Significance of nonorthogonality in tight-binding models

B. A. McKinnon and T. C. Choy
Department of Physics, Monash University, Clayton 3168, Australia
(Received 6 February 1995; revised manuscript received 24 July 1995)

The tight-binding Hamiltonian (TBH) presumes a basis of orthogonal Wannier states assumed to be localized about the lattice sites. We demonstrate that when beginning with a nonorthogonal or overlapping basis, orthogonalization introduces terms equivalent to beyond nearest-neighbor hopping terms in the TBH. We show, using the example of graphite, that the consideration of the overlap between nearest-neighbor orbitals leads to an improved and simplified parametrization of the TBH band structure. In addition, localized overlap is shown to give a simple picture of the orthogonality catastrophe.

I. INTRODUCTION

The tight-binding Hamiltonian (TBH) has the following well known form:

$$\mathbf{H} = \sum_i \epsilon_i c_i^\dagger c_i + \sum_{i,j} V_{ij} c_i^\dagger c_j + \text{H.c.}, \tag{1}$$

where the creation and destruction operators act upon a basis of orthogonal states, Wannier states, assumed to be localized about the lattice sites. ϵ_i represents the energy of an electron at site i, while V_{ij} represents the "hopping" of an electron from site i to j. Over the last three decades, this Hamiltonian has been used successfully to describe the electronic properties of systems containing localized electrons. In such systems, the electron hopping term V_{ij} is limited to hopping between near neighbors. The parameters ϵ_i and V_{ij} are optimized to fit either experimental results or more sophisticated band structure calculations.

The Wannier states form an orthogonal set and constitute the inverse Fourier transform of the crystal Bloch states. In practice, however, the Wannier states are often approximated by atomiclike Slater orbitals[1] centered about each lattice site. Obviously the overlap between such orbitals located on nearest-neighbor sites is nonzero, so that although the Slater-like orbitals form a natural basis set, they are not orthogonal. Yet the powerful formalism of second quantization requires an orthogonal basis for the proper definition of fermion creation and destruction operators and their associated Fock space.

Recently, there have been developments in the use of localized Hamiltonians with nonorthogonal bases to describe complex systems, such as defects and surfaces. Vergés and Yndurian[2] found that using a nonorthogonal Slater-like basis set, they were able to obtain a better band structure for silicon than using the conventional TBH for the same number of parameters. Menon and Subbaswamy[3] have gone on to develop a transferable parameter set for silicon as was done by Harrison[4] for TB parameters. The importance of nearest-neighbor nonorthogonality in TB calculations of electron-phonon interactions has been recognized since 1979 when a nonorthogonal basis set was introduced by Varma *et al.*[5] to calculate the electron-phonon interactions in transition metals within a TB framework. In this paper, we elucidate the connection between systems as described by an orthogonal and a nonorthogonal basis set.

In 1950, Löwdin[6] developed a transformation to obtain a Hamiltonian $\mathbf{H}$ with an orthogonal basis from a Hamiltonian $\mathcal{H}$ with a nonorthogonal basis:

$$\mathbf{H} = (\mathbf{1} + \mathbf{S})^{-1/2} \mathcal{H} (\mathbf{1} + \mathbf{S})^{-1/2}, \tag{2}$$

where

$$S_{\alpha\beta} = \int \phi_\alpha(r) \phi_\beta(r) d\tau - \delta_{\alpha,\beta} \tag{3}$$

and $\phi_\alpha(r)$ are atomiclike orbitals located at each lattice site.

In this paper, we compare the transformed Hamiltonian with the TBH. This procedure was recently carried out for a one-dimensional (1D) chain by Mirabella *et al.*,[7,8] where the Löwdin transformation was used implicitly to examine the effect of nearest-neighbor nonorthogonality. They found that the *first order* effect of the overlap between atomiclike wave functions on nearest-neighbor sites induces terms in the band structure equivalent to parametrization of the TB Hamiltonian up to *second-nearest-neighbor* hopping terms.

Using Löwdin's transformation, the results of Ref. 7 can be generalized to 2D and 3D lattices. We have examined the effect of a nonorthogonal basis set, with overlapping nearest neighbors only, upon single-particle electronic properties of square, cubic, triangular and hexagonal monatomic lattices. To obtain the same band structure starting from a TBH requires the inclusion of third-nearest-neighbor hopping terms, i.e., four parameters. Given our understanding of crystal structure in terms of bonding between overlapping orbitals, our findings indicate that a more appropriate starting point for TB calculations is a nonorthogonal localized basis set. As an example, we compare the popular Slonczewski and

Weiss and McClure[9,10] (SWMc) parametrization of the TBH for Bernal graphite with that obtained beginning from a nonorthogonal basis, and show that the use of a nonorthogonal basis leads to an improved physical appreciation of the contributions to the electronic properties of Bernal graphite.

Overlap effects also play a vital role in determining the many-body wave function. We consider a locally perturbed system, such as one in which a core electron has been excited out of the crystal, as in the x-ray edge problem. The overlap between the perturbed and unperturbed valence orbitals is shown to lead to a particularly simple demonstration of Anderson's[11] orthogonality catastrophe. In an accompanying paper, paper II, we demonstrate the effect of the incorporation of nonorthogonality into a many-body interacting system, the Hubbard Hamiltonian.

II. THE LÖWDIN TRANSFORMATION

The eigenfunctions of a system can be written as a linear combination of basis functions. The coefficients of such a linear combination of nonorthogonal basis function must satisfy the eigenvalue equation and completeness condition:

$$\mathcal{H}\mathbf{x} = (\mathbf{1}+\mathbf{S})\mathbf{x}\mathbf{E}, \tag{4}$$

$$\mathbf{1} = \mathbf{x}^{\dagger}(\mathbf{1}+\mathbf{S})\mathbf{x}, \tag{5}$$

where

$$\mathcal{H}_{\alpha\beta} = \langle x_\alpha|\mathcal{H}|x_\beta\rangle, \tag{6}$$

$$\mathbf{S}_{\alpha\beta} = \langle x_\alpha|x_\beta\rangle - \delta_{\alpha,\beta}, \tag{7}$$

and $|x_\alpha\rangle$, $|x_\beta\rangle$ are elements of the basis, $\{\mathbf{x}_i\}$, of atomiclike orbitals centered about the lattice sites α and β, and $\mathcal{H}$ is the crystal Hamiltonian.

