PHYSICAL REVIEW B VOLUME 41, NUMBER 11 15 APRIL 1990-I

Rapid Communications

The Rapid Communications section is intended for the accelerated publication of important new results. Since manuscripts submitted to this section are given priority treatment both in the editorial office and in production, authors should explain in their submittal letter why the work justifies this special handling. A Rapid Communication should be no longer than 3½ printed pages and must be accompanied by an abstract. Page proofs are sent to authors, but, because of the accelerated schedule, publication is not delayed for receipt of corrections unless requested by the author or noted by the editor.

Soft self-consistent pseudopotentials in a generalized eigenvalue formalism

David Vanderbilt
Department of Physics, Harvard University, Cambridge, Massachusetts 02138

(Received 23 January 1990)

A new approach to the construction of first-principles pseudopotentials is described. The method allows transferability to be improved systematically while holding the cutoff radius fixed, even for large cutoff radii. Novel features are that the pseudopotential itself becomes charge-state dependent, the usual norm-conservation constraint does not apply, and a generalized eigenproblem is introduced. The potentials have a separable form well suited for plane-wave solid-state calculations, and show promise for application to first-row and transition-metal systems.

The development of first-principles norm-conserving pseudopotentials by Hamann, Schlüter, and Chiang[1] (HSC) and others[2,3] has paved the way to accurate calculations of solid-state properties within the local-density approximation using plane-wave basis functions.[4] However, the utility of this approach to systems containing highly localized valence orbitals (e.g., for first-row and transition-metal atoms) has been limited, because of the difficulty of representing the pseudo-wave-functions in a plane-wave basis. The basis-set size can be reduced to some extent by constructions which insure optimal smoothness of the potential[5] or wave function[6] and by moving the cutoff radius outward. However, the norm-conserving condition[1] requires that the total pseudocharge inside the core match that of the all-electron (AE) wave function. Thus for many important cases, e.g., O $2p$ or Ni $3d$ orbitals, it has proven impossible to construct a pseudo-wave-function which is much smoother than the AE one.

These difficulties have been compensated to some extent by the development of iterative methods, which can handle very large plane-wave basis sets.[7,8] These methods are most efficient if the nonlocal part of the pseudopotential can be cast in a separable form. Some workers have taken the approach of expanding the "semilocal" (local in the radial variable only) HSC-type potential approximately as a sum of separable terms,[9,10] while Kleinman and Bylander[11] (KB) generate a fully nonlocal separable pseudopotential based on a given semilocal one.

Here, a new approach to the construction of first-principles pseudopotentials is described, in which a fully nonlocal pseudopotential is generated directly. It has the following desirable properties: (i) It takes the form of a sum of a few separable terms. (ii) It becomes local and vanishes outside the core. (iii) The scattering properties and their energy derivatives are, by construction, correct at several energies spanning the range of occupied states, and the transferability can be systematically improved by increasing the number of such energies. (iv) The norm-conserving constraint is removed so that the pseudo-wave-function can be constructed in such a way as to optimize smoothness. (v) The pseudopotential itself becomes involved in the self-consistent screening process, thereby improving transferability with respect to changes in charge configuration. Together, these features allow the cutoff radius to be increased without sacrificing transferability, even for "problem" cases such as $2p$ and d orbitals.

The construction of the new pseudopotentials will be described in three stages. In the first stage, I show that it is possible to arrive at a fully nonlocal KB-type pseudopotential by working with the wave function directly, bypassing the construction of a semilocal potential entirely. Moreover, this can be done at an arbitrary energy ε_i, as suggested by Hamann.[12] As usual, an AE calculation is carried out on a free atom in some reference configuration, leading to a screened potential $V_{\mathrm{AE}}(r)$. Cutoff radii r_{cl} and r_c^{loc} are chosen for the wave functions and local pseudopotential, respectively, and a diagnostic radius R is chosen large enough that all pseudo- and AE quantities agree at and beyond R. Some algorithm is used to generate a smooth local potential $V_{\mathrm{loc}}(r)$ which approaches $V_{\mathrm{AE}}(r)$ beyond r_c^{loc}. Now consider an AE wave function $\psi_i(\mathbf{r})$ of definite angular momentum lm, which is a solution of the Schrödinger equation, regular at the origin, at an arbitrary energy ε_i:

$$[T+V_{\mathrm{AE}}(r)]\psi_i(\mathbf{r})=\varepsilon_i\psi_i(\mathbf{r})\,. \tag{1}$$

