Generalized separable potentials for electronic-structure calculations

Peter E. Blöchl

IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 18 December 1989)

A general formalism is developed for the construction of separable potentials, either from a given pseudopotential or directly from the solutions of an atomic all-electron calculation. Such potentials are vital for the recently developed Car-Parrinello method. A subset of separable potentials, proposed by Kleinman and Bylander, has been used often with great success. However, for some atoms the previously used potentials do not reproduce the scattering properties over a sufficiently wide energy range. It is shown that the Kleinman-Bylander form represents the first term of a complete series. When additional terms are kept, the method yields highly accurate results.

The concept of nonlocal norm-conserving pseudopotentials developed by Hamann, Schlüter, and Chiang[1] ten years ago put the pseudopotential method for electronic structure calculations on firm theoretical grounds. In 1982, Kleinman and Bylander[2] proposed an alternative form for the nonlocal contributions, which leads to considerable savings while setting up the Hamiltonian. At that time, their method received only limited attention because such operations involved only a minor fraction of the computational effort for a complete self-consistent calculation. However, with the advent of the combined electronic structure and molecular-dynamics scheme of Car and Parrinello,[3] the Kleinman-Bylander (KB) form has experienced a revival. In contrast to traditional plane-wave pseudopotential calculations, the computational effort of the Car-Parrinello method is no longer dominated by matrix diagonalization, but considerable time is spent setting up the Hamiltonian and calculating total energies and forces. Therefore an efficient scheme for calculating nonlocal matrix elements, like the KB form, is absolutely essential to those calculations. However, it has recently been shown that the form for pseudopotentials chosen by Kleinman and Bylander can cause serious problems, e.g., the existence of spurious states.

In this paper we will develop a general framework for the construction of separable potentials.[4] We will show that there exists an exact transformation of any potential, local or nonlocal, into a separable form. In this context the KB potential represents the first term of the resulting series expansion. Even though such a truncation is often a valid and accurate approximation, in some cases the scattering properties of the atoms are not reproduced over a sufficiently wide energy range. Furthermore, we will describe how those problems can be cured by including a small number of additional terms into the truncated expansion.

State-of-the-art pseudopotential calculations are based on nonlocal norm-conserving pseudopotentials as proposed by Hamann, Schlüter, and Chiang,[1] which have the form

$$v_{ps}(\mathbf{r},\mathbf{r}') = v_{loc}(r)\delta(\mathbf{r}-\mathbf{r}')$$

$$+ \sum_{L}^{L_{max}} Y_L(\hat{\mathbf{r}})[v_l(r)-v_{loc}(r)]Y_L^*(\hat{\mathbf{r}}') . \quad (1)$$

Here Y_L are spherical harmonics and L denotes a combined angular momentum quantum number (l,m). The potentials $v_l(r)$ are chosen such that the resulting wave functions for the valence bands have zero nodes, but still reproduce the scattering properties of the full potential correctly. The choice of the local potential v_{loc} is arbitrary and affects only the higher partial waves with $l > l_{max}$. We will be mostly concerned with the second term in Eq. (1), namely the nonlocal potential. However, our discussion will also apply to completely general potentials. In the following we denote the nonlocal potential operator simply by v.

In order to construct the separable form we must first choose a complete but otherwise arbitrary set of functions $|\phi_i\rangle$. To be specific, we may choose a set of atomic wave functions or a complete set of partial waves[5] (which are not restricted to natural boundary conditions the way atomic wave functions are). Then the potential can be transformed exactly into the following form:

$$v = \sum_{ij} |v\phi_i\rangle U_{ij} \langle \phi_j v| , \quad (2)$$

where the matrix U satisfies the equation,

$$\sum_k U_{ik}\langle \phi_k | v | \phi_j\rangle = \delta_{ij} . \quad (3)$$

The atomic wave functions ϕ_i contain both a radial and an angular dependence. The validity of Eq. (2) can be easily shown by applying the right-hand side of Eq. (2) to an arbitrary wave function $|\psi\rangle$, which is expanded into the atomic wave functions, i.e., $|\psi\rangle = |\phi_i\rangle a_i$,

$$\sum_{ijk} |v\phi_i\rangle U_{ij}\langle \phi_j v| |\phi_k\rangle a_k = \sum_i v|\phi_i\rangle a_i = v|\psi\rangle . \quad (4)$$

The next step towards a separable form is the transformation of the atomic wave functions into a new set of functions $|\bar{\phi}_i\rangle$ for which the matrix $\langle \bar{\phi}_i | v | \bar{\phi}_j\rangle$ and hence the corresponding matrix $\bar{U}$ is diagonal. This can be achieved by means of a Gramm-Schmidt orthogonalization procedure, with a norm defined by $\langle f | v | g\rangle$ for functions f and g. We are led to a recursive equation for the

new set of functions[6] $|\bar{\phi}_i\rangle$;

$$|\bar{\phi}_i\rangle = |\phi_i\rangle - \sum_{j}^{i-1} |\bar{\phi}_j\rangle \frac{1}{c_j}\langle\bar{\phi}_j|v|\phi_i\rangle ,$$ (5)

$$c_i = \langle\bar{\phi}_i|v|\bar{\phi}_i\rangle .$$ (6)