Making the substitution, $\mathbf{x} = (\mathbf{1}+\mathbf{S})^{-\frac{1}{2}}\mathbf{c}$, the above equations reduce to the familiar

$$\mathbf{H}\mathbf{c} = \mathbf{E}\mathbf{c}, \tag{8}$$

$$\mathbf{1} = \mathbf{c}^{\dagger}\mathbf{c}, \tag{9}$$

where

$$\mathbf{H} = (\mathbf{1}+\mathbf{S})^{-1/2}\mathcal{H}(\mathbf{1}+\mathbf{S})^{-1/2}. \tag{10}$$

In writing down the system's Hamiltonian, we assume that only elements connecting the nearest neighbors are significant. The corresponding Hamiltonian is

$$\mathcal{H} = \sum_{\alpha} h_0 x_\alpha^{\dagger}x_\alpha + \sum_{\alpha\beta} h_1 x_\alpha^{\dagger}x_\beta + \text{H.c.} \tag{11}$$

where h_0 and h_1 are constants with $h_1 \neq 0$ for α,β nearest neighbors only.

The Fourier transform of the above Hamiltonian is

$$\mathcal{H} = \sum_{k}[h_0 + h_1\mu(\vec{k})]x_k^{\dagger}x_k, \tag{12}$$

where

$$\mu(\vec{k}) = \sum_{\alpha\neq\beta} e^{i\vec{k}\cdot(\vec{R}_\alpha - \vec{R}_\beta)}, \tag{13}$$

i.e., $\mu(\vec{k})$ is a function of the wave vector dependent upon the geometry of the lattice. Including nearest-neighbor overlap, S, only, the Fourier transform of the overlap matrix is

$$\mathbf{1}+\mathbf{S} = \sum_{k}[1 + S\mu(\vec{k})]x_k^{\dagger}x_k. \tag{14}$$

In single band lattices such as the 1D, square, cubic, and triangular lattices, $\mathcal{H}$ and $(\mathbf{1}+\mathbf{S})$ commute, so that

$$\mathbf{H} = \mathcal{H}(\mathbf{1}+\mathbf{S})^{-1}. \tag{15}$$

Thus, the eigenvalues of the orthogonalized Hamiltonian are given by

$$E(\vec{k}) = \frac{h_0 + h_1\mu(\vec{k})}{1 + S\mu(\vec{k})}, \tag{16}$$

which can be expressed as a series expansion in $S\mu(\vec{k})$:

$$E(\vec{k}) = h_0 + (h_1 - h_0S)\mu(\vec{k}) - S(h_1 - h_0S)\mu(\vec{k})^2 + \cdots. \tag{17}$$

One could obtain Eq. (16) by finding the eigenvalues of Eq. (4) directly. Such a technique is employed in the extended Hückel method.[12] In this paper, however, we establish the link between nonorthogonal basis calculations and TB calculations with their accompanying basis of Wannier states. This is carried out in the following section, where the elements of the series expansion in Eq. (17) are equated with their counterparts in the TB energy dispersion relations.

III. THE EFFECT OF OVERLAP ON BAND STRUCTURE

In Eq. (17), the energy dispersion is expressed as a series expansion in $S\mu(\vec{k})$. Retaining only terms up to first order in the overlap between nearest neighbors, S, the series becomes

$$E(\vec{k}) = h_0 + (h_1 - h_0S)\mu(\vec{k}) - Sh_1|\mu(\vec{k})|^2. \tag{18}$$

In Eq. (13) it can be seen that the argument of $\mu(\vec{k})$, $\vec{k}\cdot(\vec{R}_i - \vec{R}_j)$, while a continuous function of the wave vector also depends upon the displacement of nearest-neighbor lattice sites. $|\mu(\vec{k})|^2$ corresponds to two consecutive lattice displacements. In a 1D system, such as the linear chain, the $|\mu(\vec{k})|^2$ factor introduces terms into the energy dispersion corresponding to second-nearest-neighbor hopping in a TB scheme, while in 2D systems $|\mu(\vec{k})|^2$ introduces not only second-nearest-neighbor, but also third-nearest-neighbor terms. This is the crux of the relationship between nonorthogonal and orthogonal

tight-binding models. In Eq. (16), the nonorthogonal energy dispersion relation is expressed as an infinite series expansion in the nearest-neighbor overlap. Increasing powers of overlap correspond to ever larger lattice displacement factors. The series expansion is equivalent to a parametrized TBH, with hopping terms of decreasing magnitude between sites further and further apart. In the usual form of the TBH, the series is terminated after the first few terms, an approximation to the nonorthogonal system that is increasingly accurate, the more localized the electrons.

