dr .$$

CHOICE OF THE BASIS FUNCTIONS R_l^1, R_l^2

The construction of the basis functions R_l^1, R_l^2 consists of two steps. The first step is to choose two functions that allow a good representation of the true wave function in the range of energy we are interested in. In the second step we form linear combinations that satisfy the boundary conditions (2) and (3). Since this last step is trivial we will only discuss the first one.

The basis functions R_l^1, R_l^2 enter into the calculation only through the matrix elements S^1, S^2, S^{12} and H^1, H^2, H^{12} . These two sets of matrix elements allow us to calculate approximately the norm S_l and energy H_l . If the true wave function R_l were known, these two quantities would be given by the following integrals:

$$S_l = \int_0^{r_c} (R_l)^2 r^2 dr , \quad (6)$$

$$H_l = \frac{1}{2} \left[\int_0^{r_c} \left[r \frac{\partial R_l}{\partial r} \right]^2 dr + l(l+1) \int_0^{r_c} (R_l)^2 dr \right] + \int_0^{r_c} (R_l)^2 V(r) r^2 dr . \quad (7)$$

In the SFP pseudopotential the true function R_l in the above integrals is replaced by an approximate function \tilde{R}_l which is a linear superposition of R_l^1 and R_l^2 such that the \tilde{R}_l has the same value and derivative on the surface of the sphere:

$$\tilde{R}_l(r) = R_l(r_c) R_l^1(r) + R_l'(r_c) R_l^2(r) .$$

The values for the exact S_l and H_l are therefore replaced by the approximate quantities \tilde{S}_l and \tilde{H}_l :

$$\tilde{S}_l = [R_l(r_c)]^2 S_l^1 + 2R_l(r_c) R_l'(r_c) S_l^{12} + [R_l'(r_c)]^2 S_l^2 ,$$

$$\tilde{H}_l = [R_l(r_c)]^2 H_l^1 + 2R_l(r_c) R_l'(r_c) H_l^{12} + [R_l'(r_c)]^2 H_l^2 .$$

The functions R_l^1, R_l^2 must therefore be chosen such that they trace with high precision S_l and H_l of the exact solution for all the true wave functions of the atom under various chemical environments. Instead of considering S_l and H_l separately, it is usual to consider the ratio \tilde{H}_l/\tilde{S}_l , which will simply be referred to as energy in the following. For a given potential and a given r_c this quantity depends only on the logarithmic derivative

$R'_l(r_c)/R_l(r_c)$. The basic properties of the functional relationship between the energy and the logarithmic derivative are well known. In Fig. 1 we illustrate these properties for the $l=0$ component of hydrogen.² The function consists of several branches that are separated by asymptotic energy values. These energy values are solutions of the radial Schrödinger equation with the boundary condition that the wave function be zero at r_c . Choosing a different r_c will shift these asymptotics. This dependency can be seen by a comparison between Figs. 1(a) and 1(b). It is important to note that several solutions of the atomic problem may lie on the same branch. Each branch is a monotonically decreasing function of the logarithmic derivative and the lowest branch tends to infinite negative energy.

In order to parametrize the relation between the energy and the logarithmic derivative, we first have to select a certain branch of this function. In this way we choose a certain energy window and we can for instance eliminate core states.³ Then we have to decide how to parametrize the selected branch. Evidently a compromise between high accuracy in a narrow energy range and a good fit

