

Orbital constraints in the application of chemical pseudopotentials

Jayanta Kumar Rudra and W. Beall Fowler

*Department of Physics and Sherman Fairchild Laboratory, Lehigh University,
Bethlehem, Pennsylvania 18015*

(Received 3 December 1982; revised manuscript received 4 April 1983)

We have developed a modification of Anderson's chemical pseudopotential (CP) which appears to be more useful in tight-binding analyses of energy bands. We have also shown that the Anderson CP implies symmetry constraints on the orbitals which, if ignored, can lead to serious errors, especially in low-symmetry systems.

I. INTRODUCTION

Since its introduction by Anderson¹ in 1969, the chemical pseudopotential (CP) has been used both in explicit calculations^{2,3} and as a means of justifying⁴ simpler (e.g., tight-binding^{5,6}) procedures for describing energy bands. In the course of a tight-binding analysis of the valence bands of a low-symmetry solid⁷ (SiO₂, in several structures), we discovered apparent inconsistencies in applying CP concepts. We have resolved these by recognizing the existence of symmetry constraints which a given CP imposes upon the orbitals associated with that pseudopotential. We have also developed a modification of Anderson's pseudopotential¹ which appears to be more useful as a starting point for tight-binding analyses in that the constraints are less severe.

II. THEORY

The CP method is an outgrowth of the study of the Hartree-Fock equations in terms of localized orbitals. It was shown by Adams⁸ and by Gilbert⁹ that the Hartree-Fock equations may be expressed in the general form

$$(F + PDP) |A_i\rangle = \epsilon_{A_i} |A_i\rangle. \quad (1)$$

Here, the Fock operator F may be written as

$$F = T + V, \quad (2)$$

where T and V are one-electron kinetic and potential energy operators, respectively (the latter including Coulomb and exchange terms), D is an arbitrary one-electron operator, and $|A_i\rangle$ is the i th orbital localized on (or at least associated with) site A . The $|A_i\rangle$ are linear combinations of occupied Hartree-Fock orbitals; that is, they are defined over the Hartree-Fock manifold.⁹ The density operator P is given by

$$P = \sum_{B,j} \sum_{C,k} |B_j\rangle S_{B_j, C_k}^{-1} \langle C_k|, \quad (3)$$

where S^{-1} is the inverse overlap matrix defined by

$$S^{-1}S = 1, \quad (4)$$

and

$$S_{A_i, B_j} \equiv \langle A_i | B_j \rangle. \quad (5)$$

P has projection properties, e.g.,

$$P |A_i\rangle = |A_i\rangle. \quad (6)$$

The arbitrariness of the operator D has led various authors to make different choices in an attempt to localize wave functions or to otherwise simplify the one-electron equations. A common choice, for example, is $D = -U_A$, where U_A is the part of the potential energy V associated with the sites other than A ; that is, it is an "environment" potential. In terms of U_A , F may be written as

$$F = T + V_A + U_A, \quad (7)$$

where V_A is the potential energy associated with site A ; thus

$$U_A = V - V_A \equiv \sum_{B (\neq A)} V_B^A. \quad (8)$$

The Hartree-Fock equations then become, in their modified form,

$$[T + V_A + (U_A - P U_A P)] |A_i\rangle = \epsilon_{A_i} |A_i\rangle, \quad (9)$$

and indeed some of the environment potential U_A is subtracted off, presumably making $|A_i\rangle$ somewhat atomic-like.

