

Orbital nonorthogonality effects in band structures: Bond orbital model

J. Tejeda and N. J. Shevchik*

Max-Planck-Institut für Festkörperforschung, Stuttgart, Federal Republic of Germany

(Received 12 August 1975)

An approximate method to correct band structures for the effects of the nonorthogonality of a basis set consisting of degenerate orbitals is derived and applied to the bond-orbital model for tetrahedrally coordinated semiconductors. It is shown that the band structure corrected for the nonorthogonality of the basis orbitals is approximately related through a simple transformation to the band structure calculated by disregarding the nonorthogonality. Expansion of this transformation shows that nearest-neighbor overlap can produce terms in the band structure which are similar to those produced by more distant-neighbor interactions. For the bond-orbital model, the most important effect of nonorthogonality correction is to increase the curvature of the p -like bands at Γ , bringing the overall agreement of the model band structure into still better agreement with experiment.

I. INTRODUCTION

In tight-binding energy band-structure calculations, the nonorthogonality of the basis orbitals on different sites introduces computationally troublesome energy-dependent off-diagonal terms into the secular equations.¹ Here, we show that the nonorthogonality of *degenerate* atomic orbitals on adjacent sites can be taken into account approximately, so that the band structure including nonorthogonality effects can be expressed analytically in terms of that calculated by disregarding the nonorthogonality. It is shown that the effects of nonorthogonality on the band structure are similar to those induced by more distant-neighbor interactions. The most notable improvement of the energy bands of the bond-orbital model²⁻⁹ brought about by consideration of the nonorthogonality is the decrease of the effective masses of the p -like bands at Γ , while the over-all good description of the band structure achieved previously is still maintained.

II. NONORTHOGONALITY CORRECTION: FORMALISM

In this section, we discuss an approximate, but simple, method to correct the band structures for the neglect of the nonorthogonality of the degenerate basis set. Although we shall apply the results of this section to the bond-orbital model, the formalism is more general and can be applied to other cases, such as d -band metals.

We assume that the solution to Schrödinger's equation can be expressed as a linear combination of atomic orbitals. The general matrix equation operating on the coefficients of the atomic orbitals for an arbitrary point in k space including nonorthogonality of the basis set takes the form

$$I(E - E_0)\psi = [H - S(E - E_0)]\psi, \quad (1)$$

where H and S are the Coulomb-interaction and charge-density overlap matrices, respectively, be-

tween orbitals that lie on different sites. Therefore, the H and S matrices contain no diagonal terms involving the same orbitals on the same site, but do contain nonvanishing H_{ii} and S_{ii} terms which involve the same type of orbital, but each type of orbital on a different site. The symbol I is the identity matrix. The constant E_0 corresponds to the energy of the isolated atomic orbitals, which are all degenerate. The constant E corresponds to the energy eigenvalues of the matrix equation.

It is difficult to solve Eq. (1) in general since the S_{ij} terms produce off-diagonal energy-dependent terms; usually, these terms are neglected in order to simplify the calculations. The energy-dependent off-diagonal terms may be eliminated with the Löwdin scheme,¹⁰ which involves a recalculation of the interaction matrix H as a power series in the overlap parameters and therefore is still only approximate. The scheme that we shall introduce is far simpler, but not as generally applicable as Löwdin's scheme.

We now add and subtract $\rho H(E - E_0)$ from Eq. (1) to obtain

$$I(E - E_0)\psi = (H_0 + \Delta H)\psi, \quad (2)$$

where

$$H_0 = H[1 - \rho(E - E_0)],$$

$$\Delta H = \rho H(E - E_0) - S(E - E_0),$$

and ρ is a constant to be chosen so as to minimize the effects of the nonorthogonality. The motivation for manipulating the matrix equation in this manner is that there should be some degree of proportionality between the charge-density overlap matrix and the interaction matrix. When they are exactly proportional, there is some value of ρ for which ΔH vanishes and H_0 becomes the exact matrix operating on ψ .