This new set leads immediately to the general form of the separable potential

$$v = \sum_i |v\bar{\phi}_i\rangle \frac{1}{c_i}\langle\bar{\phi}_i v| .$$ (7)

The separable potentials are usually nonlocal not only in the angular dependence, but also in the radial dependence. They even represent a more general class of potentials than those of Eq. (1). Therefore, the question arises whether separable potentials can be constructed directly from an all-electron atomic calculation without first constructing a potential which is local in the radial dependence. This is indeed the case. We may choose rather arbitrarily a set of pseudo wave functions $|\phi_i\rangle$ which match the all-electron wave functions at a set of energies ε_i outside some radius. Second, we choose, again rather arbitrarily, a local potential, which matches the true atomic potential outside the same radius. Now we construct the functions $|v\phi_i\rangle$ by

$$|v\phi_i\rangle = [\varepsilon_i - (-\tfrac{1}{2}\nabla^2 + v_{\text{loc}})]|\phi_i\rangle .$$ (8)

From this point on the construction of the separable potential proceeds as outlined before.

In order to make the separable potentials a useful tool, we must truncate the series and understand the resulting errors. We note the following: Analogous to Eq. (4), it can be shown that if the summation in Eq. (4) is restricted to the lowest n atomic wave functions, then the corresponding eigenvalues and wave functions obtained from the separable form are correct. Furthermore, the energy derivatives of the logarithmic derivatives at those energies are correct as a consequence of the condition of norm conservation. Every wave function which is orthogonal to all $|v\phi_i\rangle$, where ϕ_i is one of those n lowest atomic wave functions, does not feel the nonlocal contribution at all.

This is the key to the understanding of the spurious states found in pseudopotential calculations using KB potentials. The KB potentials form a subset of separable potentials and are constructed from only one atomic bound state. If the local potential alone is much deeper than local and nonlocal parts together, there is the possibility of a bound state of the local part with one node in the region of the occupied states. Such a state is likely to be almost orthogonal to $|v\bar{\phi}_i\rangle$, where ϕ_i was used in constructing the KB potential and has no nodes. Hence, this state will persist because it barely experiences the nonlocal contribution. This will result in an inversion of the node sequence. Even though this scenario describes the extreme case, the same effect will usually result in a serious degradation of transferability.

The cure to this problem is to explicitly include more terms in the series expansion for the separable potential. A spurious state with one node will then experience the nonlocal potential through the added terms and therefore be shifted upwards in energy, out of the region of interest.

The inclusion of more terms in the expansion of the separable potential makes it possible to explicitly test and control the accuracy of the separable form.

We shall now verify our predictions with a practical example. Because of the increasing interest in III-V semiconductors, we investigate the pseudopotential for arsenic. In particular we will study the scattering properties of the s potential, while using the d component as a local potential. The pseudopotential is created within the framework of generalized norm-conserving pseudopotentials (GNCPP) by Hamann.[5] The cutoff radii are 0.9 and 1.8 a.u. for s and d potentials, respectively.

The scattering properties are best reflected in the logarithmic derivative function of the energy-dependent partial waves, which are shown in Fig. 1 with several levels of approximation. The original pseudopotential has a bound state at -0.54 a.u. The crossover from zero to one node happens just above the bound-state energy at -0.37 a.u. States with two nodes appear only above an energy of 2 a.u.

The separable potential with only one term, i.e., the KB potential, reproduces value and derivative of the logarithmic derivative at the bound-state energy, which was used in the construction of the separable potential. Also, the bound-state wave function is described correctly [Fig. 2(b)]. However, the next bound state appears already less than 0.5 a.u. above the lowest bound state. This can be seen from the singularity of the logarithmic derivative at 0 a.u., which reflects the crossover to states with two nodes [Figs. 2(c) and 2(d)]. As a consequence of this, the logarithmic derivative deviates strongly from the correct behavior if the energy moves away from the lowest bound-state energy, i.e., the transferability of the separable potential with only one term is disastrous.

If we, as outlined above, introduce a second term to the separable potential, we find the transferability greatly
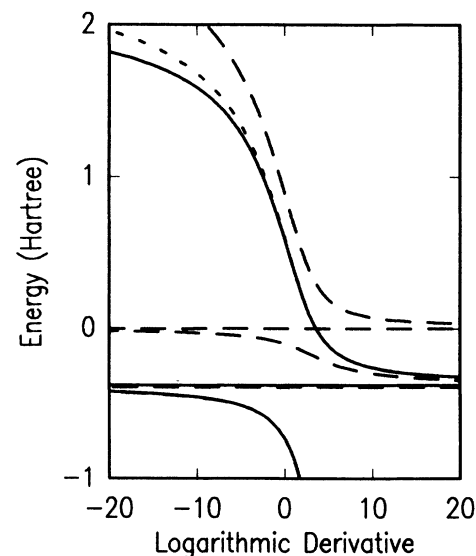