Here i is a composite index, $i=\{\varepsilon_i lm\}$, T is the kinetic-energy operator $-\frac{1}{2}\nabla^2$, and V_{AE} is the original reference screened potential, i.e., ψ_i is not determined self-consistently. (Atomic units are used throughout.) Despite the fact that ψ_i is, in general, non-normalizable, I adopt a bracket notation

$$(T+V_{\mathrm{AE}}-\varepsilon_i)\,|\psi_i\rangle=0 \tag{2}$$

as a stand-in for the previous equation. Quantities such as $\langle\psi_i|\psi_i\rangle$ are ill-defined, but I shall make use of the special notation $\langle\psi_i|\psi_j\rangle_R$ to denote the integral of $\psi_i^*(\mathbf{r})\psi_j(\mathbf{r})$ inside the sphere of radius R.

Now a pseudo-wave-function ϕ_i is constructed, subject to the constraints that it join smoothly to ψ_i at r_{cl} and that it satisfy the norm-conserving property $\langle\phi_i|\phi_i\rangle_R=\langle\psi_i|\psi_i\rangle_R$. Since the wave function

$$|\chi_i\rangle=(\varepsilon_i-T-V_{\mathrm{loc}})\,|\phi_i\rangle \tag{3}$$

is local (it vanishes at and beyond R where $V_{\mathrm{AE}}=V_{\mathrm{loc}}$ and $\phi_i=\psi_i$), the nonlocal pseudopotential operator

$$V_{\mathrm{NL}}=\frac{|\chi_i\rangle\langle\chi_i|}{\langle\chi_i|\phi_i\rangle} \tag{4}$$

is well defined. It is straightforward to verify that $|\phi_i\rangle$ is an eigenvector of $T+V_{\mathrm{loc}}+V_{\mathrm{NL}}$, and that the scattering properties and their energy derivatives are correct at ε_i in the usual way (see below).

The second stage of the new pseudopotential scheme is arrived at by generalizing the previous construction to the case of *two or more* energies ε_i at which the scattering properties will be correct, as follows. For a given angular momentum l, some number (usually between one and three) of energies which span the energy of occupied states of a target (e.g., bulk crystal) calculation are chosen. Now the set of pseudo-wave-functions $|\phi_i\rangle$ are constructed from the AE wave functions $|\psi_i\rangle$ as before, except that they should satisfy the generalized norm-conserving condition $Q_{ij}=0$, where

$$Q_{ij}=\langle\psi_i|\psi_j\rangle_R-\langle\phi_i|\phi_j\rangle_R\,. \tag{5}$$

Forming the matrix $B_{ij}=\langle\phi_i|\chi_j\rangle$ and defining a set of local wave functions

$$|\beta_i\rangle=\sum_j(B^{-1})_{ji}\,|\chi_j\rangle\,, \tag{6}$$

which are dual to the $|\phi_i\rangle$, the nonlocal pseudopotential operator can be chosen as

$$V_{\mathrm{NL}}=\sum_{i,j}B_{ij}\,|\beta_i\rangle\langle\beta_j|\,. \tag{7}$$

Then it can easily be shown that $|\phi_i\rangle$ satisfies the secular equation $(H-\varepsilon_i)\,|\phi_i\rangle=0$, where $H=T+V_{\mathrm{loc}}+V_{\mathrm{NL}}$.

I now show that the matrix B_{ij}, and therefore the operator V_{NL}, are Hermitian when $Q_{ij}=0$. Taking $u_i(r)/r$ to be the radial wave function associated with $\phi_i(\mathbf{r})$,

$$B_{ij}=\int_0^R dr\,u_i^*(r)\left\{\varepsilon_j+\frac{1}{2}\frac{d^2}{dr^2}-\frac{l(l+1)}{2r^2}-V_{\mathrm{loc}}(r)\right\}u_j(r)\,. \tag{8}$$

