

Explicit, First-Principles Tight-Binding Theory

O. K. Andersen and O. Jepsen

Max-Planck-Institut für Festkörperforschung, D-7000 Stuttgart 80, Federal Republic of Germany

(Received 18 July 1984)

The minimal base of muffin-tin orbitals is transformed exactly into a tight-binding base. The linear transformations, the orbitals, and the Hamiltonian, overlap, and Green's function matrices are expressed in terms of one matrix, the canonical structure matrix S_{ij} . It vanishes beyond second-nearest neighbors and is tabulated. Tight-binding two-center forms with transfer integrals proportional to S_{ij} are derived.

PACS numbers: 71.10.+x, 71.25.Cx

The simplest and most widely used one-electron Hamiltonian is that of the *tight-binding* (TB) method with a *minimal* base, that is, with at most one s , three p , and five d orbitals per atom, and in its *two-center* approximation.¹ Yet, this method has remained an empirical rather than a first-principles method of the kind needed for self-consistent density-functional calculations. This, for instance, is a reason why no such calculations exist for amorphous solids or for infinite solids with extended defects, where a TB recursion,² moment, or Green's-function³ technique seems to be required.

We have now succeeded in transforming the minimal but long-ranged base of muffin-tin orbitals (MTO's)⁴ frequently used in first-principles band-structure calculations into a TB base. Since the transformation is *exact*, all results hitherto obtained with the linear MTO method⁴—or with the almost identical augmented spherical wave method⁵—would be reproduced by the new TB method. The TB-MTO's have an almost *universal*¹ decay and, with the so-called atomic spheres approximation (ASA),⁴ our TB method can be cast in two-center form. In the following we give a self-contained derivation of our theory and discuss it.⁶

A *minimal* set of orbitals (like the set of partial waves in scattering theory) possesses only *one* radial function per site, $\bar{\mathbf{R}}$, and angular momentum, lm . Each orbital must therefore approximately satisfy Schrödinger's differential equation in the region *between* the atoms. Here the potential is flat on a scale of 1 Ry and, since the energy range of interest begins near the point where the electron can pass classically between the atoms and extends upwards by about 1 Ry, it seems natural to choose orbitals which have zero kinetic energy, i.e., satisfy the Laplace equation, in the interstitial region. This is the choice made for solid-state MTO's. A *conventional* MTO, χ_{lm}^0 , is proportional to the 2^l -pole field $r^{-l-1}Y_{lm}(\hat{r})$ in the interstitial region. However, an MTO need not be a solution of Laplace's equation

inside the neighboring atoms where it will be modified through augmentation. We therefore try to *localize the MTO's by screening with multipoles added on the neighboring sites*.

The formalism is as follows: We define the regular and irregular radial Laplace solutions $J_l^0(r) = [2(2l+1)]^{-1}(r/w)^l$ and $K_l(r) = (r/w)^{-l-1}$, as well as a modified function, $J_l^\alpha(r) = J_l^0(r) - \alpha_l K_l(r)$, to which the amount $-\alpha$ of irregular solution has been added. w is some length scale introduced to make the functions and the *screening constants* α dimensionless. We now denote $|J_{\bar{\mathbf{R}}lm}^0\rangle$, $|K_{\bar{\mathbf{R}}lm}\rangle$, and $|J_{\bar{\mathbf{R}}lm}^\alpha\rangle$ as the radial functions multiplied by Y_{lm} and centered at site $\bar{\mathbf{R}}$. These functions are furthermore defined to vanish outside some Wigner-Seitz cell or sphere centered at $\bar{\mathbf{R}}$. The *bare* 2^l -pole field at $\bar{\mathbf{R}}_j$, extending in *all* space, is denoted $|K_j^0\rangle$, where j is a shorthand notation for $\bar{\mathbf{R}}_j l_j m_j$. This bare field equals $|K_j\rangle$ inside its own cell, and inside any other cell, say the one at $\bar{\mathbf{R}}_i$, it may be expanded in an $l_i m_i$ series of the regular Laplace solutions $|J_i^0\rangle$, i.e.,

$$|K_j^0\rangle = \sum_i (|K_i\rangle \delta_{ij} - |J_i^0\rangle S_{ij}^0). \quad (1)$$