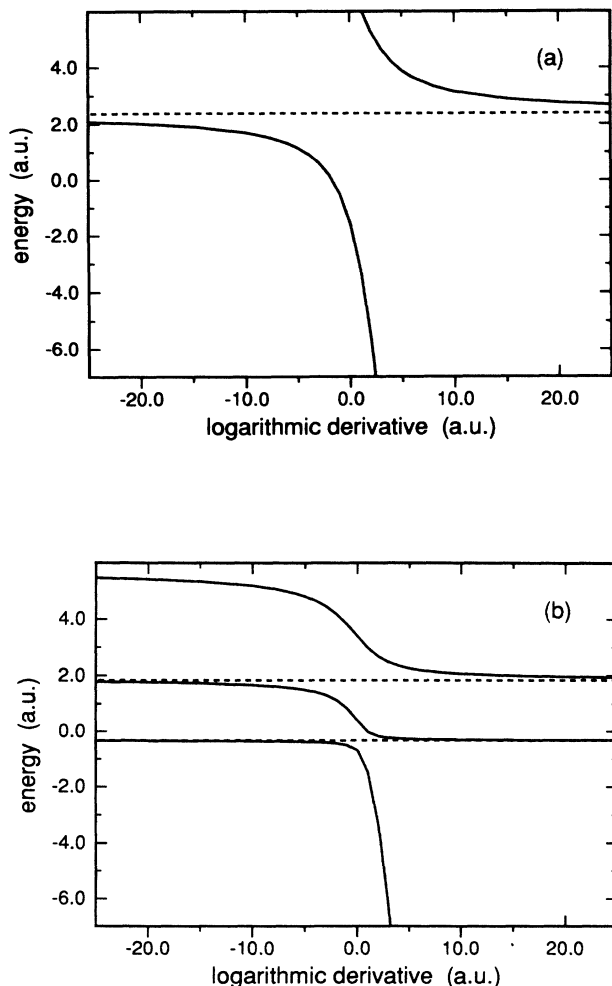


FIG. 1. Logarithmic derivative vs energy for the $l=0$ component in hydrogen. (a) $r_c = 1$ a.u., (b) $r_c = 2.5$ a.u.

over a wider energy range must be accepted. A choice which will give optimal accuracy around a fixed energy is used in the linear augmented-plane-wave (LAPW) method,^{4,5} where one takes the true solution at a fixed energy and its derivative with respect to the energy as the basis functions. If it should turn out that the energy range obtained with this method is not large enough, the coefficients $S^1, S^2, S^{12}, H^1, H^2, H^{12}$ could be determined by a least-squares fit if one is willing to accept a larger error at the reference energy. Expressed in the language of pseudopotential theory, this means that we have constructed in the first case a very accurate pseudopotential with poor transferability, whereas in the second case we have constructed a pseudopotential with good transferability but reduced accuracy at the reference energy.

Another parameter that must be chosen is the radius r_c . This parameter gives us considerable flexibility, since it allows us to shift the asymptotics of the energy versus logarithmic derivative function. Decreasing r_c will shift the asymptotics upwards. Evidently, it is impossible to describe several branches of the energy versus logarithmic derivative curve with one set of coefficients $S^1, S^2, S^{12}, H^1, H^2, H^{12}$. Therefore, we must choose r_c such that one branch of the energy versus logarithmic derivative curve opens the whole energy interval we are interested in,³ i.e., the considered energy range must not contain any asymptotics of the energy versus logarithmic derivative function. This excludes choosing r_c close to a zero of the wave function. If we choose r_c very large, the energy versus logarithmic derivative curve becomes very flat and our energy window too small. The region between the last node and the outermost maximum of the valence wave function therefore seems also appropriate for this type of pseudopotential. If a particularly soft pseudopotential is needed, an r_c beyond the outermost maximum may be chosen. However, in this case the potential within the sphere may not be spherical in all chemical environments and corrections for the nonsphericity as in full-potential LAPW-based methods will be necessary.

SYSTEMATIC IMPROVEMENT OF THE ACCURACY AND TRANSFERABILITY

From the theory of finite elements, it is known that a basis set for a second-order differential equation must at least be continuous, whereas its derivative need not be continuous. We have required the continuity of the derivative since it allows one to parametrize the energy as a function of the logarithmic derivative. If we had only required the continuity of the wave function, the energy density in the core region would have been a constant, evidently a bad approximation. One possibility to improve the transferability or accuracy would be to require higher derivatives to be continuous, too. Unfortunately the mathematics becomes fairly cumbersome in this case. An easier method would be to include additional basis functions in the core region which have the property that their values and their radial derivatives are zero at r_c . In addition to the plane-wave expansion coefficients, this would introduce additional degrees of freedom, which are the coefficients of these additional basis functions in the

core region. Since typically several hundred plane waves are used per atom, this slight increase in the size of the matrix is negligible.

NUMERICAL EFFICIENCY

Iterative diagonalization methods have turned out to be the most efficient in electronic-structure calculations.⁶ With these methods, the numerical effort for generalized eigenvalue problems is roughly two times larger than for simple eigenvalue problems. They all require the repeated calculation of matrix times vector products. This operation can be performed very efficiently for the SFP pseudopotential, because the matrices consist of sums of the following three types of matrices.

(i) Convolution matrices: $H_{\mathbf{K},\mathbf{K}'} = f(\mathbf{K} - \mathbf{K}')$. Using two fast Fourier transformations,⁷ matrix times vector products can be calculated with $O(n \ln(n))$ operations, where n is the number of plane waves.

(ii) Separable matrices: $H_{\mathbf{K},\mathbf{K}'} = f(\mathbf{K})g(\mathbf{K}')$. Matrix times vector products can be calculated with $O(n)$ operations.