The expression for B_{ji}^* is identical except that ε_j is replaced by ε_i, and the derivative d^2/dr^2 acts to the left. After one integration by parts on each expression,

$$B_{ij}-B_{ji}^*=(\varepsilon_j-\varepsilon_i)\langle\phi_i|\phi_j\rangle_R+\tfrac{1}{2}\,[u_i^*(R)u_j'(R)-u_i'^*(R)u_j(R)]\,. \tag{9}$$

A similar expression can be derived for the AE wave functions; subtracting this from (9), and noting that the pseudo- and AE wave functions and their derivatives match at R, one obtains

$$B_{ij}-B_{ji}^*=(\varepsilon_i-\varepsilon_j)Q_{ij}\,, \tag{10}$$

which vanishes when $Q_{ij}=0$.

Again one may verify that $(d\ln u/dr)_R$ and its energy derivative match the corresponding AE quantities *at each* ε_i. Thus, by increasing the number of energies ε_i at which the construction is done, the scattering properties of the pseudopotential can be made to reproduce those of the AE potential *with arbitrary accuracy* over the energy range of interest.[13]

One could stop here, and still have a useful scheme. However, I now show that the constraint $Q_{ij}=0$ is *unnecessary,* if one is willing to adopt a generalized eigenvalue formalism in which an overlap operator appears. In this third stage, I define a nonlocal overlap operator

$$S=1+\sum_{i,j}Q_{ij}\,|\beta_i\rangle\langle\beta_j|\,, \tag{11}$$

and redefine the nonlocal potential operator to be

$$V_{\mathrm{NL}}=\sum_{i,j}D_{ij}\,|\beta_i\rangle\langle\beta_j|\,, \tag{12}$$

where

$$D_{ij}=B_{ij}+\varepsilon_jQ_{ij}\,, \tag{13}$$

and Q_{ij} is as given in Eq. (5). Note that with these definitions,

$$\langle\phi_i|S|\phi_j\rangle_R=\langle\psi_i|\psi_j\rangle_R\,. \tag{14}$$

Then $|\phi_i\rangle$ is easily shown to be a solution of the generalized eigenvalue problem $(H-\varepsilon_iS)\,|\phi_i\rangle=0$. Now it follows from Eqs. (5), (10), and (13) that Q and D are Hermitian matrices, even though B is not. Thus H and S are Hermitian operators. Moreover, it follows from the identity

$$0=\left[\frac{d}{d\varepsilon}\langle\phi_\varepsilon|\,T+V_{\mathrm{loc}}+V_{\mathrm{NL}}-\varepsilon S\,|\phi_\varepsilon\rangle_R\right]_{\varepsilon=\varepsilon_i} \tag{15}$$

that

$$-\frac{1}{2}u_i^2\frac{d}{d\varepsilon}\frac{d}{dr}\ln u_\varepsilon(r)\bigg|_R=\langle\phi_i|\phi_i\rangle_R+Q_{ii}=\langle\psi_i|\psi_i\rangle_R\,, \tag{16}$$

so that the matching of the AE and pseudologarithmic derivatives follows in the usual way.

The relaxation of the constraint $Q_{ij}=0$ means that each ψ_i can be made into a pseudo-wave-function ϕ_i independently, with the *only constraint* being the matching of $\phi(r)$ to $\psi(r)$ at the cutoff radius. Thus it becomes possible to choose the cutoff radius to be *well beyond the radial wave-function maximum,* as illustrated in Fig. 1. A consequence of this freedom is that a generalized eigenvalue problem has to be solved in the target solid-state calculation. However, within iterative approaches to the eigenvector problem, the time-dominant step is the multiplication of $H-\varepsilon S$ by a trial vector $\phi_{n\mathbf{k}}$. In this case the operation count need hardly increase at all, because the identical form of the nonlocal parts of S and H allows them to be consolidated into a single operator. Incidentally, the current pseudopotential bears a formal resemblance to the original Phillips-Kleinman pseudopotential.[14] The latter can be cast in the form of Eqs. (11) and (12) (with the $|\beta_i\rangle$ being just the core orbitals), but does not have an adjustable cutoff radius.