Kunz¹⁰ pointed out that D could be chosen as

$$D = -P_A U_A P_A, \quad (10)$$

where

$$P_A \equiv \sum_k |A_k\rangle \langle A_k|. \quad (11)$$

Because $PP_A = P_A P = P_A$, this becomes

$$[T + V_A + (U_A - P_A U_A P_A)] |A_i\rangle = \epsilon_{A_i} |A_i\rangle \quad (13a)$$

or

$$\left[T + V_A + \sum_{B \neq A} (V_B^A - P_A V_B^A P_A) \right] |A_i\rangle = \epsilon_{A_i} |A_i\rangle. \quad (13b)$$

Anderson¹ showed that the basic equation [Eq. (1)] could be written in non-Hermitian form, since $P |A_i\rangle = |A_i\rangle$:

$$(F + PD) |A_i\rangle = \epsilon_{A_i} |A_i\rangle. \quad (14)$$

Anderson then formed a pseudopotential (still exact¹⁰) which arises by choosing

$$D = - \sum_{B (\neq A)} P_B V_B^A. \quad (15)$$

This yields

$$\left[T + V_A + U_A - \sum_{B (\neq A)} P_B V_B^A \right] |A_i\rangle = \epsilon_{A_i} |A_i\rangle, \quad (16)$$

or, since

$$U_A = \sum_{B (\neq A)} V_B^A, \quad (17)$$

$$\left[T + V_A + \sum_{B (\neq A)} (V_B^A - P_B V_B^A) \right] |A_i\rangle = \epsilon_{A_i} |A_i\rangle. \quad (18)$$

We have chosen to use still another form of the pseudo-

potential; we choose

$$D = - \sum_C P_C U_A \quad (19a)$$

$$= - \left[\sum_C P_C \right] \left[\sum_{B (\neq A)} V_B^A \right] \quad (19b)$$

$$= - \sum_{B (\neq A)} P_B V_B^A - \sum_{B (\neq A)} \sum_{C (\neq B)} P_C V_B^A, \quad (19c)$$

which yields

$$\left[T + V_A + \sum_{B (\neq A)} \left[V_B^A - \sum_C P_C V_B^A \right] \right] |A_i\rangle = \epsilon_{A_i} |A_i\rangle. \quad (20)$$

Each of Eqs. (9), (13), (18), and (20) is an exact Hartree-Fock equation in the sense that the $|A_i\rangle$ span the Hartree-Fock manifold and can be used with F to obtain Hartree-Fock results. The reason for developing alternate equations is to obtain approximate solutions; which version is best to use depends upon circumstances.

We concentrate here on the two CP versions, Eqs. (18) and (20). Our version [Eq. (20)] contains a term on the left-hand side of the equation which is absent in Anderson's CP [Eq. (18)],

$$- \sum_{B (\neq A)} \sum_{C (\neq B)} P_C V_B^A |A_i\rangle \quad (21a)$$

$$= - \sum_{B (\neq A)} \sum_{C (\neq B)} \sum_k |C_k\rangle \langle C_k | V_B^A |A_i\rangle \quad (21b)$$

$$= - \sum_{B (\neq A)} \sum_k \left[|A_k\rangle \langle A_k | V_B^A |A_i\rangle + \sum_{C (\neq A, B)} |C_k\rangle \langle C_k | V_B^A |A_i\rangle \right] \quad (21c)$$

$$\approx - \sum_{B (\neq A)} \sum_k |A_k\rangle \langle A_k | V_B^A |A_i\rangle, \quad (22)$$

neglecting the last part which contains only three-centered integrals.

Since our version subtracts more from the environment potential than Anderson's,¹ we expect the self-consistent solutions of Eq. (20) to be more localized and thus more atomiclike than the corresponding solutions of Eq. (18). This point has yet to be investigated in detail. Both versions, it should be noted again, are non-Hermitian.¹¹ The non-Hermitian aspect of CP's, which we do not treat in this paper, has been discussed by Weeks *et al.*¹²