If the effect of ΔH on the eigenvalues of H_0 can be neglected, the eigenvalues Eq. (1) would be given by

$$I(E - E_0)\psi = H_0\psi. \quad (3)$$

Since $E - E_0$ and $1 - \rho(E - E_0)$ are constants, we can divide both sides of Eq. (3) by the factor $1 - \rho(E - E_0)$ to obtain

$$IE'\psi = H\psi, \quad (4)$$

where

$$E' = \frac{(E - E_0)}{1 - \rho(E - E_0)}.$$

Equation (4) is precisely the form of the matrix equation obtained under the assumption that the overlap terms are zero, i. e., the basis set is orthogonal. We have chosen the zero point of the energy E to be equal to that of the isolated orbital. Since the energy of the atomic orbital E_0 has been absorbed into the definition of E' , E' cannot be arbitrarily shifted but is now defined by the absolute values obtained by solving Eq. (4). The eigenvalues of Eq. (3) are related to those of Eq. (4) by

$$E = \frac{E'}{1 + \rho E'} + E_0. \quad (5)$$

Thus, the charge-density overlap is very easy to include, since the band structure calculated, taking into account the nonorthogonality, can be expressed in terms of that calculated disregarding the nonorthogonality. The nonorthogonality correction does not depend upon the detailed nature (e. g., symmetry) of the wave function, but only upon its energy. Furthermore, the eigenfunctions of Eqs. (3) and (4) are identical.

A crucial restriction needed in the derivation of the transformation is that all of the isolated atomic orbitals have the same energy and that crystal-field splitting does not remove their degeneracy. If nondegenerate orbitals were considered, the matrix IE_0 would be replaced by another diagonal matrix in which the elements would not all be the same. For this case, it is obvious that the simple transformation given in Eq. (4) is not applicable.

We have been unable to derive a simple transformation for the general case of an arbitrary number of nondegenerate orbitals. A quadratic transformation can be derived when sets of orbitals having two energies are used, provided that these two sets of orbitals interact with each other, but not among themselves. It appears that the general transformation, if it exists, is a polynomial of order N , where N is the number of different orbital energies.

We now discuss a procedure by which ρ can be chosen. When the eigenfunctions for Eq. (4) have been calculated, we might wish to minimize some average moment of the matrix ΔH over the energy bands. For example, by requiring that

$$\Delta = \sum_{i,j} |\langle \psi_i | \Delta H | \psi_j \rangle|^2$$

be a minimum, we obtain

$$\rho = \frac{\sum_i E'_i S_{ii}}{\sum_i E_i'^2}.$$

This definition of ρ depends only upon the diagonal components of S represented in terms of the eigenfunctions of (4). In the case for which

$$S = \lambda H,$$

i. e., the overlap matrix is proportional to the interaction matrix, choosing $\rho = \lambda$ makes ΔH vanish and the overlap correction as given by Eq. (5) becomes exact. If the interaction and overlap matrices are dominated by one term, S_{\max} and H_{\max} , then ρ can be taken approximately as

$$\rho \cong S_{\max} / H_{\max}.$$

From this simple expression we see that ρ is expected to be a negative quantity, since a positive charge-density overlap usually implies a corresponding negative interaction integral.

III. GENERAL FEATURES OF THE NONORTHOGONALITY CORRECTION

In this section, we discuss the general features of the nonorthogonality correction as given by Eq. (5). Section IV shall present some numerical examples on the bond-orbital model which demonstrate these features more explicitly.