FIG. 1. Logarithmic derivative function of an As pseudopotential in the original form (solid), and the corresponding separable potentials truncated after the first (long dashes) and the second (short dashes) term.
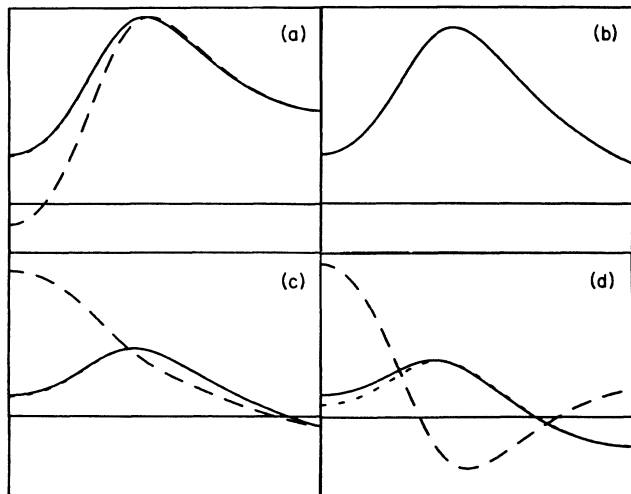
FIG. 2. S-type partial waves of an As pseudopotential in the original form (solid), and the corresponding separable potentials truncated after the first (long dashes) and the second (short dashes) term. The radial distance from the center is shown on the horizontal axis and the amplitude on the vertical axis. The energies are (a) -0.7 a.u., (b) -0.54 a.u., (c) -0.25 a.u., and (d) 0.5 a.u. Partial waves which are not visible fall on top of the correct solution shown as solid lines.

enhanced, and the logarithmic derivative of the separable potential coincides with that of the original potential over a wide energy range. This second term is derived from a partial wave 0.01 Hartree above the bound-state energy. A partial wave with a higher energy would be an even better choice for the wide energy range shown in Fig. 1.

Even though this example does not yet show an inversion of the node sequence as suggested previously, we have actually found such an inversion for a one-dimensional model potential.

The transferability of the KB potential depends somewhat on the choice of cutoff radii and usually some improvement can already be achieved by simply choosing a softer pseudopotential. However, in all cases a large improvement of the transferability has been observed when the second term is included in the separable form. Our feeling is that two terms in the separable potential will already be sufficient for all relevant cases.

In order to optimize the transferability in the energy range of interest, a good choice for the function $|\phi_i\rangle$ is one partial wave and its energy derivatives. This choice is equivalent to choosing a set of partial waves from a small energy range. The first energy derivative of a partial wave can be easily obtained either numerically or by solving the

inhomogeneous Schrödinger equation:

$$(-\tfrac{1}{2}\nabla^2 + v_{\mathrm{loc}} + v)\,|\dot{\phi}(\varepsilon)\rangle = |\phi(\varepsilon)\rangle. \tag{9}$$

Here $\phi(\varepsilon)$ denotes a partial wave obtained by outward integration of the Schrödinger equation at the energy ε and $\dot{\phi} = \partial\phi/\partial\varepsilon$ is the corresponding energy derivative.

It is worth mentioning that the partial waves of truncated separable potentials frequently show a peculiar behavior. In our numerical studies we find states with one or more nodes below the states with zero nodes [Fig. 2(a)]. However, we have not observed any bound states corresponding to those states and the logarithmic derivative function is not affected by the appearance of more nodes. The reason is that those nodes do not cross the sphere surface but are pulled in from the center. Hence they do not affect the behavior of the partial waves in the outer region of the atom which is relevant for the electronic properties. This behavior is interesting insofar as that the number of nodes alone no longer uniquely determines the main quantum number. In addition, it may have consequences for the plane-wave convergence of calculations using separable potentials truncated after the first term. The reason for this is that states with more nodes are harder to describe by a plane-wave expansion and their position, therefore, strongly depends on the plane-wave cutoff. If a node enters from the center even in the region of the valence states, this will result in a relatively hard pseudopotential. This effect, too, is greatly reduced if a second term is included in the separable form [Fig. 2(a)], and from our studies we expect once again, that two terms in the separable potential will be sufficient for all practical purposes.

In conclusion, we have presented a general framework for the construction of separable potentials from a given arbitrary potential or directly from the solutions of an all-electron atomic calculation. This concept generalizes the potential of Kleinman and Bylander in a natural way. We have pointed out problems which occur in connection with the KB potentials, and have shown how these problems can be easily cured within the framework of separable potentials. In addition, this framework provides an efficient way to control and improve the accuracy of truncated separable potentials such as the KB potentials.

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

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

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

[4]It came to our attention that D. Vanderbilt has recently independently developed a similar concept of separable potentials.

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

[6]If any function $\bar{\phi}_i$ vanishes in Eq. (5) it shall be excluded from the set.