This is readily illustrated by the example of the infinite one-dimensional chain, for which with interatomic spacing a, $\mu(\vec{k}) = 2\cos ka$. If the only terms in the Hamiltonian are the overlap between nearest-neighbor atomiclike orbitals S and the elements of the nonorthogonal Hamiltonian connecting the site to itself h_0 and its nearest neighbor h_1, then the energy dispersion to the first order in S is

$$E(\vec{k}) = (h_0 - 2Sh_1) + 2(-Sh_0 + h_1) \times \cos ka - 2Sh_1 \cos 2ka\,. \tag{19}$$

This corresponds to a parametrized TB 1D chain energy dispersion, with hopping between *next-nearest* neighbors, i.e.,

$$E(\vec{k}) = \alpha + 2\beta\cos ka + 2\gamma\cos 2ka. \tag{20}$$

The effect of nonorthogonality is clearly greatest where $\mu(\vec{k})$ is a maximum, namely, at the band edges. In Fig. 1(a) it can be seen that, as was noted by Ref. 7, the nearest-neighbor orthogonality makes the band structure asymmetric, corresponding to the effective mass of the electrons at the top of the band being smaller than that of the electrons at the bottom of the band. When the band is half filled, $\mu(\vec{k}) = 0$, so the neglect of overlap effects is valid. Figure 1(a) shows that for a small overlap, the series expansion converges rapidly to the all orders result.

Figure 1(b) illustrates the effect of nearest-neighbor nonorthogonality in the square lattice. Here, the maximum value of $\mu(\vec{k})$ is twice that of the linear chain, so that the effect of nonorthogonality for the same nearest-neighbor overlap is greater. The other important feature is that the simulation of the effect of nearest-neighbor nonorthogonality to first order introduces TB hopping terms between up to the third nearest neighbors. The correspondence between the TBH parameters and the coefficients for the energy dispersion in the triangular, square, and cubic lattices with nonorthogonal bases is contained in Appendix A.

The hexagonal lattice is more complex, being composed of two interpentrating triangular sublattices. It is convenient to define two reciprocal space states, corresponding to the Fourier sum over each sublattice. In this representation, the Hamiltonian and overlap matrices are, respectively,

$$\mathcal{H} = \begin{bmatrix} h_0 & h_1\mu(\vec{k}) \\ h_1\mu^*(\vec{k}) & h_0 \end{bmatrix}, \tag{21}$$

$$\mathbf{1} + \mathcal{S} = \begin{bmatrix} 1 & S\mu(\vec{k}) \\ S\mu^*(\vec{k}) & 1 \end{bmatrix}, \tag{22}$$

where $\mu(\vec{k}) = e^{i(k_x a + k_y b)} + e^{i(-k_x a + k_y b)} + e^{-i2k_y b}$ [see Fig. 2(a)]. The Löwdin transformation can be expressed as a series expansion in $\mathcal{S}$:

$$\mathbf{H} = (\mathbf{1} + \mathcal{S})^{-\frac{1}{2}}\mathcal{H}(\mathbf{1} + \mathcal{S})^{-\frac{1}{2}} \tag{23}$$

$$= \mathcal{H} - \frac{1}{2}(\mathcal{S}\mathcal{H} + \mathcal{H}\mathcal{S}) + \frac{1}{4}\mathcal{S}\mathcal{H}\mathcal{S} + \frac{3}{8}(\mathcal{S}\mathcal{S}\mathcal{H} + \mathcal{H}\mathcal{S}\mathcal{S}) + \cdots. \tag{24}$$

Keeping terms to the first order in S, the orthogonalized

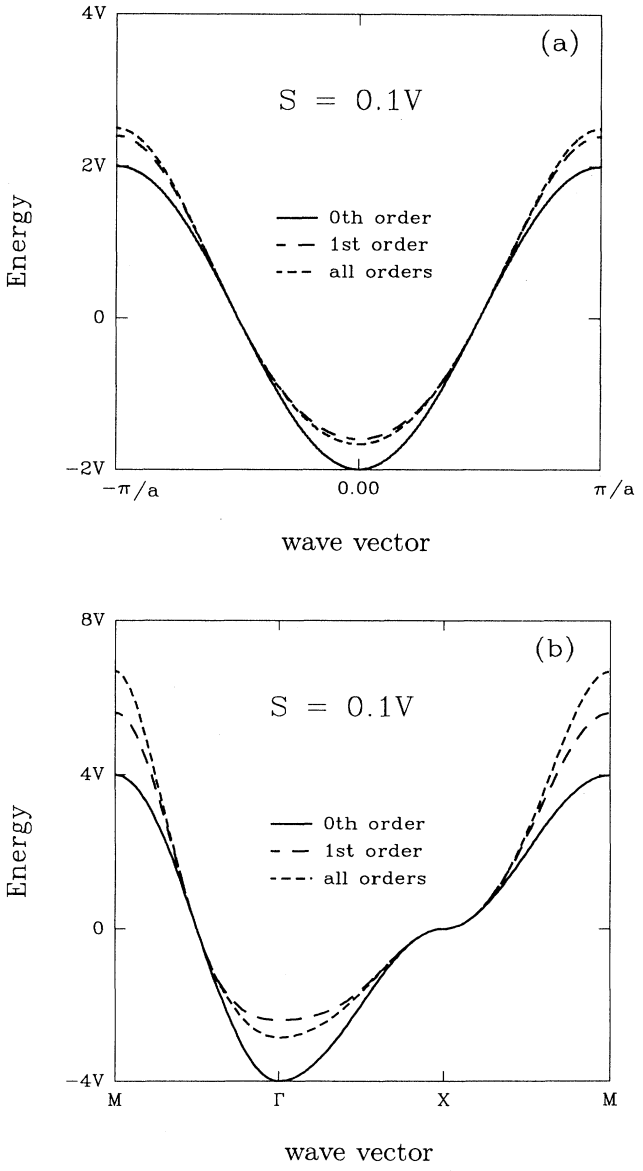

FIG. 1. Band structure to the 0th, 1st, and all orders in overlap for a (a) 1D chain and (b) square lattice.

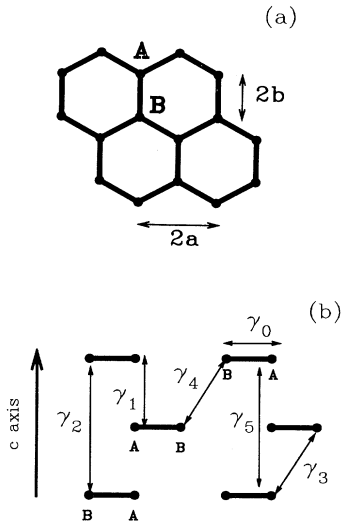

FIG. 2. (a) The hexagonal lattice, (b) SWMc parameters for Bernal graphite.