As a check on our formalism, we consider the trivial case of two equivalent orbitals interacting via a potential matrix element V (e. g., the hydrogen molecule); the resulting energies are according to Eq. (5)

$$E = \pm V(1 \mp \rho V). \quad (6)$$

This gives back the well-known result that the effects of the overlap are to increase the energy of the antibonding states more than that of the bonding states.¹¹

Since the correction for nonorthogonality depends only upon the energy, the density of states including nonorthogonality corrections, $N(E)$, is conveniently related to the old density of states deduced by disregarding the nonorthogonality, $N'(E')$, by the expression

$$N(E) = N' \left(\frac{E - E_0}{1 - \rho(E - E_0)} \right) \frac{1}{[1 - \rho(E - E_0)]^2}. \quad (7)$$

Since ρ is negative, the bands are stretched in energy where $E(k) > E_0$, and flattened where $E(k) < E_0$. As consequence of the fact that the bands are stretched more than they are flattened, the center of mass of the corrected density of states no longer corresponds to that of the isolated atom, but is shifted upward in energy by $\sim -\rho \langle (E')^2 \rangle$.

It is easy to show by differentiation of Eq. (5)

that the effective masses at all critical points corrected for the nonorthogonality are given by

$$\frac{1}{m} = \left(\frac{1}{m}\right)' \frac{1}{(1 + \rho E')^2}, \quad (8)$$

where $(1/m)'$ and E' are the effective mass and energy calculated ignoring the overlap terms. In general, for $E'(0) > 0$, the overlap correction increases both the bandwidths and their curvature. However, comparison of Eqs. (5) and (8) shows that the curvature at the top of the band is increased by a larger factor than the corresponding bandwidth.

In the tight-binding approximation for the s band of a linear chain of atoms, with interactions restricted to nearest neighbors only, the form of the energy dispersion neglecting the nonorthogonality of the s orbitals is

$$E'(k) = V \cos ka,$$

where a is the nearest-neighbor separation. Taking into account the nonorthogonality, the band structure becomes via Eq. (5)

$$E(k) = \frac{V \cos ka}{1 + \rho V \cos ka}. \quad (9)$$

By expanding the denominator, one finds that the band structure as given by Eq. (9) contains all Fourier components, and not just a single $\cos ka$ term. When the nonorthogonality is neglected, such as in the Slater-Koster method,¹ higher-order Fourier terms arise only when interactions between more distant neighbors are taken into account. As this example demonstrates, however, similar higher-order Fourier components can arise from nearest-neighbor overlap terms. This points out the possible dangers of interpreting correctly calculated band structures in terms of a simplified tight-binding model in which overlap effects are not properly taken into account.

IV. OVERLAP CORRECTIONS IN THE BOND-ORBITAL MODEL

In the remainder of this paper, we shall illustrate the effects of the overlap correction by considering the band structures of Si derived from the bond-orbital model. Recent works^{7,9} have shown that the bond-orbital model, with just a few adjustable parameters extended to second-neighbor bonds, gives a surprisingly accurate description of the energy bands calculated with much larger secular equations. The only major defects of the model are in its failure to describe the middle p -like band along the Σ axis and the effective masses of the p bands at Γ .^{3,10} Pantelides and Harrison⁷ suggest that these defects are due to the neglect of interactions between more distant bond orbitals. Here we show that inclusion of the nonorthogonality brings the

effective masses at Γ into better agreement with more sophisticated calculations.

The basis set of the bond-orbital model consists of four tetrahedral bonds in the unit cell, which connect each atom with its four nearest neighbors. In terms of atomic orbitals, such bonding orbitals are constructed by taking tetrahedral sp^3 hybrids on each type of atom and forming the bonding combination between nearest-neighbor atoms. The antibondinglike combination is projected to much higher energy and can therefore be neglected when considering only the valence bands.

The valence-band structure of Ge and Si which we shall consider here has already been presented in Ref. 9 using only two interaction parameters: V_0 , the interaction between two bond orbitals meeting at an atomic vertex, and V_2 , the interaction between second-neighbor bond orbitals that are parallel to one another.

In the energy bands of the bond-orbital model for Si to be presented, we have chosen a value of $\bar{\rho} = 0.1 \text{ eV}^{-1}$, which is about twice as large as that which we calculate for the nearest-neighbor bond overlap of Si using the overlap tables of Mulliken *et al.*¹² We may consider this parameter to be absorbing to some extent the effects of interactions with more-distant neighbors and of basis functions not explicitly included. Our primary purpose here is not for quantitative accuracy, however, but rather only to display the qualitative effects of the nonorthogonality correction.