III. APPLICATIONS

CP's have been applied to a variety of problems. For example, Bullett^{2,6} has shown how Anderson's¹ version [Eq. (18)] may be applied to an extended system (molecule or solid): If the eigenfunctions are written as linear combinations of the $|A_i\rangle$,

$$|\psi_\alpha\rangle = \sum_{A_i} |A_i\rangle C_{A_i, \alpha}, \quad (23)$$

then the eigenvalues E_α and the coefficients $C_{A_i, \alpha}$ are solutions of a secular equation

$$\sum_{A_i} (D_{B_j, A_i} - E_\alpha \delta_{AB} \delta_{ij}) C_{A_i, \alpha} = 0, \quad (24)$$

where the diagonal elements of the matrix D are given by

$$D_{A_i, A_i} = \epsilon_{A_i}, \quad (25)$$

and the off-diagonal elements are given by

$$D_{B_j, A_i} = \langle B_j | V_B^A |A_i\rangle, \quad B \neq A \quad (26)$$

$$D_{A_j, A_i} = 0, \quad j \neq i. \quad (27)$$

This is, in fact, a slight extension of the original Anderson work as applied by Bullett^{2,3} for the many-orbital-per-atom case. Anderson considered only one orbital per atom (a π -electron system such as benzene), in which case Eq. (27) would be irrelevant, and j would be equal to i in Eq. (26).

The same procedure applied to our CP [Eq. (20)] yields

$$D_{A_i, A_i} = \epsilon_{A_i} + \langle A_i | U_A |A_i\rangle = \epsilon_{A_i} + \left\langle A_i \left| \sum_{C (\neq A)} V_C^A \right| A_i \right\rangle, \quad (28)$$

$$D_{B_j, A_i} = \langle B_j | U_A |A_i\rangle = \left\langle B_j \left| \sum_{C (\neq A)} V_C^A \right| A_i \right\rangle = \langle B_j | V_B^A |A_i\rangle, \quad B \neq A \quad (29)$$

(in a two-center approximation) and

$$D_{A_j, A_i} = \langle A_j | U_A | A_i \rangle = \left\langle A_j \left| \sum_{C (\neq A)} V_C^A \right| A_i \right\rangle, \quad j \neq i. \quad (30)$$

The apparent differences in the matrix elements between our approach and Anderson's¹ would seem to belie our assertion that both approaches, along with others, are in principle exact (in the Hartree-Fock sense) and therefore equivalent. To solve this problem one must note that, in general, the orbitals $|A_i\rangle$ and the energy parameters ϵ_{A_i} will not be the same in different pseudopotential schemes. Rather, they are solutions of the particular pseudopotential

$$\langle A_j | T + V_A | A_i \rangle + \sum_{B (\neq A)} \langle A_j | (V_B^A - P_B V_B^A) | A_i \rangle = \epsilon_{A_i} \delta_{ij}. \quad (31)$$

Since $\langle A_j | T + V_A | A_i \rangle = \epsilon_{A_i}^0 \delta_{ij}$, where $\epsilon_{A_i}^0$ is the orbital energy for the free atom, this becomes

$$\epsilon_{A_i}^0 \delta_{ij} + \sum_{B (\neq A)} \left[\langle A_j | V_B^A | A_i \rangle - \sum_k \langle A_j | B_k \rangle \langle B_k | V_B^A | A_i \rangle \right] = \epsilon_{A_i} \delta_{ij}, \quad (32a)$$

or

$$\epsilon_{A_i}^0 + \sum_{B (\neq A)} \langle A_i | V_B^A | A_i \rangle - \sum_{\substack{B, k \\ (B \neq A)}} \langle A_i | B_k \rangle \langle B_k | V_B^A | A_i \rangle = \epsilon_{A_i} \quad (32b)$$

and

$$\sum_{B (\neq A)} \langle A_j | V_B^A | A_i \rangle - \sum_{\substack{B, k \\ (B \neq A)}} \langle A_j | B_k \rangle \langle B_k | V_B^A | A_i \rangle = 0, \quad j \neq i \quad (\text{Anderson}). \quad (32c)$$