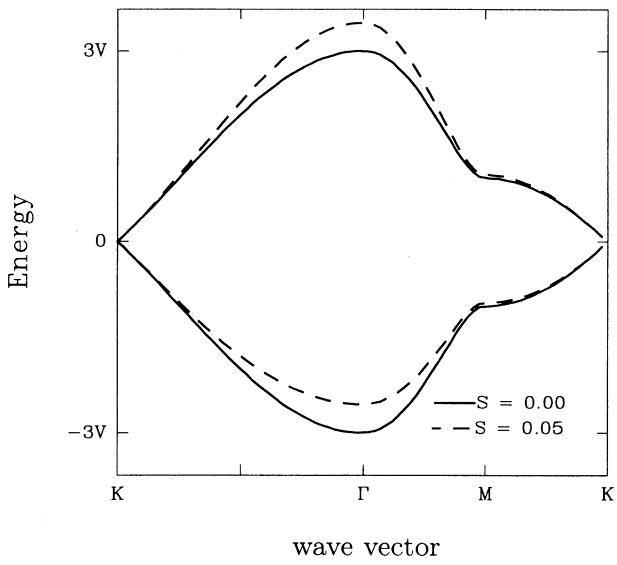

FIG. 3. Band structure for hexagonal lattice: $S = 0, S = 0.05$.

Hamiltonian $\mathbf{H}$ is

$$\mathbf{H} = \begin{bmatrix} h_0 - Sh_1|\mu(\vec{k})|^2 & (h_1 - Sh_0)\mu(\vec{k}) \\ (h_1 - Sh_0)\mu^*(\vec{k}) & h_0 - Sh_1|\mu(\vec{k})|^2 \end{bmatrix} . \quad (25)$$

The TBH for the hexagonal lattice, including nearest- and next-nearest-neighbor hopping terms V_0 and V_1, is

$$\mathbf{H} = \begin{bmatrix} \epsilon + V_1[|\mu(\vec{k})|^2 - 3] & V_0\mu(\vec{k}) \\ V_0\mu^*(\vec{k}) & \epsilon + V_1[|\mu(\vec{k})|^2 - 3] \end{bmatrix} . \quad (26)$$

A comparison of Eqs. (25) and (26) shows that the inclusion of next-nearest neighbor hopping terms in the TBH model of the hexagonal lattice corresponds to modeling nearest-neighbor overlap to the first order. The structure of the hexagonal lattice means that upon displacement by two lattice sites, one arrives at the next-nearest neighbor of the initial lattice site. The energy dispersion to first order in nearest-neighbor nonorthogonality is

$$E(\vec{k}) = h_0 - h_1 S|\mu(\vec{k})|^2 \pm (h_1 - Sh_0)|\mu(\vec{k})| , \quad (27)$$

from which it can be seen that the band minimum and maximum are very sensitive to the overlap. This is illustrated in Fig. 3, with $h_0 = 0$ and $h_1 = V_0$.

A. Bernal graphite

In Bernal graphite, the layers are stacked alternately $ABAB$. The effect is that the "B" sites are seated in the hollows of the hexagons of the adjacent layers, unlike the "A" sites, which bond weakly with the "A" sites in the neighboring planes above and below. The bands about the Fermi level in graphite arise from $2p_z$ orbitals participating in π bonding within the layer plane and weak bonding between adjacent layers. The overlap with the intraplane hybridized sp_2 bonded bands is small.[13] We now describe the electronic structure of Bernal graphite beginning from a nonorthogonal basis where the only significant overlap is between in plane and interplane nearest neighbors, s_0 and s_1. The only nonzero elements of the Hamiltonian are taken to be h_0, the matrix element between nearest neighbors in the layer plane, and h_1, the matrix element between the nearest neighbor on adjacent planes. The respective Hamiltonian and overlap matrices are[14]

$$\mathcal{H} = \begin{bmatrix} 0 & h_0\mu(\vec{k}) & 2h_1\nu & 0 \\ h_0\mu^*(\vec{k}) & 0 & 0 & 0 \\ 2h_1\nu & 0 & 0 & h_0\mu^*(\vec{k}) \\ 0 & 0 & h_0\mu(\vec{k}) & 0 \end{bmatrix} , \quad (28)$$

$$\mathbf{1} + \mathcal{S} = \begin{bmatrix} 1 & s_0\mu(\vec{k}) & 2s_1\nu & 0 \\ s_0\mu^*(\vec{k}) & 1 & 0 & 0 \\ 2s_1\nu & 0 & 1 & s_0\mu^*(\vec{k}) \\ 0 & 0 & s_0\mu(\vec{k}) & 1 \end{bmatrix} , \quad (29)$$

where $\nu = \cos k_z c$, and s_0 and s_1 are the overlap integrals between in plane and interplane nearest neighbors, respectively. The following orthogonalized Hamiltonian is obtained upon performing a Löwdin transformation and retaining only terms to first order in the overlaps s_0 and s_1:

$$\mathbf{H} = \begin{bmatrix} -(s_0h_0|\mu|^2 + 4s_1h_1\nu^2) & h_0\mu & 2h_1\nu & -(s_1h_0 + s_0h_1)\nu\mu^* \\ h_0\mu^* & -s_0h_0|\mu|^2 & -(s_1h_0 + s_0h_1)\nu\mu^* & 0 \\ 2h_1\nu & -(s_1h_0 + s_0h_1)\nu\mu & -(s_0h_{\alpha\beta}|\mu|^2 + 4s_1h_1\nu^2) & h_0\mu^* \\ -(s_1h_0 + s_0h_1)\nu\mu & 0 & h_0\mu & -s_0h_0|\mu|^2 \end{bmatrix} . \quad (30)$$

Table I contains values of the band structure in graphite obtained from angle-resolved photoemission spectroscopy experiments. We have fitted our nonorthogonality parameters h_0, h_1, s_0, and s_1 to a conduction band Γ point of 10.0 eV and valence band Γ points of -7.0 and -8.5 eV, obtaining $h_0 = -3.0$, $h_1 = -0.37$, $s_0 = 0.044$, and $s_1 = -0.047$.