(iii) Mixed-type matrices:

$$H_{\mathbf{K},\mathbf{K}'} = f(\mathbf{K})g(\mathbf{K} - \mathbf{K}')h(\mathbf{K}').$$

Matrix times vector products can be calculated with $O(n + n \ln(n))$ operations.

The SFP pseudopotential therefore allows us to calculate eigenvectors and eigenvalues nearly as fast as a local pseudopotential, where this can be done with $O(n \ln(n))$ operations.

COMPARISON WITH THE LAPW METHOD

The SFP pseudopotential and the LAPW method are based on the same idea. The differences are therefore rather quantitative. In the LAPW method the radius r_c is chosen as large as possible and spherical harmonics for the lowest ten l values are typically considered. In pseudopotential theory r_c is chosen much smaller. The region where we have chemical bonding is described by plane waves. High- l components in this small sphere are less important and typically the lowest three l components are sufficient for the nonlocal pseudopotential. Since the number of plane waves increases linearly with the volume of the valence region, the SFP pseudopotential method requires many more plane waves than the LAPW method. As has been discussed above, the separability of the Hamiltonian and overlap matrices allows us to use huge numbers of plane waves within the context of iterative diagonalization.

ACCURACY

The LAPW method is known to be very accurate for electronic-structure calculations. It is to be expected that the same accuracy can be obtained with the SFP pseudopotential. There are two principal sources of errors. The parametrization of the overlap and energy and the finite number of plane waves. We have checked the accuracy for the case of the hydrogen atom where all quantities can be calculated analytically. Using a plane-wave cutoff of 10 hartrees, an r_c of 1 a.u., and an energy window of 1 hartree, we have found that the repartition of charge be-

tween the core and valence region is given with an error of less than 1 ppm. A detailed analysis of the errors in the LAPW method is given in Refs. 4 and 5.

It is well known that so-called ghost eigenvalues may appear in a LAPW calculation. Within the LAPW method these ghost eigenvalues arise if a solution has a logarithmic derivative that is outside the region where the parametrization of the logarithmic-derivative versus energy curve is accurate. This problem has also been detected with separable pseudopotentials.⁸ A recent analysis⁹ shows that the separable Kleinmann-Bylander pseudopotential does not correctly reproduce the energy versus logarithmic-derivative function. Since the SFP pseudopotential was constructed under the requirement that it reproduce the energy versus logarithmic-derivative curve correctly, ghost eigenvalues cannot occur for eigenvalues in the allowed range.

RESULTS FOR OXYGEN

Results for the $2s$ and $2p$ wave function of the oxygen atom are shown in Fig. 2. The logarithmic-derivative versus energy function was parametrized by the curves shown in Fig. 3. The pseudo-wave-function in the core region is defined as the analytical continuation of the

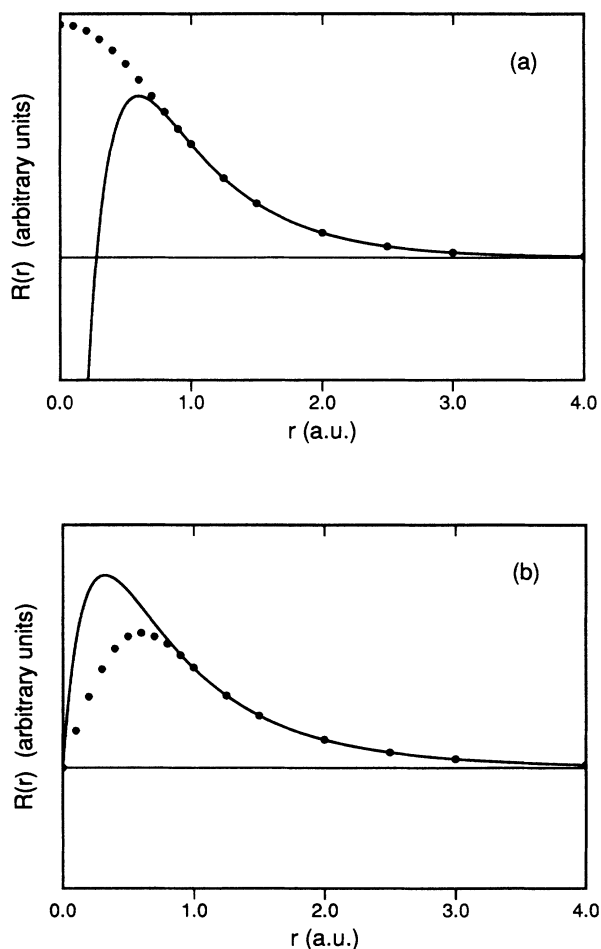


FIG. 2. True wave functions (solid line) and pseudo-wave-functions (solid circles) for oxygen. (a) $2s$ functions; (b) $2p$ functions.

plane-wave solution into the core region. The calculation was done in a cubic supercell of sidelength 24 a.u. with a plane-wave energy cutoff of 20 hartrees. This corresponds to roughly 130 000 plane waves.

