

Trial functions in pseudopotential theory

R. D. Zwicker

Department of Physics, University of the Witwatersrand, Johannesburg, South Africa

(Received 26 February 1975)

It is shown that the use of variational trial functions with non-Hermitian pseudopotentials can lead to energies which do not represent accurate upper bounds to the true energy. Such difficulties can be resolved in principle by choosing a trial function which satisfies the same criterion used to define the pseudopotential. Accurate results can be obtained more simply by direct use of the Phillips-Kleinman pseudopotential, which minimizes the difference between the trial function and the exact pseudo wave function.

I. INTRODUCTION

The pseudopotential theory¹⁻⁵ provides a valuable technique for studying and interpreting the properties of valence electrons in atoms, molecules, and crystals. This method takes full advantage of one's knowledge of the core states ψ_c to simplify the search for valence eigenfunctions ψ and energies E satisfying the wave equation

$$(T + V)\psi = E\psi. \quad (1)$$

By defining a function $\phi = \psi + \sum_c \alpha_c \psi_c$, where the parameters $\alpha_c = (\psi_c, \phi)$ are arbitrary constants, one obtains the Phillips-Kleinman (PK) equation

$$T\phi + V\phi + \sum_c (E - E_c)(\psi_c, \phi)\psi_c = E\phi. \quad (2)$$

We assume here that the functions ψ_c and ψ form an orthonormal set, and that the original Hamiltonian $H = T + V$ is Hermitian. The second and third terms of Eq. (2) are taken together to form the pseudopotential V_p .

The cancellation between the attractive and repulsive contributions to the pseudopotential is cited to justify various simple models for atomic and solid-state systems. This cancellation was examined in detail by Cohen and Heine² (CH), who noted that since the coefficients α_c are arbitrary one is allowed to impose further constraints on ϕ , such as the requirement that ϕ be a smooth function. We can write such constraints in the form

$$(E - E_c)(\psi_c, \phi) = (F_c, \phi), \quad (3)$$

where F_c is chosen to suit the conditions imposed. Then Eq. (2) becomes

$$T\phi + V\phi + \sum_c (F_c, \phi)\psi_c = E\phi. \quad (4)$$

A further simplification is achieved if ϕ is smooth enough to be regarded as constant over the core region, and hence can be taken out of integrals whose integrands are negligible outside the core.

The pseudopotential theory was originally developed in the context of the orthogonalized plane

wave (OPW) method of Herring,⁶ and has since been used to further simplify energy-band calculations.⁷ It has also been applied successfully to problems (e.g., defect-center calculations⁸⁻¹⁰) for which plane waves are not appropriate to represent the pseudo wave function ϕ . The variational method with adroitly chosen trial functions provides a convenient alternative approach to the solution of Eq. (4). However, some difficulties arise in this method due to the possibly non-Hermitian nature of the pseudopotential. In Sec. II we examine the accuracy of the variationally determined energy and obtain a general expression for the energy correction in terms of the difference Δ between the trial function ϕ_t and the true pseudo wave function. This result is applied in Sec. III to the specific cases of the CH and Austin pseudopotentials, where it is found that the energy correction is generally of first order in Δ . In Sec. IV we show that better energies can be obtained by direct use of the PK pseudopotential, which adjusts itself to accommodate the trial function in such a way as to minimize the function Δ . Section V is reserved for concluding remarks.

II. ENERGY CORRECTION

In applying the variational method we first select a trial function ϕ_t which we suppose will yield a reasonable approximation to the true pseudo wave function ϕ . The trial function is then used to evaluate an approximate energy

$$E_t = (\phi_t, H_p \phi_t) / (\phi_t, \phi_t), \quad (5)$$

where H_p is the pseudo Hamiltonian of Eq. (4).