In a self-consistent calculation, the "deficit" of valence charge in the core region associated with a pseudo-wave-function such as that of Fig. 1 will have to be restored. The solutions of the generalized eigenvalue problem should be normalized according to

$$\langle\phi_{n\mathbf{k}}|S|\phi_{n'\mathbf{k}}\rangle=\delta_{nn'}\,, \tag{17}$$

which is automatic in the usual methods of solution. Taken together with Eq. (14), Eq. (17) ensures that the pseudosolution has the same amplitude as the AE one at and beyond R. To make up the charge deficit, the valence charge density is defined to be

$$n_v(\mathbf{r})=\sum_{n,\mathbf{k}}\phi^*_{n\mathbf{k}}(\mathbf{r})\phi_{n\mathbf{k}}(\mathbf{r})+\sum_{i,j}\rho_{ij}Q_{ji}(\mathbf{r})\,, \tag{18}$$

where

$$\rho_{ij}=\sum_{n,\mathbf{k}}\langle\beta_i|\phi_{n\mathbf{k}}\rangle\langle\phi_{n\mathbf{k}}|\beta_j\rangle\,, \tag{19}$$

$$Q_{ij}(\mathbf{r})=\psi^*_i(\mathbf{r})\psi_j(\mathbf{r})-\phi^*_i(\mathbf{r})\phi_j(\mathbf{r})\,. \tag{20}$$

It follows from Eqs. (11) and (17) that $\int d^3r\,n_v(\mathbf{r})=N_v$ exactly, where N_v is the number of valence electrons in the unit cell.

In order to make a variational theory, the total energy

$$E_{\text{tot}}=\sum_{n,\mathbf{k}}\langle\phi_{n\mathbf{k}}|\left(T+V^{\text{ion}}_{\text{loc}}+\sum_{ij}D^{\text{ion}}_{ij}|\beta_i\rangle\langle\beta_j|\right)|\phi_{n\mathbf{k}}\rangle+E_H[n_v]+E_{\text{xc}}[n_v+n_c] \tag{21}$$

is to be minimized subject to the constraint (17). Here n_c is a frozen-core density included to improve transferability.[15] Defining

$$V_{\text{Hxc}}(\mathbf{r})=V_H^{[n_v]}(\mathbf{r})+V_{\text{xc}}^{[n_v+n_c]}(\mathbf{r})\,, \tag{22}$$

$$D^{\text{Hxc}}_{ij}=\int d^3r\,V_{\text{Hxc}}(\mathbf{r})Q_{ij}(\mathbf{r})\,, \tag{23}$$

the secular equation becomes

$$(T+V_{\text{loc}}+V_{\text{NL}}-\varepsilon_{n\mathbf{k}}S)|\phi_{n\mathbf{k}}\rangle=0\,, \tag{24}$$

with V_{NL} and S given by Eqs. (11)-(13), $V_{\text{loc}}=V^{\text{ion}}_{\text{loc}}+V_{\text{Hxc}}$, and $D_{ij}=D^{\text{ion}}_{ij}+D^{\text{Hxc}}_{ij}$. The $V^{\text{ion}}_{\text{loc}}$ and D^{ion}_{ij} must be obtained by unscreening the V_{loc} and D_{ij} of the generating atomic configuration in the usual way. Expressions for the Hellmann-Feynman forces and stresses will be given elsewhere. The dependence of D_{ij} upon n_v through V_{Hxc} implies that the pseudopotential itself must be updated as part of the self-consistent screening process.

In the limit that enough energies ε_i are chosen to reproduce the scattering properties over the entire energy range of occupied states, n_v approaches the exact AE valence density. Thus there is reason to expect that the pseudopotential calculation will match the AE one step for step during the self-consistent screening process. In practice, it will generally be necessary to replace the charge densities $n_c(r)$ (Ref. 15) and $Q_{ij}(r)$ by pseudo-versions below some relatively small cutoff radius r_0 determined by the ability to represent the *charge density* in the solid-state calculation.