The same procedure applied to Eq. (20) yields for our case

$$\epsilon_{A_i}^0 \delta_{ij} + \sum_{B (\neq A)} \left[\langle A_j | V_B^A | A_i \rangle - \sum_{C, k} \langle A_j | C_k \rangle \langle C_k | V_B^A | A_i \rangle \right] = \epsilon_{A_i} \delta_{ij} \quad (33)$$

or

$$\epsilon_{A_i}^0 \delta_{ij} + \sum_{B (\neq A)} \left[- \sum_{\substack{C, k \\ (C \neq A)}} \langle A_j | C_k \rangle \langle C_k | V_B^A | A_i \rangle \right] = \epsilon_{A_i} \delta_{ij}, \quad (34)$$

or, in the two-center approximation,

$$\epsilon_{A_i}^0 \delta_{ij} + \sum_{B (\neq A)} \left[- \sum_k \langle A_j | B_k \rangle \langle B_k | V_B^A | A_i \rangle \right] = \epsilon_{A_i} \delta_{ij} \quad (\text{ours}). \quad (35a)$$

That is,

$$\epsilon_{A_i}^0 - \sum_{\substack{B, k \\ (B \neq A)}} \langle A_i | B_k \rangle \langle B_k | V_B^A | A_i \rangle = \epsilon_{A_i} \quad (35b)$$

and

$$\sum_{\substack{B, k \\ (B \neq A)}} \langle A_j | B_k \rangle \langle B_k | V_B^A | A_i \rangle = 0, \quad j \neq i \quad (\text{ours}). \quad (35c)$$

Thus when atomic orbitals are used for the $|A_i\rangle$'s, the resulting constraints are (35b) and (35c) with our CP, while with Anderson's CP Eqs. (32b) and (32c) should be satisfied. We consider first Eqs. (35b) and (32b) for the

tial equations in which they appear.

While this point is rather obvious,⁶ these constraints on the $|A_i\rangle$ and ϵ_{A_i} are easily overlooked in semiquantitative applications of the CP method, and they can indeed have important consequences. In such applications the CP concept might be used to justify the use of atomic orbitals to evaluate the matrix elements of the overlap-free secular equation (24) without actually solving the CP equation. Then, our CP is better suited for this purpose, in that the resulting constraints are less severe, as shown in the following discussion.

If we assume the $|A_i\rangle$'s to be atomic orbitals, and multiply Eq. (18) by $\langle A_j |$, we have for the Anderson case

two different CP's. These equations define the eigenvalues of the CP equations. If we assume the atomic orbitals to be fairly localized, so that the wave-function overlaps are small, we have for Eq. (35b) $\epsilon_{A_i}^0 \approx \epsilon_{A_i}$, neglecting the second-order terms. In the case of Eq. (32b), in addition to the second-order terms, $\epsilon_{A_i}^0$ and ϵ_{A_i} differ by a crystal-field term

$$\sum_{B (\neq A)} \langle A_i | V_B^A | A_i \rangle,$$

which may not be small even when wave-function overlaps are small. Secondly, the constraint equation (35c) in our case is not, in general, as severe as Eq. (32c) is for Anderson's case, due to the presence of the off-diagonal crystal-field term

$$\sum_{B (\neq A)} \langle A_j | V_B^A | A_i \rangle, j \neq i$$

in Eq. (32c). These crystal-field terms will generally be large compared to the wave-function-overlap terms.

Thus we see that the atomic orbitals $|A_i\rangle$, as solutions, are more compatible with our CP than with Anderson's, and there will be cases for which such $|A_i\rangle$ will not be appropriate to use with the Anderson-Bullett scheme [Eqs. (25)–(27)]. In cases of high symmetry, both CP's will yield similar results with atomic $|A_i\rangle$'s. In such cases the crystal-field terms in Eq. (32c) will vanish, and Eqs. (32c) and (35c) will become identical; the crystal-field terms in Eq. (32b) for an atomically degenerate state will be independent of i , so that the eigenvalues ϵ_{A_i} in Eqs. (32b) and (35b) will differ only by a constant. Then both schemes [Eqs. (25)–(27) and Eqs. (28)–(30)] will give the same result, except for a rigid shift in the energy scale.