The expansion coefficients S^0 form a Hermitian matrix which is dimensionless and independent of the scale of the structure. This is the so-called (bare) *canonical structure matrix*.⁴ It vanishes for $\bar{\mathbf{R}}_i = \bar{\mathbf{R}}_j$ and, as a function of the interatomic distance $d = |\bar{\mathbf{R}}_i - \bar{\mathbf{R}}_j|$, it decays like $(w/d)^n$ with $n = l_i + l_j + 1$. In analogy to (1) we define the screened field by

$$|K^\alpha\rangle = |K\rangle - |J^\alpha\rangle S^\alpha, \quad (2)$$

where we have dropped the subscripts and regarded the functions as components of row vectors. We see that the screened field is given by the superposition

$$|K^\alpha\rangle = |K^0\rangle (1 + \alpha S^\alpha) = |K^0\rangle (1 - \alpha S^0)^{-1} \quad (3)$$

of bare fields, and that the relation between the bare and screened structure matrices is

$$S^\alpha = S^0(1 - \alpha S^0)^{-1}, \quad (4)$$

where α is now regarded as a diagonal matrix with elements $\alpha_{\vec{R}_l}$. From the first Eq. (3) we identify αS^α as the "screening charge," that is, $\alpha_l S_{ij}^\alpha$ is the relative strength of the 2^l pole at \vec{R}_i screening the 2^j pole at \vec{R}_j . As a function of d/w the decay of S^α must be exponential for sufficiently small but positive α values because the spectrum of S^0 (which is essentially the so-called canonical bands⁴) is upwards bound.

We now determine α such that S^α has the shortest possible range. Since normally we only want to include the set of s , p , and d MTO's in the superposition (3), we choose $\alpha = 0$ for $l > 2$. Moreover, since we consider infinite systems and S^0 has infinite range, we must, in order to perform the matrix inversion in (4), first consider periodic structures. We thus let α be the same on all lattice sites and are left with three parameters, α_s , α_p , and α_d which we determine numerically by trial and error. For an assumed lattice and α , we invert the small Hermitian matrix $\alpha^{-1} - S^0(\vec{k})$ as a function of the Bloch vector \vec{k} . Subsequent Fourier transformation then yields S^α in \vec{R} space. Our investigations for the fcc, bcc, and sc lattices yield three remarkable results:

(1) The range of S^α may be limited to essentially *second-nearest neighbors*. (2) The corresponding α

(Table I) is independent of the structure provided that w is taken as the average Wigner-Seitz radius and, hence, measures the *density* of multipoles. (3) The optimally screened structure matrix (Table I) expressed as two-center integrals with the z axis along the interatomic vector, \vec{d} , depends almost only on d/w and hardly on the structure. In other words, S^α is almost *universal*.¹

As a consequence, one may obtain a first estimate of S^α for any reasonably homogeneous structure simply by interpolation⁷ in Table I, and this has been used for amorphous systems.⁸ For $1.6 < d/w < 2.1$ the matrix elements of S^α decay roughly like $\exp(-\lambda d/w)$ with $\lambda = 3.9$ ($ss\sigma$), 3.5 ($sp\sigma$), 3.0 ($sd\sigma$), 3.3 ($pp\sigma$), 6.7 ($pp\pi$), 3.0 ($pd\sigma$), 6.1 ($pd\pi$), 3.1 ($dd\sigma$), 5.4 ($dd\pi$), and 7.9 ($dd\delta$). Plots of $S^\alpha(d/w)$, and a more accurate interpolation formula, will be published in Ref. 6. Since S^α specifies the screened field through the one-center expansions (2) and since this field, appropriately normalized, will be the envelope of the tail of the TB orbital, the latter has a nearly universal decay. This decay is relevant for the dependence of the electronic structure on shear (w constant) whereas the dependence on compression (\vec{R}/w constant) enters through the w dependence of the orbital normalization factor N^α to be defined below.

For a given α , which needs not be the one for optimal screening, we now form a set of MTO's by augmenting the corresponding set of multipoles (3)

TABLE I. Localized canonical two-center integrals, S^α , with $\alpha_s = 0.3485$, $\alpha_p = 0.05303$, and $\alpha_d = 0.0107$.