The SWMc TB parametrization of Bernal graphite consists of seven parameters, six of which are illustrated in Fig. 2(b). The seventh parameter γ_6 represents the small difference between the on site energy of "A" and "B" sites. McClure[10] showed that of the seven parameters, the four most significant are γ_0, γ_1, γ_2, and γ_3. The hopping term between "B" sites on every second layer, γ_2, lifts the degeneracy of the highest valence band and lowest conduction band along the K-H line of the Brillouin zone, creating the Fermi surface with its characteristic hole and electron pockets. In contrast to γ_2, the hopping term between "A" sites on every second layer, γ_5, has little effect on the band structure. γ_3 introduces asymmetry into the band structure and causes a trigonal warping of the Fermi surface; however, the effect of γ_4 is small. The difference in on site energies between the "A" and "B" sites, γ_6, shifts the bands slightly with respect to one another.

TABLE I. Experimental energy values at M, K, and Γ (in eV).

		Law *et al.*[a]	Marchand *et al.*[b]	Takahashi[c] *et al.*
Γ	conduction			10
	valence	-6.5, -8.0	-7.6 ±0.1, -9.0±0.1	-8,-10
K	conduction			
	valence	0, -0.5	-0.35±0.1, -1.05±0.1	-1
M	conduction			2
	valence	-2.0, -3.0	-2.4±0.1, -3.0±0.1	-2.7

[a]Reference 21.
[b]Reference 22.
[c]Reference 23.

In a reciprocal space representation with four basis vectors corresponding to the respective Fourier sums over the "A" and "B" sites of the two layers in the unit cell, the SWMc TBH becomes

$$\mathbf{H} = \begin{bmatrix} 2\gamma_5(2\nu^2-1)+\gamma_6 & \gamma_0\mu & 2\gamma_1\nu & 2\gamma_3\nu\mu^* \\ \gamma_0\mu^* & 2\gamma_2(2\nu^2-1) & 2\gamma_3\nu\mu^* & 2\gamma_4\nu\mu \\ 2\gamma_1\nu & 2\gamma_3\nu\mu & 2\gamma_5(2\nu^2-1)+\gamma_6 & \gamma_0\mu^* \\ 2\gamma_3\nu\mu & 2\gamma_4\nu\mu^* & \gamma_0\mu & 2\gamma_2(2\nu^2-1) \end{bmatrix}. \tag{31}$$

Comparison of the TBH with that obtained beginning from a nonorthogonal basis [Eq. (30)] shows that $\gamma_0 \leftrightarrow h_0$, $\gamma_1 \leftrightarrow h_1$, $\gamma_3 \leftrightarrow -(s_1h_0+s_0h_1)$, and $\gamma_5 \leftrightarrow -s_1h_1$. The next-nearest-neighbor hopping terms γ_3 and γ_5 can be seen to model nearest-neighbor overlap in the nonorthogonal system. However, the orthogonalized Hamiltonian has no equivalent of γ_2 or γ_4. Yet γ_2 determines the existence of the Fermi surface in Bernal graphite. Therefore, the Hamiltonian matrix element between "B" sites on the next-nearest layers, h_2, also needs to be included in the nonorthogonal Hamiltonian to obtain the experimentally observed Fermi surface. Density functional calculations[16] indicate that in Bernal graphite there is delocalization of some of the π bond charge density from the "A" sites to the neighboring in plane "B" sites and interlayer regions. Such a redistribution of charge density leads to an increase in the significance of the matrix element between "B" sites on every second layer.

From the correspondence between the nonorthogonal and TB parameters obtained above, we find that our nonorthogonal parameters correspond to a tight-binding set of $\gamma_0 = -3.0$, $\gamma_1 = -0.37$, $\gamma_3 = -0.13$, and $\gamma_5 = -0.02$. Table II contains SWMc parameters determined from more sophisticated theoretical band structure calculations and experimental measurements, with which our values of γ_0, γ_1, and γ_5 are consistent. Inclusion of the second-nearest-neighbor terms in the nonorthogonal Hamiltonian would introduce terms corresponding to γ_2 and γ_4 as well as adding to γ_3, improving the agreement with the experimental parameters. Experimentally, γ_3 is found to be much larger than γ_4, which is consistent with our finding that the primary role of γ_3 is to account for the nonorthogonality of the nearest neighbors. The relative magnitudes of γ_3 and γ_4 indicate that contributions from the next-nearest-neighbor terms in the Hamiltonian are much smaller than those from the nearest neighbors.

In Fig. 4, we compare band structures obtained using the two theoretical SWMc parameter sets from Table II and our orthogonalized band structure. Figure 5 shows a more sophisticated double-ζ basis set LCAO (Ref. 17) calculation. The band structure obtained from the orthogonalized Hamiltonian is in better general agreement

TABLE II. SWMc parameters, theoretical and experimental in eV (from Ref. 15).