In Fig. 1, we compare the band structure calculated for $V_0 = -1.5 \text{ eV}$, $V_2 = 0$, with $\bar{\rho} = 0$ and -0.1 eV^{-1} . Since only one interaction parameter is involved, there is some $\bar{\rho}$ for which Eq. (5) is exact. Without the interaction between second-neighbor bond orbitals, V_2 , two flat p bands are obtained, as has been noted in previous works.²⁻⁹ The lower two bands are not in good agreement with photoemission experiments.¹³ The separation of the

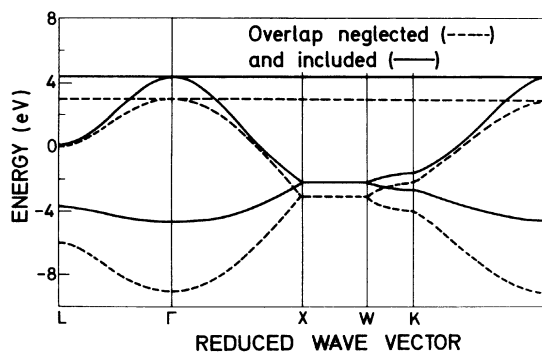


FIG. 1. Valence bands of the bond-orbital model calculated for $V_1 = -1.5 \text{ eV}$, $V_2 = 0$, with $\bar{\rho} = 0.0$ (dashed line) $\bar{\rho} = -0.1$ (solid line).

lower two bands at L is equal to one-half the total bandwidth, whereas it should be $\frac{1}{4}$ this amount; the point X at which the lower two bands meet lies midway between the two Γ points. The correction for the nonorthogonality does not remove the flat p bands, but it does decrease the separation of the bands at L and lower the X point, tending to bring the shape of the bands into better agreement with experiment. (Rescaling of the width of the corrected band structure should bring it into very good agreement with the experiment.) As will be seen shortly, the second-neighbor interaction V_2 produces similar changes in the lower bands.

If this overlap correction is primarily responsible for these improvements, then it is less likely that dihedral rotations, as suggested earlier,⁹ are important in smearing the two-peak structure in the lower valence bands of the amorphous form. The Weaire-Thorpe model,⁸ therefore, has an even greater degree of validity for the valence bands than previously believed.

In Fig. 2, the band structure calculated without the nonorthogonality correction (henceforth referred to as I) was adjusted to fit the orthogonalized-plane-wave (OPW) band structure of Herman.¹⁴ Considering the simplicity of the model, the agreement is remarkably good, except in the Σ direction where the middle p band does not dip to low enough energies. When this band structure is corrected for nonorthogonality (henceforth referred to as II) it becomes greatly distorted. The Γ_{25} and Γ_1 points move up in energy by 6 and 2.5 eV, respectively, accounting for an increase of 3.5 eV (25%) in the total bandwidth. The curvatures of the bands at Γ_{25} have more than doubled. Although the widths of the p bands have doubled, that of the s -band has decreased to about one-half its previous value. The separation of the two lower bands at L has diminished to about 1 eV and the X point is now too low in energy. It is apparent that

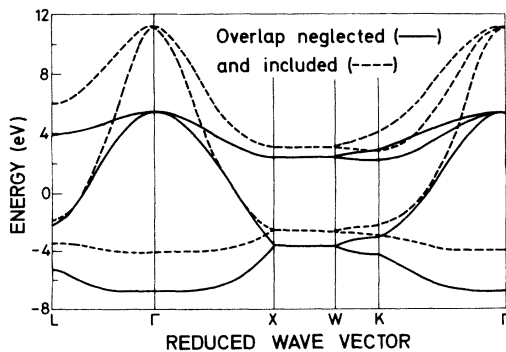


FIG. 2. Valence bands of the bond-orbital model calculated for $V_1 = -1.5$ eV, $V_2 = 0.75$ eV, with $\rho = 0$ (solid line, band structure I) and $\rho = -0.1$ (dashed line, band structure II).