lattice	shell	d/w ^a	$ss\sigma$	$sp\sigma$	$sd\sigma$	$pp\sigma$	$pp\pi$	$pd\sigma$	$pd\pi$	$dd\sigma$	$dd\pi$	$dd\delta$
sc	1	on-site	3.72			3.71				5.89 ^b	1.42 ^c	
bcc	1	on-site	3.09			2.79				1.30 ^b	2.71 ^c	
fcc	1	on-site	3.05			2.74				1.67 ^b	2.37 ^c	
sc	6	1.6120	-0.961	1.84	-2.10	3.62	-0.76	-4.41	1.61	-6.17	3.52	-0.54 ^d
bcc	8	1.7589	-0.593	1.18	-1.42	2.36	-0.36	-2.93	0.82	-3.84	1.85	-0.19
fcc	12	1.8049	-0.484	0.98	-1.22	2.00	-0.25 ^d	-2.52	0.60 ^d	-3.32	1.42 ^d	-0.13 ^d
bcc	6	2.0310	-0.203	0.44	-0.60	0.93	-0.05	-1.29	0.13	-1.76	0.36	-0.02 ^d
sc	12	2.2797	-0.057	0.12	-0.15	0.26	-0.03 ^d	-0.34	0.06 ^d	-0.50	0.11 ^d	-0.02 ^d
fcc	6	2.5589	-0.020	0.04	-0.06	0.09	0.00	-0.14	-0.01	-0.23	-0.02	-0.00 ^d

^a(Interatomic distance)/(lattice Wigner-Seitz radius).

^b e_g .

^c t_{2g} .

^dAverage value (Ref. 7).

inside each atom: In the one-center expansions (2) we substitute for each radial function J_l^α by some regular function, $\bar{J}_{\bar{R}l}^\alpha$, and each radial function K_l by a linear combination of $\bar{J}_{\bar{R}l}^\alpha$ and the regular solution, $\phi_{\bar{R}l}(E)$, at energy E of the radial Schrödinger equation for the potential spherically averaged around site \bar{R} . The functions match continuously and differentiably at a MT sphere surrounding the atom so that $|K\rangle = |\phi(E)\rangle N^\alpha(E) + |J^\alpha\rangle P^\alpha(E)$, at and near the sphere surface, and the elements of the diagonal matrices P and N are

$$P^\alpha(E) = \frac{\langle \phi(E), K \rangle}{\langle \phi(E), J^\alpha \rangle} = \frac{P^0(E)}{1 - \alpha P^0(E)}, \quad (5)$$

$$N^\alpha(E) = \frac{\langle J^\alpha, K \rangle}{\langle J^\alpha, \phi(E) \rangle} = (w/2)^{1/2} \dot{P}^\alpha(E)^{1/2}. \quad (6)$$

Here $\{. . .\}$ denotes a Wronskian at the sphere, $\dot{X} = \partial X / \partial E$, and $\phi(E)$ is normalized to unity in its sphere such that $\langle \phi(E) | \phi(E) \rangle = 0$ and $\langle \phi(E), \phi(E) \rangle = -1$. $P_{\bar{R}l}^0(E)$ is the conventional potential function,⁴ which is essentially the cotangent of the phase shift. The set of energy-dependent MTO's, $|\chi^\alpha(E)\rangle$, thus equals $|K^\alpha\rangle$ in the interstitial region and

$$|\phi(E)\rangle N^\alpha(E) + |\bar{J}^\alpha\rangle [P^\alpha(E) - S^\alpha] \quad (7)$$

in the spheres; the linear combination $|\chi^\alpha(E)\rangle u^\alpha$, specified by a column vector u^α , is seen to be a solution of Schrödinger's equation at energy E for the MT potential if it equals the one-center expansions $|\phi(E)\rangle N^\alpha(E) u^\alpha$ in the spheres, i.e., if the set of linear homogeneous equations $[P^\alpha(E) - S^\alpha] u^\alpha = 0$ has a proper solution. This is the generalization of the so-called tail cancellation or Korringa-Kohn-Rostoker (KKR) condition.⁴ Each set, $|\chi^\alpha(E)\rangle$, is thus complete⁹ for the MT potential, which equals the spherical average in the MT spheres and the constant E in between. The non-MT deviations may be treated variationally by including them in the Hamiltonian and overlap matrices given by (14) and (15) below. For closely packed structures it usually suffices to include the deviation $V - E$ in the interstitial region and, in the ASA,⁴ this is simply done by substituting the MT spheres by "space-filling" atomic Wigner-Seitz spheres whose overlap is neglected such that the KKR condition remains valid.

The secular matrix $P^\alpha(E) - S^\alpha$ depends on the potential only through the potential functions along the diagonal and, for the most localized set, it has the TB two-center form with S^α playing the role of the transfer integrals. The corresponding Green's-

function matrix is

$$g^\alpha(E) = [P^\alpha(E) - i\epsilon - S^\alpha]^{-1}. \quad (8)$$

The potential functions may be parametrized like

$$\alpha + P^\alpha(E)^{-1} = P^0(E)^{-1} \approx \gamma + \Delta / (E - C), \quad (9)$$