	Theoretical values		Experimental
γ_0	-2.598 ± 0.015[a]	-2.9[b]	-3.16 ± 0.05[c]
γ_1	-0.364 ± 0.020	-0.27	0.39 ± 0.01[d]
γ_2	0.014 ± 0.008	0.22	0.020 ± 0.002[e]
γ_3	-0.319 ± 0.020	-0.14	-0.315 ± 0.015[f]
γ_4	-0.177 ± 0.025	-0.10	-0.044 ± 0.024[g]
γ_5	-0.036 ± 0.013	-0.0063	-0.038 ± 0.005[d]
γ_6	0.026 ± 0.010	-0.0079	0.008 ± 0.002[c]
ϵ_F	-0.013 ± 0.010	-0.027	-0.024 ± 0.002[d]

[a]Reference 15.
[b]Reference 24.
[c]Reference 25.
[d]Reference 26.
[e]Reference 27.
[f]Reference 28.
[g]Reference 29.

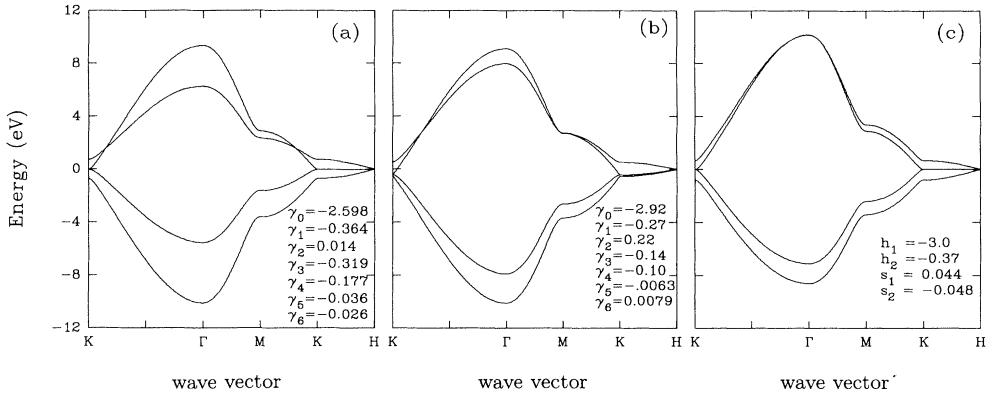

FIG. 4. Parametrized band structures for Bernal graphite (a) SWMc parameters from Ref. 15, (b) SWMc parameters from Ref. 24, (c) nonorthogonal parameters.

with that of Ref. 17 than the other two parameter sets, chiefly because of the inclusion of the nearest-neighbor in plane overlap.

It is interesting to compare the values of the overlap integrals s_0 and s_1 with those obtained by approximating the $2p_z$ orbital by a single-ζ Slater[1] orbital. ζ is a length scale factor accounting for the screening of the nucleus by core electrons. $s_0 = 0.044$ corresponds to $\zeta \approx 2.6$, while $s_1 = -0.05$ corresponds to $\zeta \approx 1.6$, which is the free carbon atom value.[18] The inconsistency of the two values reflects the anisotropy of the system.[19] In the layer plane, the wave function is considerably more localized than in the free atom, whereas from an interlayer perspective, the wave function resembles that of the free atom. This picture agrees with the finding by Ref. 16 of significant interlayer charge density, indicating that parametrization of the overlap can lead to an appreciation of the effect of the crystal field upon the charge distribution.

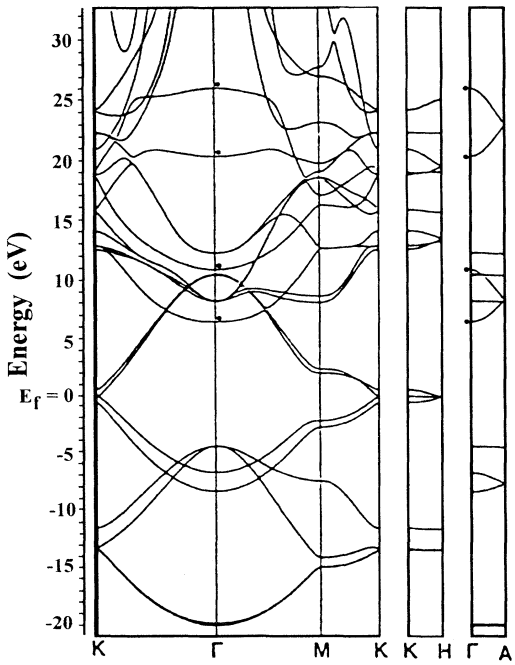

FIG. 5. Double-ζ basis set LCAO band structure calculation for Bernal graphite (from Ref. 17).

While the SWMc parameterization has four principal parameters, γ_0, γ_1, γ_2, and γ_3, to model the Fermi surface properties of Bernal graphite, beginning from a nonorthogonal basis requires five parameters: h_0, h_1, h_2, as well as the overlap parameters s_0, s_1. Although in comparison with the minimal TB model the number of parameters is not reduced, the overall agreement of the band structure with experiment and more sophisticated calculations is improved, principally through the incorporation of the in plane nearest-neighbor overlap, a feature which is neglected in the TB model and which would require an additional parameter.

This example illustrates two advantages of beginning with an overlapping basis and transforming to the more easily handled orthogonal basis. First, it enables interpretation of the need for beyond nearest-neighbor terms when fitting the parametrized band structure to more sophisticated calculations. Second, the overlap itself can be treated as a parameter, giving information about the crystal field effects within the structure. Note that inclusion of nonorthogonality effects in TB density of states calculations[20] would also improve the DOS, while not introducing extra complexity to the calculation.

IV. OVERLAP AND THE ORTHOGONALITY CATASTROPHE

The orthogonality catastrophe[11] is a many-body effect that arises when a system experiences a local change in potential. Although the single-particle states of the perturbed system are not orthogonal to the original single-particle states, the respective N-particle ground states are orthogonal in the limit that the number of particles, N, becomes infinite. Such orthogonality signals the failure of perturbation theory to describe the transition between states and is thus "catastrophic."