In practice, the trial function never has the exact form of the true pseudo wave function, so that ϕ_t will always differ from ϕ by some presumably small amount.¹¹ We define this quantity as

$$\Delta = \phi_t - \phi. \quad (6)$$

For a Hermitian Hamiltonian it is easily shown¹² that the variationally determined energy E_t differs from the true energy by a positive amount of sec-

ond order in Δ . However, this does not hold for the generally non-Hermitian pseudo Hamiltonian H_p . For this case we wish to calculate the energy correction $E_t - E$ in terms of the difference Δ . For simplicity in what follows we will regard all quantities as real. This does not affect our conclusions, but it shortens the equations considerably. We will also assume that the pseudo wave function ϕ is normalized to unity.

To obtain an expression for the energy correction we first note that the function F_c can depend on ϕ and define the difference function

$$\Delta F_c = F_c^t - F_c,$$

where F_c^t is evaluated with respect to the trial function. Then using the definition (6) in Eq. (5) and noting that the original Hamiltonian H is Hermitian, we arrive at an equation for E_t valid to first order in Δ :

$$\begin{aligned} E_t = & \left((\phi, H\phi) + 2(\Delta, H\phi) + \sum_c (F_c, \phi)(\phi, \psi_c) \right. \\ & + \sum_c (F_c, \Delta)(\phi, \psi_c) + \sum_c (F_c, \phi)(\Delta, \psi_c) \\ & \left. + \sum_c (\Delta F_c, \phi)(\phi, \psi_c) \right) [1 - 2(\Delta, \phi)]. \end{aligned}$$

This expression can be greatly simplified by noting that

$$2(\Delta, H\phi) + 2 \sum_c (F_c, \phi)(\Delta, \psi_c) = 2E(\Delta, \phi).$$

Then collecting terms of like order gives us the final result for the energy correction,

$$\begin{aligned} E_t - E = & \sum_c [(F_c, \Delta)(\phi, \psi_c) - (F_c, \phi)(\Delta, \psi_c)] \\ & + \sum_c (\Delta F_c, \phi)(\phi, \psi_c). \end{aligned} \quad (7)$$

This quantity is generally nonzero, and there is no clear method for determining its sign. The existence of a first-order energy discrepancy stems from the fact that the trial function does not in general satisfy the criteria used to derive the pseudopotential. We can remedy this situation by choosing ϕ_t such that

$$(E_t - E_c)(\psi_c, \phi_t) = (F_c^t, \phi_t). \quad (8)$$

In this case we can show that the first-order correction vanishes by subtracting this equation from Eq. (3) and substituting the resulting first-order expression into Eq. (7). Although this choice of ϕ_t is simple in principle, it can lead to some dif-

ficulties in practice. Concrete examples are discussed in Sec. III.

III. CH AND AUSTIN PSEUDOPOTENTIALS

The CH pseudopotential² is derived by requiring that ϕ be the smoothest possible pseudo wave function according to the criterion that $\int |\nabla \phi|^2 d\tau / (\phi, \phi)$ be a minimum. The pseudopotential is then defined by

$$F_c = (\bar{V}_p - V)\psi_c,$$

where $\bar{V}_p = (\phi, V_p \phi) / (\phi, \phi)$.

In this case F_c depends on ϕ through \bar{V}_p , so that $\Delta F_c = \Delta \bar{V}_p \psi_c$ is of first order. Then the energy correction from Eq. (7) can be written

$$\begin{aligned} E_t - E = & - \sum_c [(\psi_c, V\Delta)(\phi, \psi_c) - (\psi_c, V\phi)(\Delta, \psi_c)] \\ & + \Delta \bar{V}_p \sum_c |(\psi_c, \phi)|^2. \end{aligned} \quad (9)$$

If ϕ and Δ are smooth enough to be assumed constant over the core region, the first sum in Eq. (9) vanishes, but the second remains. Hence, for the CH pseudopotential the energy correction is of first order in Δ . Furthermore, this correction is not necessarily a positive quantity, so that the variationally determined E_t does not assure an upper bound on the true energy.