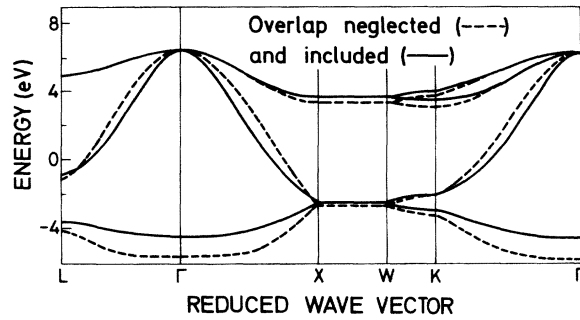


FIG. 3. Valence bands of the bond-orbital model calculated for $V_1 = -1.5$ eV, $V_2 = +0.3$ eV, $\rho = -0.1$ (solid line, band structure III) compared to that calculated in Fig. 2 for $\rho = 0$ (dashed line).

the combined effects of V_2 and the overlap correction on the X and L points are too strong, and thus some readjustment of the parameters is necessary to bring the band structure back into agreement with the results of more elaborate calculations. The overlap correction does not improve the middle p -like band along the Σ direction, as evident from inspection [Eq. (5)]. Apparently, this feature is in part due to neglected basis functions or longer-range interactions.^{15,16}

After several attempts to adjust V_2 and V_0 to bring the nonorthogonality corrected band structure II into agreement with the uncorrected one (I), we concluded that a good fit was achieved by simply reducing V_2 from 0.75 to 0.3 eV, and leaving V_0 unchanged. The readjusted band structure (henceforth referred to as III) is compared in Fig. 3 with the one calculated previously without the correction. To aid in comparison of the bands, we made the Γ_{25} point of the readjusted band coincide with that of the earlier calculation. The adjusted band structure III matches band structure I very well at the symmetry points and thus the agreement with more sophisticated band structures is partially restored. The lower s -like bands are now about 1 eV narrower than those calculated by Herman.¹⁴ We believe that inclusion of the neglected antibonding s -like orbitals might increase the width of the lower s -like bands. A slightly larger value of V_0 could have also moved the Γ_1 point of band structure III to coincide with that of band structure I, but this would have destroyed the good agreement at the X and L points. The primary improvement brought about by reducing V_2 is in the reduction of the effective masses of the p bands at Γ_{25} to about one-half of their previous values.

The shapes of the uppermost p bands of III are nearly the same as those as I, but the lower p band has a different form throughout the zone. Since the primary effect of the overlap correction is to

TABLE I. Calculated effective masses at the Γ point for different parameter values in the directions indicated (free-electron mass identical to 1).

V_0	-1.5	-1.5	-1.5	-1.5
V_2	0.75	0.75	0.3	0.3
ρ	0	-0.1	0	-0.1
Λ				
m_1	0.948	0.231	2.32	0.863
m_2	0.948	0.231	2.32	0.863
m_3	0.239	0.053	0.378	0.140
Δ				
m_1	0.641	0.144	1.608	0.598
m_2	0.641	0.144	1.608	0.598
m_3	0.318	0.072	0.453	0.169
Σ				
m_1	1.264	0.285	3.33	1.239
m_2	0.632	0.142	1.55	0.576
m_3	0.254	0.057	0.395	0.147

change the curvature of the bands, we have listed in Table I the effective masses at Γ along the Δ , λ , and Σ axes before and after the overlap correction for the various interaction parameter values used. Inspection of this table shows that the overlap correction can account for a large portion of the effective masses, particularly for the light- p -hole band. The net change (i.e., difference between band structures I and III) in the heavy mass due to the overlap correction is only 10%, since it was necessary to reduce V_2 by one-half of its original value to maintain the correct p band width, however, the overlap corrected light mass is one-half of the uncorrected one. Both the heavy- and light-hole masses are now in better agreement with experimental values.¹⁷ A different choice of parameters could reproduce the small effective masses in Ge, but at the expense of flattening the s bands. In terms of this model, the much smaller effective masses of Ge compared to Si might be explained by a larger overlap parameter; the effective masses are extremely sensitive to ρ , partic-