Here, we give a simple demonstration of how the orthogonality catastrophe arises from the overlap between the initial and final localized single-particle states. Let the initial conduction band system be described in terms of Wannier states in a TB picture:

$$\mathbf{H}_0 = -t\sum_{i,j} c_i^\dagger c_j + \text{H.c.}, \tag{32}$$

where the sum is over nearest neighbors, j, of i. The system is altered in some manner, e.g., through emission of a core state electron, causing a change in the wave function at site μ, such that

$$\mathbf{H} = -t\sum_{i,j} c_i^\dagger c_j + t\sum_i (c_\mu^\dagger - d_\mu^\dagger)c_i + \text{H.c.}, \tag{33}$$

where $d_\mu^\dagger$ is the operator creating the new wave function localized about μ. In the case of emission of an electron from a core state, the difference between the old and new wave functions of the valence electrons is small, so that $c_\mu^\dagger - d_\mu^\dagger \approx (1-S)c_\mu^\dagger = \lambda c_\mu^\dagger$, S being the overlap between the new and old conduction electron wave functions. The Fourier transform of Eq. (33) gives

$$\mathbf{H} = \frac{t}{N}\sum_k \epsilon(\vec{k})c_k^\dagger c_k + \frac{t\lambda}{\sqrt{N}}\sum_k \epsilon(\vec{k})c_\mu^\dagger c_k + \text{H.c.}, \tag{34}$$

which we canonically transform into a diagonal representation by introducing the new operator:

$$\zeta_k^\dagger = uc_k^\dagger + vc_\mu^\dagger, \tag{35}$$

where $u = \frac{1}{\sqrt{2}}(1 + \frac{1}{\sqrt{1+4\lambda^2}})^{\frac{1}{2}}$, $v = \frac{1}{\sqrt{2}}(1 - \frac{1}{\sqrt{1+4\lambda^2}})^{\frac{1}{2}}$. Clearly, as λ approaches zero, u and v become unity and zero, respectively, i.e., the original system is restored.

The new ground state is

$$|\Psi\rangle = \prod_k^{k_F} \zeta_k^\dagger |0\rangle. \tag{36}$$

The overlap between $|\Psi\rangle$ and the ground state of the original Hamiltonian

$$|\Phi\rangle = \prod_k^{k_F} c_k^\dagger |0\rangle \tag{37}$$

is

$$\mathcal{S} = \langle\Phi|\Psi\rangle \tag{38}$$

$$= \langle 0|c_{k_F}\ldots c_{k_1}\zeta_{k_1'}^\dagger \ldots \zeta_{k_F'}^\dagger|0\rangle, \tag{39}$$

which through the use of the commutation relations $\{c_{k'}, c_k^\dagger\} = \delta_{k',k}$ and $\{c_{k'}, c_\mu^\dagger\} = e^{ik'\mu}N^{-\frac{1}{2}}$ becomes

$$\mathcal{S} = \prod_k^{k_F}(u + ve^{i\vec{k}\mu}N^{-\frac{1}{2}}). \tag{40}$$

Hence

$$\ln\mathcal{S} = \sum_k \ln(u + ve^{i\vec{k}\mu}N^{-\frac{1}{2}}) \tag{41}$$

$$= \sum_k \left\{\ln u + \ln\left(1 + \frac{v}{u}e^{ik\mu}N^{-\frac{1}{2}}\right)\right\} \tag{42}$$

$$\approx \sum_k \left\{\ln u + \frac{v}{u}e^{ik\mu}N^{-\frac{1}{2}}\right\} \tag{43}$$

$$= N\ln u + \frac{v}{u}\sqrt{N}\delta_{\mu,0}, \tag{44}$$

so that

$$\mathcal{S} = \mathcal{A}e^{N\ln u}. \tag{45}$$

If $\lambda = 0$, i.e., original and new ground states are the same, then $v = 0$, $u = 1$, and $\mathcal{S} = 1$. However, for any finite value of λ, $u < 1$, hence in the limit of an infinite number of particles, $\mathcal{S} = 0$. This is the orthogonality catastrophe. The effect of a localized impurity is to create a ground state orthogonal to that of the system with no impurity, so that there can be no adiabatic transition from one to the other, resulting in a breakdown of any perturbative description of the transition.

V. CONCLUSION

In conclusion, we have found that nonorthogonality effects cast light on the physics of band structure calculations. The use of the Löwdin transformation to obtain the Hamiltonian for an orthogonal system from that of a nonorthogonal system is shown to elucidate the origin of beyond nearest-neighbor terms in the TBH and to lead to an improved parametrization of the band structure of Bernal graphite. In addition, we have shown that localized nonorthogonality between states of the system before and after a transition, leads to a nonperturbative ground state for a many-body fermionic system, demonstrating in a simple way the orthogonality catastrophe originally found by Anderson.[11]

ACKNOWLEDGMENTS

B.A.M. acknowledges financial support from the Australian Department of Education, Employment, and Training. T.C.C. acknowledges Professor M.P. Das and the workshop on high-T_c superconductivity at the Australian National University where part of Sec. IV was written.

APPENDIX A: CORRESPONDENCE BETWEEN ENERGY DISPERSION OF TB AND NONORTHOGONAL HAMILTONIANS

1. Square and cubic lattices

The series expansion of the energy dispersion in the overlap between nearest neighbors is

$$E(\vec{k}) = h_0 + (h_1 - h_0S)\mu(\vec{k}) - S(h_1 - h_0S)\mu(\vec{k})^2 + \cdots. \tag{A1}$$

In the square lattice $\mu(\vec{k}) = 2(\cos k_x a + \cos k_y a)$, where

a is the lattice parameter. Hence the energy dispersion to first order in nearest-neighbor overlap is

$$E(\vec{k}) = (h_0 - 4Sh_1) + (h_1 - Sh_0)(2\cos k_x a + 2\cos k_y a) - 2Sh_1[2\cos(k_x a + k_y a) + \cos(k_x a - k_y a)] - Sh_1(2\cos 2k_x a + 2\cos 2k_y a). \quad \text{(A2)}$$