As mentioned in Sec. II, the first-order contribution to $E_t - E$ can be made to vanish if we use a trial function which satisfies the same criterion applied in deriving the pseudopotential. We can accomplish this by calculating a new trial function $\chi_t = \phi_t + \sum_c b_c \psi_c$, where the b_c are adjusted so that χ_t satisfies the CH smoothness criterion. Requiring that $\int |\nabla \chi_t|^2 d\tau / (\chi_t, \chi_t)$ be a minimum leads to a set of nonlinear equations in the parameters b_c which can usually be solved by an iterative process. The resulting χ_t satisfies Eq. (8), and the variational energy E_t evaluated with respect to χ_t then differs from the true energy by terms of second order in Δ . Although this procedure is straightforward in application, the increased numerical difficulties involved tend to limit its usefulness.¹³

The Austin pseudopotential³ is also obtained by a smoothness condition on ϕ . In this case one requires that $\int |\nabla \phi|^2 d\tau$ be a minimum, and the resulting pseudopotential is defined by $F_c = -V\psi_c$. Then $\Delta F_c = 0$, and $E_t - E$ is given by the first sum in Eq. (9). Again the energy correction is of first order in Δ , but if the assumption of constant ϕ and Δ is valid, then $E_t - E$ becomes of second order. It can also be shown that if the trial function is smoother than the exact pseudo wave function in the sense that $|E(\psi_c, \phi_t)| > |(\psi_c, T\phi_t)|$, then $E_t > E$. Thus, to the extent that ϕ_t and Δ are smooth functions, the variational method may give an accurate

upper bound to the energy E . However, the Austin pseudopotential does not lead to the smoothest possible ϕ , and the assumption that Δ is constant over the core region is not necessarily justifiable.

As in the previous case, it is again possible to obtain second-order accuracy by calculating a new trial function which satisfies the Austin smoothness criterion. This criterion is somewhat easier to apply since the resulting equations are linear in the coefficients b_c . Again, however, some difficulty is encountered in evaluating the energy with respect to such a trial function.

IV. PK PSEUDOPOTENTIAL

Both the CH and Austin pseudopotentials are obtained by requiring that the pseudo wave function be in some sense smooth. Hence the variational energy E_t approximates the true energy only to the extent that ϕ_t approximates the unique ϕ which satisfies Eq. (4). Instead of specifying ϕ by a smoothness condition, it would appear more advantageous in this instance to define the pseudopotential by requiring that ϕ be that unique pseudo wave function which most closely approximates the trial function. That is, we seek to minimize the difference function Δ by requiring that

$$\delta \int |\phi_t - \phi|^2 d\tau = 0. \quad (10)$$

Solving this equation gives us a set of uniquely defined parameters a_c given by

$$a_c = (\psi_c, \phi_t). \quad (11)$$

This apparently trivial result has an important interpretation. The PK pseudopotential, when used with a variational trial function, adjusts itself to accommodate the trial function in such a way that the exact solution ϕ is that unique ϕ which most closely approximates the trial function. We can clarify this point somewhat by rewriting the PK equation with the parameters a_c defined as above:

$$T\phi + V\phi + \sum_c (E - E_c)(\psi_c, \phi_t)\psi_c = E\phi. \quad (12)$$

This equation has the same form as Eq. (2), but it has a different meaning in that the quantities (ψ_c, ϕ_t) are now uniquely defined. Although ϕ_t is not an exact solution to Eq. (12), the pseudopotential used is nevertheless a valid one, and the true solution is that unique ϕ which satisfies the criterion (10). Also, it is easily shown that the difference $E_t - E$ is of second order in Δ , and that $E_t > E$ (assuming the core states are accurately known), so that the variational energy gives an upper bound to the true energy.

The ability of the pseudopotential to adjust to the trial function may also allow us to simplify the calculations by giving ϕ_t whatever desirable proper-

ties we can reasonably justify. For example, if it is certain that a smooth ϕ exists and we wish to approximate ϕ by a constant in the core region, we need only choose a reasonably flexible, smooth ϕ_t to ensure that the corresponding ϕ is smooth. Other simplifying properties besides smoothness might also be assigned to ϕ_t if we can show that an exact ϕ should exist which has those properties. However, if such approximations are to be used in a variational calculation, care must be taken to ensure that in minimizing the energy we are not in fact maximizing the error in the approximations.