CONSTRUCTION OF THE PSEUDOPOTENTIAL WITHIN A SELF-CONSISTENT LDA CALCULATION

Self-consistent pseudopotential calculations traditionally rely on the frozen-core approximations; i.e., it is supposed that the core electrons do not readjust their density distribution in response to changes in the valence distribution. This approximation is usually well obeyed¹⁰ and

we will therefore use it here. Fixing the positions of the nuclei, the total energy is given by

$$E = \sum_i \int \frac{1}{2} \nabla \psi_i^* \nabla \psi_i + \psi_i^* V \psi_i + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int e_{xc}(\rho(\mathbf{r})) d\mathbf{r},$$

where $V(r)$ is the potential of the nuclei. If we distinguish between the core functions ψ_{ic} and the valence functions ψ_{iv} , and between the corresponding core and valence densities ρ_c and ρ_v , we obtain

$$E = \sum_{ic} \int \frac{1}{2} \nabla \psi_{ic}^* \nabla \psi_{ic} + \psi_{ic}^* V \psi_{ic} + \frac{1}{2} \int \int \frac{\rho_c(\mathbf{r})\rho_c(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \sum_{iv} \int \frac{1}{2} \nabla \psi_{iv}^* \nabla \psi_{iv} + \psi_{iv}^* V \psi_{iv} + \frac{1}{2} \int \int \frac{\rho_v(\mathbf{r})\rho_v(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int \int \frac{\rho_v(\mathbf{r})\rho_c(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int e_{xc}(\rho_c(\mathbf{r}) + \rho_v(\mathbf{r})) d\mathbf{r}.$$

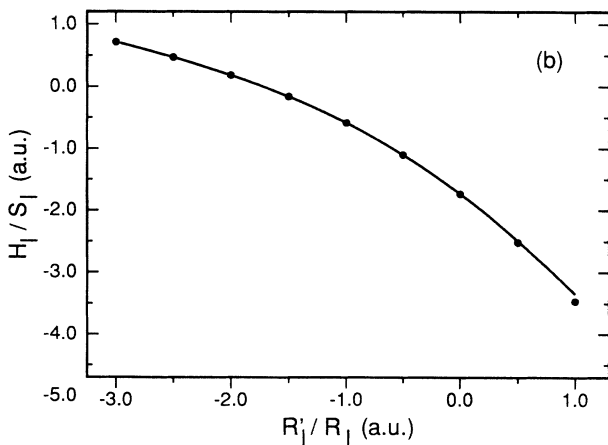
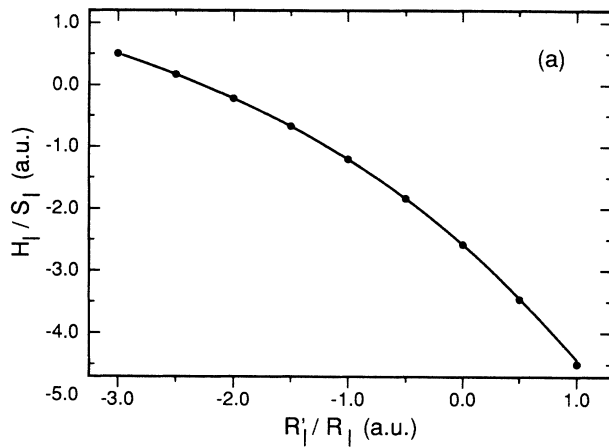


FIG. 3. Logarithmic derivative vs energy curves for the $l=0$ and $l=1$ components in oxygen. (a) Case $l=0$: Only the second branch is shown. The first branch is much lower in energy (at around -19). (b) Case $l=1$: The first branch is shown. Exact values are given by solid circles, the solid line is the parametrization.