as obtained from (5) by expansion of the radial wave function in a Taylor series of energy to first order about some energy, E_ν , at the center of interest. $C_{\bar{R}l}$, $\Delta_{\bar{R}l}$, and $\gamma_{\bar{R}l}$ are the conventional potential parameters^{4,10} describing, respectively, the position, width, and shape of the $\bar{R}l$ band. These potential parameters have been tabulated in Refs. 4 and 6 for nearly all elemental metals and at different lattice constants. Under uniform compression Δ_l , for instance, decays roughly as w^{-2l-1} . The KKR equations have the form of an eigenvalue problem if P^α is a linear function of E . This is true if we choose α equal to the potential parameter γ in which case the effective two-center Hamiltonian is seen to be

$$H_{ij}^\gamma = C_i \delta_{ij} + (\sqrt{\Delta_i}) S_{ij}^\gamma (\sqrt{\Delta_j}). \quad (10)$$

This will turn out to be the MT or AS Hamiltonian in a base, $|\chi^\gamma\rangle$, of energy-independent, nearly orthogonal MTO's. S^γ depends on the potential through γ , and its exponential decay exhibits potential- and structure-dependent, damped oscillations as may be seen in Fig. 5 of Ref. 6. For crystals, where the matrix inversion in (4) can be performed, one may obtain S^γ from S^0 , or from S^α .

In order to treat extended defects in crystals one would compute the crystalline $g^\gamma(Z) = \sqrt{\Delta} (Z - H^\gamma)^{-1} \sqrt{\Delta}$ by standard means, and then transform it into $g^\alpha(Z)$ in the TB representation, needed for treating the extended perturbation with the techniques of Ref. 3. This transformation is

$$g^\alpha(E) = (\alpha - \gamma) \frac{P^\gamma(E)}{P^\alpha(E)} + \frac{P^\gamma(E)}{P^\alpha(E)} g^\gamma(E) \frac{P^\gamma(E)}{P^\alpha(E)}$$

which, since P is diagonal, involves no matrix multiplications but merely a rescaling of each element.¹¹

In order to obtain an eigenvalue problem also when $\alpha \neq \gamma$ energy-independent orbitals are needed. Now, $|\chi^\alpha(E)\rangle$ is independent of E in the interstitial region and, in the spheres, its first energy derivative at E_ν will vanish if we choose \bar{J}^α proportional to the energy derivative function

$$\begin{aligned} |\dot{\phi}^\alpha\rangle &= [\partial |\phi(E)\rangle N^\alpha(E) / \partial E] / N^\alpha \\ &= |\dot{\phi}^\gamma\rangle + |\phi\rangle \dot{N}^\alpha / N^\alpha = |\dot{\phi}^\gamma\rangle + |\phi\rangle o^\alpha \end{aligned} \quad (11)$$

as is seen by differentiation of (7). Omission of the variable E means that it is fixed at E_ν . Moreover, we have labeled $|\dot{\phi}\rangle$, for which $\langle \phi | \dot{\phi} \rangle = 0$, by the superscript γ because $|\dot{\phi}^\gamma\rangle$ is seen from (11) to cor-

respond to the α value for which \dot{N}^α vanishes and, according to (6) and (9), this is γ . The overlap $\langle \phi | \phi^\alpha \rangle = o^\alpha$ is a potential parameter. The set $|\chi^\alpha\rangle$ is thus complete to first order in $E - E_\nu$ and can therefore yield variational energy estimates correct to third order. If we normalize $|\chi^\alpha\rangle$ to equal $|K^\alpha\rangle N^\alpha$ in the third interstitial region it is seen to be

$$|\chi^\alpha\rangle = |\phi\rangle + |\dot{\phi}^\alpha\rangle h^\alpha, \quad (12)$$

$$h^\alpha = - (P^\alpha / \dot{P}^\alpha) + (\dot{P}^\alpha)^{-1/2} S^\alpha (\dot{P}^\alpha)^{-1/2}, \quad (13)$$

in the spheres. In this base the MT or AS parts of the overlap and Hamiltonian matrices are

$$\langle \chi | \chi \rangle = (1 + ho)(1 + oh) + hph, \quad (14)$$

$$\langle \chi | H - E_\nu | \chi \rangle = h(1 + oh), \quad (15)$$

because $(H - E)|\phi(E)\rangle = 0$ and $p = \langle (\dot{\phi}^\gamma)^2 \rangle$. Here, and in the following, all superscripts α are dropped.