A simple example of a low-symmetry system in which the differences in the CP's are important is the following: We consider a hypothetical linear water molecule in which only oxygen p orbitals $|A_i\rangle$ are used. Since there are no orbitals for which $B \neq A$, the Anderson¹ constraints [Eqs. (32b) and (32c)] become

$$\epsilon_{A_i}^0 + \sum_{B (\neq A)} \langle A_i | V_B^A | A_i \rangle = \epsilon_{A_i} \quad (36a)$$

and

$$\sum_{B (\neq A)} \langle A_j | V_B^A | A_i \rangle = 0, j \neq i. \quad (36b)$$

Equation (36b) is satisfied *only* if the orbitals $|A_i\rangle$ have the proper symmetry; this will occur in the present case if one p orbital is quantized along the H–O–H direction, but for arbitrary orientation of the quantization coordinates, this equation will not be satisfied. In Eq. (36a), ϵ_{A_i} and $\epsilon_{A_i}^0$ will differ by

$$\sum_{B (\neq A)} \langle A_i | V_B^A | A_i \rangle$$

(which is never zero) and since, in general, it depends on i , this difference is more than just a rigid shift of the energy for various i 's. It *will* be a rigid shift if the H–O–H direction makes equal angles with the x and y , and z axes of the quantization coordinate, but in that case Eq. (36b) is not satisfied. Our CP applied to this example, on the other hand, gives $\epsilon_{A_i}^0 = \epsilon_{A_i}$, and the constraint equation (35c) is automatically satisfied for *arbitrary* quantization coordinate orientations.

If the water molecule is not linear, the two OH bonds at an arbitrary angle, there is not even a preferred symmetry axis of quantization for oxygen p orbitals to make the diagonal crystal-field term independent of i , or to satisfy Eq. (36b). In such a case the orbitals $|A_i\rangle$ will have to be considerably different from atomlike to satisfy the above constraints.

Thus in the semiquantitative application of the CP approach, it is essential that the constraint equations be satisfied, or that the violations are not severe. (Of course, if the CP equations *are* solved self-consistently in a detailed calculation, there will be no constraint conditions, and the different CP's will yield equivalent results.)

The above example is realistic for many solid-state applications. In most forms of SiO_2 the oxygen is twofold coordinated, and in all cases its environment is of low symmetry. Many fluorides have similar symmetries, as do many other oxides. One anticipates that in such systems the constraint equations for Anderson's pseudopotential will remain unsatisfied for any particular quantization of the anion p orbitals. For idealized β -cristobalite¹³ for example, Eq. (36b) is satisfied only if the orbitals are quantized parallel and perpendicular to the O–Si axis. With our CP, on the other hand, arbitrary quantization is allowed. In the case of rutile-type SiO_2 (stishovite), where oxygen is threefold coordinated, Eq. (36b) cannot be satisfied for oxygen p orbitals, whereas with our CP, such a severe constraint does not appear for atomic orbitals.

These points are relevant to tight-binding and Hückel-type procedures. In both cases it is convenient to take the overlap matrix as unity, and the CP theory may be used⁴ to justify this, since the overlap matrix does not appear in the CP matrix elements.

The potential pitfall is now apparent. In general, in these procedures one uses the atomic orbitals and quantizes them in the most convenient way, i.e., p_x, p_y, p_z with respect to fixed axes. *But*, if this is done, then one must include crystal-field matrix elements, as in Eqs. (28) and (30). If Anderson's CP is used,¹ in which these terms are omitted, the results may be in serious error. This has occurred in several cases, including the following.

Hückel theory, in general, does not include off-diagonal crystal-field terms

$$\sum_{B (\neq A)} \langle A_j | V_B^A | A_i \rangle, i \neq j.$$

Hückel calculations¹⁴ of the SiO_2 valence bands give results in very poor agreement with other calculations which are thought to be approximately correct.¹⁵

Tight-binding calculations of rutile-type ionic crystals have recently appeared.¹⁶ In these calculations, only anion orbitals have been included and crystal-field terms have been neglected. These calculations appear to give unreliable results as compared with both experiment and theory.¹⁷ The calculations⁷ which stimulated the present investigation involve the valence bands of rutile compounds (in particular, the stishovite form of SiO_2) where we used only anion wave functions but included crystal-field terms. These results agree well with those of Ref. 17.