The corresponding TB energy dispersion is

$$E(\vec{k}) = \alpha + \beta(2\cos k_x a + 2\cos k_y a) + \gamma[2\cos(k_x a + k_y a) + 2\cos(k_x a - k_y a)] + \delta(2\cos 2k_x a + 2\cos 2k_y a). \quad \text{(A3)}$$

In the cubic lattice, the expansion of the energy dispersion to first order in the nearest-neighbor overlap also corresponds to the inclusion of third-nearest-neighbor TB hopping terms. The correspondence is as follows:

$$\alpha \leftrightarrow h_0 - 6Sh_1,$$
$$\beta \leftrightarrow h_1 - Sh_0,$$
$$\gamma \leftrightarrow -2Sh_1,$$
$$\delta \leftrightarrow -Sh_1. \quad \text{(A4)}$$

2. Triangular lattice

In the triangular lattice $\mu(\vec{k})$ has the following form:

$$\mu(\vec{k}) = 2\cos k_x 2a + 2\cos(k_x a + k_y b) + 2\cos(k_x a - k_y b). \quad \text{(A5)}$$

As in the previous section, the expansion of $E(\vec{k})$ to first order in nearest-neighbor overlap, S, gives

$$E(\vec{k}) = (h_0 - 6Sh_1) + [(1-2S)h_1 - Sh_0] \times (2\cos 2k_x a + 4\cos k_x a \cos 2k_y b) - 2Sh_1(2\cos 2k_y b + 4\cos 3k_x a \cos k_y b) - Sh_1(2\cos 4k_x a + 4\cos 2k_x a \cos 2k_y b). \quad \text{(A6)}$$

The TB energy dispersion for the triangular lattice, including the third-nearest-neighbor hopping terms is

$$E(\vec{k}) = \alpha + \beta(2\cos 2k_x a + 4\cos k_x a \cos k_y b) + \gamma(2\cos 2k_y b + 4\cos 3k_x a \cos k_y b) + \delta(2\cos 4k_x a + 4\cos 2k_x a \cos 2k_y b). \quad \text{(A7)}$$

Comparison of Eqs. (A6) and (A7) shows that as for the square lattices, nonorthogonality effects persist up to the third nearest neighbor. For the hexagonal lattice, see Sec. III.

[1] J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1963), Vol. 4.
[2] J. Vergés and F. Yndurian, Phys. Rev. B **37**, 4333 (1988).
[3] M. Menon and K. Subbaswamy, Phys. Rev. B **50**, 11 577 (1994).
[4] W. Harrison, *Electronic Structure and Properties of Solids* (Freeman, San Francisco, 1980).
[5] C. Varma, E. Blount, P. Vashishta, and W. Weber, Phys. Rev. B **19**, 6130 (1979).
[6] P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950).
[7] D. A. Mirabella, C. M. Aldao, and R. R. Deza, Am. J. Phys. **62**, 162 (1994).
[8] D. Mirabella, C. Aldao, and R. Deza, Phys. Rev. B **50**, 12 152 (1994).
[9] J. Slonczewski and P. Weiss, Phys. Rev. **109**, 272 (1958).
[10] J. McClure, Phys. Rev. **108**, 612 (1957).
[11] P. W. Anderson, Phys. Rev. Lett. **18**, 1049 (1967).
[12] The extended Hückel method [R. Hoffmann, J. Chem. Phys. **39**, 1397 (1963)] makes the additional approximation of assuming that the off-diagonal elements of the Hamiltonian are proportional to the overlap between nearest neighbors.
[13] G. S. Painter and D. E. Ellis, Phys. Rev. B **1**, 4747 (1970).
[14] Note that the energies are all measured with respect to the on site Hamiltonian matrix element $\langle k_A|\mathcal{H}|k_A\rangle$, which is taken to be equal to $\langle k_B|\mathcal{H}|k_B\rangle$.
[15] J.-C. Charlier, X. Gonze, and J.-P. Michenaud, Phys. Rev. B **43**, 4579 (1991).
[16] J.-C. Charlier, X. Gonze, and J.-P. Michenaud, Carbon **32**, 289 (1994).
[17] C. Frétigny, R. Saito, and H. Kamimura, J. Phys. Soc. Jpn. **58**, 2098 (1989).
[18] E. Clementi and L. Roetti, At. Data Nucl. Data Tables **14**, 178 (1974).
[19] Such anisotropy presents a problem for TB band structure calculations, such as that employed by Whittle *et al.*, J. Phys. Condens. Matter **5**, 6555 (1993), where ζ is treated as a parameter to be determined self-consistently.
[20] B. A. McKinnon and T. C. Choy, Aust. J. Phys. **46**, 601 (1993).
[21] A. Law, M. T. Johnson, and H. Hughes, Phys. Rev. B **34**, 4289 (1986).
[22] D. Marchand *et al.*, Phys. Rev. B **30**, 4788 (1984).
[23] T. Takahashi, H. Tokailin, and T. Sagawa, Phys. Rev. B **32**, 8317 (1985).
[24] R. C. Tatar and S. Rabii, Phys. Rev. B **25**, 4126 (1982).
[25] W. W. Toy, M. S. Dresselhaus, and G. Dresselhaus, Phys. Rev. B **15**, 4077 (1977).
[26] A. Misu, E. Mendez, and M. S. Dresselhaus, J. Phys. Soc. Jpn. **47**, 208 (1979).
[27] D. E. Soule, J. W. McClure, and L. B. Smith, Phys. Rev. **134**, A453 (1964).
[28] R. E. Doezema, W. R. Datars, H. Schaber, and A. Van Schyndel, Phys. Rev. B **19**, 4224 (1979).
[29] E. Mendez, A. Misu, and M. S. Dresselhaus, Phys. Rev. B **21**, 827 (1980).