The first two terms are constant in the frozen-core approximation and will therefore be dropped in the following. We now still split up the remaining Coulomb integrals into integrals over the core and the valence region:

$$\int \int \frac{\rho_v(\mathbf{r})\rho_v(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' = \int_v \int_v \frac{\rho_v(\mathbf{r})\rho_v(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int_c \int_c \frac{\rho_v(\mathbf{r})\rho_v(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + 2 \int_v d\mathbf{r} \rho_v(\mathbf{r}) \int_c \frac{\rho_v(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'.$$

Since we assume the valence charge in the core region to be spherical, we can write

$$\frac{Q_v}{r} = \int_c \frac{\rho_v(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}',$$

where Q_v is the valence charge in the core region. We apply the same decomposition for the core-valence Coulomb integral:

$$\int \int \frac{\rho_v(\mathbf{r})\rho_c(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' = \int_v d\mathbf{r} \rho_v(\mathbf{r}) \int \frac{\rho_c(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \int_c d\mathbf{r} \rho_v(\mathbf{r}) \int \frac{\rho_c(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'.$$

Because of the frozen-core approximation, the inner integral can be replaced by the fixed potential V_c :

$$V_c(r) = \int \frac{\rho_c(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'.$$

The total energy is now a sum of two terms, E_1 and E_2 , with

$$\begin{aligned}
E_1 &= \sum_{iv} \int_c \frac{1}{2} \nabla \psi_{iv}^* \nabla \psi_{iv} + \int_c \rho_v(\mathbf{r}) [V(\mathbf{r}) + V_c(\mathbf{r})] d\mathbf{r} \\
&\quad + \frac{1}{2} \int_c \int_c \frac{\rho_v(\mathbf{r}) \rho_v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
&\quad + \int_c e_{xc}(\rho_c(\mathbf{r}) + \rho_v(\mathbf{r})) d\mathbf{r} , \\
E_2 &= \sum_{iv} \int_v \frac{1}{2} \nabla \psi_{iv}^* \nabla \psi_{iv} + \int_v \rho_v(\mathbf{r}) \left[V(\mathbf{r}) + \frac{Q_v}{r} + V_c(\mathbf{r}) \right] d\mathbf{r} \\
&\quad + \frac{1}{2} \int_v \int_v \frac{\rho_v(\mathbf{r}) \rho_v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
&\quad + \int_v e_{xc}(\rho_v(\mathbf{r}) + \rho_c(\mathbf{r})) d\mathbf{r} .
\end{aligned}$$

Conventional pseudopotential calculations usually neglect the contribution of the core electrons to the exchange-correlation energy in the valence region, i.e., the exchange-correlation energy in the valence region is approximated by

$$\int_v e_{xc}(\rho_v(\mathbf{r})) d\mathbf{r} .$$

This approximation is often poor,¹¹ but it can easily be avoided within the context of the SFP pseudopotential.

The energy E_1 is a function of the values and deriva-

tives of the valence functions on the sphere. In an analogous way to the non-self-consistent case, we must again find a parametrization for E_1 . As we see, the unscreened ionic potential is never needed in the SFP pseudopotential method. The expression for E_2 can be handled in the usual way. The fact that the integrals in the expression for E_2 extend only over the valence region poses no problem. They can be replaced by integrals over the whole space, where the argument is multiplied by a Θ function that is zero in the core region and unity outside. In Fourier space this multiplication will merely introduce an additional convolution.

CONCLUSIONS

We have constructed a separable first-principles pseudopotential which has the following promising properties. (i) The separability of the Hamiltonian and overlap matrices allows the calculation of eigenstates with an efficiency that is comparable to the case of a local potential. (ii) The exact valence wave function is known in the core region. Therefore, physical phenomena that require the correct charge density near the core are accessible. (iii) The transferability and accuracy can systematically be improved.

¹G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

²In Figs. 1 and 3, H_i was actually calculated using $\int -rR_i(\partial^2/\partial r^2)(rR_i)dr$ instead of $\int (r\partial R_i/\partial r)^2 dr$ as in (7). The two integrals differ only by a surface term.

³If the system has extended core states (so-called semicore states), it is necessary to include them in the calculation. To describe the semicore states and the valence states, the energy window must be chosen large enough to include both. This can be achieved by a sufficiently small r_c .

⁴D. D. Koelling and G. O. Arbman, J. Phys. F **5**, 2041 (1975).

⁵O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

⁶E. R. Davidson, Comp. Phys. Commun. **53**, 49 (1989).

⁷J. L. Martins and M. L. Cohen, Phys. Rev. B **37**, 6134 (1988).

⁸L. Kleinmann and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

⁹X. Gonze, P. Kaeckell, and M. Scheffler, Phys. Rev. B **41**, 12 264 (1990).

¹⁰W. E. Pickett, Comp. Phys. Rep. **9**, 117 (1989).

¹¹S. G. Louie, S. Froyen, and M. L. Cohen, Phys. Rev. B **26**, 1738 (1982).