The important quantities (8) and (12)–(15) are all expressed in terms of only *one* nondiagonal matrix, namely, the canonical structure matrix S or, equivalently, the *effective two-center Hamiltonian*, h , defined in terms of S and potential parameters in (13). $h + E_\nu$ has the same form as the Hamiltonian (10) and it reduces to it when $\alpha \equiv \gamma$. With α as in Table I, $h + E_\nu$ is the effective two-center TB Hamiltonian. This has shorter range than (14) and (15) which are quadratic functions of h^α . Since all energy-independent MTO sets are obtained by substitution of each radial (K, J^0) Hilbert space by the (ϕ, ϕ^γ) space, *all MTO sets span the same Hilbert space* and transform into each other according to (3) and (6). An equivalent way of expressing this linear transformation follows by inserting (11) in (12): The γ set, characterized by having $o^\gamma = 0$, is

$$|\chi^\gamma\rangle = |\chi\rangle (1 + oh)^{-1} = |\chi\rangle (1 - oh^\gamma) \quad (16)$$

in terms of any other set. According to (14), *the γ representation is orthogonal* to first order in h^γ and its Hamiltonian matrix is simply

$$\begin{aligned} \langle \chi^\gamma | H - E_\nu | \chi^\gamma \rangle &= h^\gamma = h(1 + oh)^{-1} \\ &= h - hoh + \dots \end{aligned}$$

which is (10). By applying

$$\langle \chi^\gamma | \chi^\gamma \rangle^{-1/2} = 1 - h^\gamma ph^\gamma / 2 + \dots$$

from the right- and left-hand sides we may transform into a fully orthogonal base where *the Hamiltonian is seen to be expressible as a power series in the effective two-center TB Hamiltonian, h* . Even the spectrum of h , the first term in this series, is usually a good approximation to the true spectrum

over an energy range of about 0.5 Ry around E_ν . This will be illustrated for bcc Fe in Fig. 7 of Ref. 6. This power series, together with the fact that the eigenvectors of h transform the orbital base (12) into one-center expansions for the wave functions (tail cancellation), make self-consistent calculations with recursion² or moment techniques feasible. Most recently such calculations were presented⁸ for amorphous $\text{Fe}_{80}\text{B}_{20}$.

We are grateful to O. Gunnarsson for performing the first, preliminary screening calculations.

¹See, for instance, W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

²R. Haydock, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1980), Vol. 35.

³A. R. Williams, P. J. Feibelman, and N. D. Lang, *Phys. Rev. B* **26**, 5433 (1982).

⁴O. K. Andersen, *Phys. Rev. B* **12**, 3060 (1975), and *The Electronic Structure of Complex Systems* (Plenum, New York, 1984); H. L. Skriver, *The LMTO Method* (Springer, New York, 1984).

⁵A. R. Williams, J. Kübler, and C. D. Gelatt, *Phys. Rev. B* **19**, 6094 (1979).

⁶A different and more lengthy derivation of some of the present results may be found in O. K. Andersen, O. Jepsen, and D. Glötzel, "Enrico Fermi International School of Physics, Course LXXXIX" (to be published).

⁷The reduction of S^α to axially symmetric two-center integrals is not exact because the screened field has lower than lm symmetry as a result of its halo at other sites. The full structure matrices are tabulated in Ref. 6. As Table I demonstrates, the *on-site* elements exhibit structure-dependent crystal-field splittings. For a general structure these elements may, however, be obtained from the interpolated nearly universal off-site elements and S^0 , by use of (4) in the form $S_{on}^\alpha = S_{off}^0 \alpha S_{off}^\alpha$.

⁸T. Fujiwara, *J. Non-Cryst. Solids* **61&62**, 1039 (1984).

⁹For MTO's, $\chi_{R,l}^\alpha$, *not* included in the set the formalism assumes that $P_{R,l}^\alpha(E) = \infty$ (no scattering).

¹⁰ γ is $(s/w)^{2l+1}$ times the γ defined in Refs. 4 and 6. The parametrization (9) is correct to order $(E - E_\nu)^2$ and substitution of E by $E + (E - E_\nu)^3 p$ makes it valid to third order.

¹¹The transformation follows from (4) and the result

$$N^\alpha(E)/N^\beta(E) = P^\alpha(E)/P^\beta(E) = 1 + (\alpha - \beta)P^\alpha(E)$$

derived from (5) and (6). After having obtained $g^\alpha(E)$ for the defect structure in the ASA one may wish to include deviations from the AS potential in the region of the defect via a Dyson equation in the TB representation (12). This requires a transformation from the effective (8) to the real (14) and (15) TB representation, and the result obtained by use of (8), (9), (11), (13), and (16) is

$$\langle \chi | Z - H | \chi \rangle^{-1} = \dot{P}^{1/2} [g(Z) - g(\infty)] \dot{P}^{1/2}.$$