Table I shows some transferability tests of the new pseudopotentials for the oxygen atom. The pseudopotential was generated in the ground s^2p^4 configuration using two reference energies $\varepsilon_1=\varepsilon_s^{\text{AE}}$ and $\varepsilon_2=\varepsilon_p^{\text{AE}}$ for each of $l=0$ and $l=1$, and $r_0=0.6$ a.u., $r_c^{\text{loc}}=1.0$ a.u., and $r_{cs}=r_{cp}=1.8$ a.u. (The pseudoization was done by a Kerker-like[2] procedure so that ϕ matches ψ exactly at and

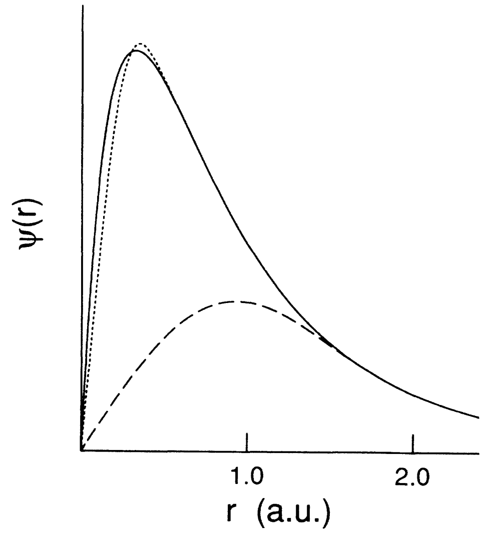

FIG. 1. Oxygen $2p$ radial wave function (solid line), and corresponding pseudo-wave-functions generated using HSC (dotted line) and current (dashed line) methods.

TABLE I. Valence s and p levels, and excitation energy ΔE_{tot} from the ground state, for the AE case and for pseudopotentials generated in the ground (s^2p^4) state using HSC and present methods. Units are in Ry.

State		AE	HSC	Present
s^1p^5	s	−1.7662	−1.7649	−1.7653
	p	−0.6981	−0.6982	−0.6979
	ΔE_{tot}	1.0658	1.0651	1.0654
s^0p^6	s	−1.7987	−1.7957	−1.7969
	p	−0.7262	−0.7261	−0.7256
	ΔE_{tot}	2.1361	2.1331	2.1348
s^2p^3	s	−2.8738	−2.8737	−2.8753
	p	−1.7909	−1.7904	−1.7928
	ΔE_{tot}	1.2066	1.2065	1.2075

beyond r_c.) Also shown are the results from a high-quality HSC pseudopotential using parameters ($r_{cs}=0.5$ a.u., $r_{cp}=0.35$ a.u., $\lambda=3.5$) very similar to those suggested by Bachelet, Hamann, and Schlüter.[16] The *p* pseudo-wave-functions are those shown in Fig. 1. It is clear from Table I that, in spite of the radically larger cutoff radii used in the new method, the overall quality of the two pseudopotentials is very similar. (Tests of logarithmic derivatives versus energy also indicate comparable energy transferability.) In fact, a pseudopotential of similar quality can be formed using $r_{cs}=r_{cp}=2.5$ a.u. if *three* reference energies are used. However, such a large r_c may be precluded by overlap of the core regions of neighboring atoms. (As a benchmark, the SiO_2 midbond distance is 1.5 a.u.; compare Fig. 1.)

In conclusion, it is hoped that the present method will allow pseudopotentials to be applied to first-row atom and transition-metal systems using modest plane-wave cutoffs for the first time.

Support for this work was provided by National Science Foundation Grant No. DMR-88-17291 and by the Alfred P. Sloan Foundation.

---

[1] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[2] G. Kerker, J. Phys. C **13**, L189 (1980).

[3] A. Zunger and M. L. Cohen, Phys. Rev. B **18**, 5449 (1978); **20**, 4082 (1979).

[4] W. E. Pickett, Comput. Phys. Rep. **9**, 115 (1989).

[5] D. Vanderbilt, Phys. Rev. B **32**, 8412 (1985).

[6] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990).

[7] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[8] R. Natarajan and D. Vanderbilt, J. Comput. Phys. **82**, 218 (1989).

[9] I. Štich, R. Car, M. Parrinello, and S. Baroni, Phys. Rev. B **39**, 4997 (1989).

[10] X. Gonze, J. P. Vigneron, and J.-P. Michenaud, J. Phys. Condens. Matter **1**, 525 (1989).

[11] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[12] D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).

[13] For an alternative approach, see E. L. Shirley, D. C. Allan, R. M. Martin, and J. D. Joannopoulos, Phys. Rev. B **40**, 3652 (1989).

[14] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959).

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

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