IV. CONCLUSIONS

The symmetry constraints imposed on the orbitals by the CP equations must be recognized when the CP method is used to justify a tight-binding fit of valence bands in nonmetals. These constraints require, in general, that both crystal-field and overlap terms be included. When this is done, good energy bands can be obtained in low-symmetry systems even though only anion orbitals are used.

V. ACKNOWLEDGMENT

This research was supported by the U.S. Office of Naval Research, Contract No. N00014-76-C-0125.

- ¹P. W. Anderson, Phys. Rev. Lett. 21, 13 (1968); Phys. Rev. 181, 25 (1969).
- ²D. W. Bullett, J. Phys. C 8, 2695 (1975); 8, 2707 (1975).
- ³D. W. Bullett, Solid State Commun. 17, 965 (1975).
- ⁴W. Lambrecht, B. Djafari-Rouhani, M. Lannoo, and J. Vennik, J. Phys. C 13, 2503 (1980).
- ⁵E. Doni, R. Girlanda, and L. Resca, J. Phys. Soc. Jpn. 49, Suppl. A, 89 (1980).
- ⁶D. W. Bullett, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1980), Vol. 35, p. 129; V. Heine, in *Solid State Physics*, Vol. 35, p. 1.
- ⁷J. K. Rudra and W. Beall Fowler, following paper, Phys. Rev. B 28, 1061 (1983).
- ⁸W. H. Adams, J. Chem. Phys. 34, 89 (1961); 37, 2009 (1962).
- ⁹T. L. Gilbert, in *Molecular Orbitals in Chemistry, Physics, and Biology*, edited by P. O. Löwdin and B. Pullman (Academic, New York, 1964), p. 405.
- ¹⁰A. B. Kunz, J. Phys. B 6, L47 (1973).
- ¹¹Kunz has recently carried out band-structure calculations using local orbitals derived from a modification of the approach used in Ref. 10. His D [Eq. (1)] is similar to ours, except that his is Hermitian [A. B. Kunz, Phys. Rev. B 26, 2056 (1982)].
- ¹²J. D. Weeks, P. W. Anderson, and A. G. H. Davidson, J. Chem. Phys. 58, 1388 (1973).
- ¹³P. M. Schneider and W. Beall Fowler, Phys. Rev. Lett. 86, 425 (1976); Phys. Rev. B 18, 7122 (1978).
- ¹⁴A. Breeze and P. G. Perkins, J. Chem. Soc. Faraday Trans. II 69, 1237 (1973).
- ¹⁵Early mixed-basis and tight-binding (Ref. 13) and bond-orbital [S. T. Pantelides, Phys. Lett. A54, 401 (1975); S. T. Pantelides and W. A. Harrison, Phys. Rev. B 13, 2667 (1976).] calculations on β -cristobalite gave similar results whose validity is discussed in these references. Similar agreement exists among calculations on quartz: J. R. Chelikowsky and M. Schluter, Phys. Rev. B 15, 4020 (1977); E. Calabrese and W. Beall Fowler, *ibid.* 18, 2888 (1978).
- ¹⁶V. K. Bashenov, I. Baumann, and V. V. Timofeenko, Phys. Status Solidi B 81, K55 (1977); V. K. Bazhenov, Yu. K. Markolenko, and V. V. Timofeenko, Ukr. Fiz. Zh. (Russ. Ed.) 26, 677 (1981); C. Jouanin, J. P. Albert, and C. Gout, J. Phys. (Paris) 37, 595 (1976).
- ¹⁷J. Robertson [J. Phys. C 12, 4767 (1979)] has obtained valence bands for a number of rutile compounds using a tight-binding approach which includes both cation and anion orbitals, quantized with respect to the bonding unit. He presents convincing arguments in favor of his results, which differ considerably from those of Ref. 16. His assertion that earlier tight-binding results are unreliable because of the absence of cation orbitals is not correct, however; rather, as the present paper emphasizes, the error is in the absence of crystal-field terms.