We can write the pseudo-Hamiltonian of Eq. (12) in the form of a linear homogeneous operator by making explicit our assumptions concerning the norms of ϕ and ϕ_t . In order to display homogeneity we allow the norm of ϕ to vary by letting $\phi = \alpha_v \psi_v + \sum_c a_c \psi_c$, where α_v is arbitrary. We then write the trial function as $\phi_t = N\phi_{tn}$, where ϕ_{tn} is the function we actually choose and it has whatever norm we assign to it—unity, for our purposes. Adjusting N to satisfy the criterion (10) then gives $N = (\phi_{tn}, \phi)$, and Eq. (12) becomes

$$T\phi + V\phi + \sum_c (E - E_c)(\psi_c, \phi_{tn})(\phi_{tn}, \phi)\psi_c = E\phi.$$

The solution ϕ is then unique within a multiplicative constant, and it is that solution which minimizes Δ according to Eq. (10) when the norms are properly chosen.

V. CONCLUSIONS

Due to the non-Hermiticity of the general pseudopotential of Eq. (4), variationally determined energies generally differ from the true energy by an amount of first order in Δ and of indeterminate sign. Second-order accuracy can be obtained by requiring the trial function to satisfy the same criteria (3) that hold for the true pseudo wave function. However, this increases considerably the numerical difficulties involved in evaluating the energy.

The indeterminacy of the PK pseudopotential is a problem in principle only. If a trial function is used to approximate the pseudo wave function, then the pseudopotential is uniquely defined by the trial function.¹⁴ Direct use of the PK pseudopotential then has two important advantages over the more restricted pseudopotentials:

(i) The variational energy is greater than the true energy by a positive amount of second order in the minimized quantity Δ ; and

(ii) the pseudo wave function ϕ which ϕ_t must approximate is that unique ϕ which most closely reproduces ϕ_t .

Added to this is the fact that the PK equation is generally easier to solve because it avoids matrix elements of the type $(\psi_c, V\phi)$.

If the core states are not accurately known, an

upper bound on the energy can still be obtained by a method suggested by Brown.⁴ In this approach the orthogonality conditions $(\psi_v, \psi_c) = 0$ are omitted and the parameters a_c are determined variationally. The resulting secular determinant will yield not only a true upper bound to the energy E but improved core states as well.

The self-adjusting property of the PK pseudopotential can also be used to explain the success of OPW calculations from a pseudopotential point of view. In this case the pseudo wave function is

expanded in terms of a few plane waves of low wave number and good results are obtained in spite of the fact that no smoothness condition is invoked to define the pseudopotential. In the light of the discussion of Sec. IV, we see that the act of selecting the plane waves in itself defines the pseudopotential in the most advantageous way possible.

ACKNOWLEDGMENT

The author would like to acknowledge the helpful comments of Professor E. Brown.

¹James C. Phillips and Leonard Kleinman, Phys. Rev. 116, 287 (1959).

²Morrel H. Cohen and V. Heine, Phys. Rev. 122, 1821 (1961).

³B. J. Austin, V. Heine, and L. J. Sham, Phys. Rev. 127, 276 (1962).

⁴E. Brown, Phys. Rev. 126, 421 (1962).

⁵J. Rubio and F. Garcia-Moliner, Proc. Phys. Soc. Lond. 91, 739 (1967).

⁶C. Herring, Phys. Rev. 57, 1169 (1940).

⁷W. A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).

⁸R. H. Bartram, A. M. Stoneham, and Philip Gash,

Phys. Rev. 176, 1014 (1968).

⁹W. Weber and B. G. Dick, Phys. Status Solidi 36, 723 (1969).

¹⁰A. H. Harker, J. Phys. C 6, 2993 (1973).

¹¹We assume that the ratio of norms of ϕ and ϕ_t is chosen to minimize this difference by some criterion.

¹²C. Schwartz, Ann. Phys. (N. Y.) 2, 170 (1959).

¹³Since χ_t satisfies Eq. (8), this approach is analytically equivalent to the simpler procedure of Sec. IV.

¹⁴The pseudopotential is "uniquely defined" in the sense that the F_c of Eq. (4) are chosen such that Eq. (3) is not trivially satisfied.