ularly when $1 + \rho E \rightarrow 0$, as can be seen in Eq. (8). However, as already mentioned, other factors, such as interactions between more distant neighbors³ and interactions with higher-lying orbitals are probably still contributing significantly to the effective masses. In this case, ρ loses its physical significance and becomes nothing more than a parameter which attempts to absorb the effects of these interactions.

V. SUMMARY AND CONCLUSIONS

It has been shown that the nonorthogonality of degenerate orbitals can be incorporated into band-structure calculations, in an approximate and simple manner. When the overlap matrix is proportional to the interaction matrix, the procedure outlined here is exact. The primary effect of this correction is to change the functional form of the bands so that the effective masses of the top of the bands are decreased. We suggest that the failure of the bond-orbital models to predict correctly the effective masses at Γ is in part due to the neglect of the charge-density overlap, and not entirely due to the neglect of interactions between more-distant neighbor atoms and higher-lying bands.

From the present results, we conclude that the bond-orbital model, including nonorthogonality of the basis set, is indeed faster converging (in the usual tight-binding sense) than originally believed. The omission of such nonorthogonality terms must be picked up on interactions between more-distant neighbors in a Slater-Koster-type scheme, which ignores the overlap. We suggest that the overlap correction introduced here should also be useful for bands derived from d orbitals.

ACKNOWLEDGEMENTS

We thank Dr. G. Guntherodt, Dr. P. Kemeny, Dr. L. Ley, and Professor M. Cardona for reading the manuscript and for their useful comments.

*Present address: Physics Dept., State University of New York, Stony Brook, N. Y. 11794.

¹J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

²G. G. Hall, Philos. Mag. **43**, 338 (1952); **3**, 429 (1958).

³D. Stocker, Proc. Ro. Soc. Lond. **270**, 397 (1962).

⁴G. Lehman and J. Friedel, J. Appl. Phys. Suppl. **1**, 281 (1962).

⁵W. A. Harrison, Phys. Rev. B **10**, 4487 (1973).

⁶W. A. Harrison and S. Ciraci, Phys. Rev. B **10**, 1516 (1974).

⁷S. T. Pantelides and W. A. Harrison, Phys. Rev. B **11**, 3006 (1975).

⁸D. Weaire and M. F. Thorpe, Phys. Rev. B **4**, 2508 (1971).

⁹N. J. Shevchik, J. Tejada, and M. Cardona, Phys. Rev. B **9**, 2627 (1974).

¹⁰P. O. Löwdin, J. Chem. Phys. **18**, 365 (1970).

¹¹C. J. Ballhausen and H. B. Gray, *Molecular Orbital Theory* (Benjamin, New York, 1964), p. 20.

¹²R. S. Mulliken, C. A. Rieke, D. Orloff, and H. Orloff, J. Chem. Phys. **17**, 1248 (1949).

¹³L. Ley, S. Kowalczyk, R. Pollak, and D. A. Shirley, Phys. Rev. Lett. **29**, 1088 (1972).

¹⁴F. Herman, R. L. Kortum, C. D. Kuglin, and J. P. Van Dyke, in *Methods in Computation Physics*, edited by B. Adler, S. Fernbach, and M. Rotenberg (Academic, New York, 1968).

¹⁵E. A. Kraut, Solid State Commun. **13**, 923 (1973).

¹⁶D. J. Chadi and M. L. Cohen, Phys. Status Solidi B **68**, 405 (1975).

¹⁷The shape of the constant-energy surface has been accurately determined for Si by cyclotron resonance and the average value for the light and heavy masses are 0.50 and 0.15 (free-electron mass identical to 1) [J. C. Hensel and G. Feher, Phys. Rev. **129**, 1041 (1963)].