Institute for Advanced Simulation

An Introduction to the Tight Binding Approximation – Implementation by Diagonalisation

Anthony T. Paxton

published in

Multiscale Simulation Methods in Molecular Sciences,
J. Grotendorst, N. Attig, S. Blügel, D. Marx (Eds.),
Institute for Advanced Simulation, Forschungszentrum Jülich,
NIC Series, Vol. **42**, ISBN 978-3-9810843-8-2, pp. 145-176, 2009.

http://www.fz-juelich.de/nic-series/volume42

An Introduction to the Tight Binding Approximation – Implementation by Diagonalisation

Anthony T. Paxton

Atomistic Simulation Centre
School of Mathematics and Physics
Queen's University Belfast
Belfast BT1 7NN, UK
E-mail: Tony.Paxton@QUB.ac.uk

1 What is Tight Binding?

"Tight binding" has existed for many years as a convenient and transparent model for the description of electronic structure in molecules and solids. It often provides the basis for construction of many body theories such as the Hubbard model and the Anderson impurity model. Slater and Koster call it the tight binding or "Bloch" method and their historic paper provides the systematic procedure for formulating a tight binding model.[1] In their paper you will find the famous "Slater–Koster" table that is used to build a tight binding hamiltonian. This can also be found reproduced as table 20–1 in Harrison's book and this reference is probably the best starting point for learning the tight binding method.[2] Building a tight binding hamiltonian yourself, by hand, as in Harrison's sections 3–C and 19–C is certainly the surest way to learn and understand the method. The rewards are very great, as I shall attempt to persuade you now. More recent books are the ones by Sutton,[3] Pettifor[4] and Finnis.[5] In my development here I will most closely follow Finnis. This is because whereas in the earlier literature tight binding was regarded as a simple empirical scheme for the construction of hamiltonians by placing "atomic-like orbitals" at atomic sites and allowing electrons to hop between these through the mediation of "hopping integrals," it was later realised that the tight binding approximation may be directly deduced as a rigorous approximation to the density functional theory. This latter discovery has come about largely through the work of Sutton *et al.*[6] and Foulkes;[7] and it is this approach that is adopted in Finnis' book from the outset.

In the context of atomistic simulation, it can be helpful to distinguish schemes for the calculation of interatomic forces as "quantum mechanical," and "non quantum mechanical." In the former falls clearly the local density approximation (LDA) to density functional theory and nowadays it is indeed possible to make molecular dynamics calculations for small numbers of atoms and a few picoseconds of time using the LDA. At the other end of the scale, classical potentials may be used to simulate millions of atoms for some nanoseconds or more. I like to argue that tight binding is the simplest scheme that is genuinely quantum mechanical. Although you will read claims that the "embedded atom method" and other schemes are LDA-based, tight binding differs from these in that an explicit calculation of the electron *kinetic energy* is attempted either by diagonalising a hamiltonian, which is the subject of this lecture; or by finding its Green function matrix elements which is the subject of the lecture by Ralf Drautz.[8] The enormous advantage of the latter is that calculations scale in the computer linearly with the number of atoms, while diagonalisa-

tion is $\mathcal{O}(N^3)$. At all events, tight binding is really the cheapest and simplest model that can capture the subtleties in bonding that are consequences of the quantum mechanical nature of the chemical bond. Some well-known examples of these quantum mechanical features are magnetism, negative Cauchy pressures, charge transfer and ionic bonding; and of course bond breaking itself which is not allowed by simple molecular mechanics models. At the same time tight binding will reveal detailed insight into the nature of the bonds and origin of interatomic forces in the system you are studying.

1.1 The two centre approximation

In density functional calculations, the hamiltonian is constructed after making a choice of functions used to represent the wavefunctions, charge density and potential. If these are atom centred, for example gaussians, "fire balls" or Slater type orbitals rather than plane waves, then matrix elements of the hamiltonian may become spatial integrals of three such functions. An explicit formula taken from the LMTO method is displayed in equation (26) in section 3.2 below. This can be the most time consuming part of a bandstructure calculation, compared to the subsequent diagonalisation. In the tight binding approximation, we side step this procedure and construct the hamiltonian from a parameterised look up table. But the underlying theory has the same structure. Each hamiltonian matrix element is conceived as a integral of three functions, one potential and two orbitals centred at three sites. (We have made the *Ansatz* that the effective potential may be written as a sum of atom centred potentials.) If all are on the same site, this is a one centre, or *on-site* matrix element; if the orbitals are on different sites and are "neighbours" while the potential is on one of these sites we have a two centre matrix element, or "hopping integral." All other possibilities, namely three centre terms and overlap of orbitals on distant sites are neglected. This forms a central tenet of the tight binding approximation—the nearest neighbour, two centre approximation. The canonical band theory[9] allows us to isolate these terms explicitly and to predict under what circumstances these are indeed small (see section 3.2). The two centre approximation is more than just a convenient rejection of certain terms; it is implicit in the Slater–Koster table and in the calculation of interatomic force that the hamiltonian can be written in parameterised two centre form. This allows one to express the dependence of hopping integrals upon distance analytically. It is a feature of the quantum mechanical method that whereas the hamiltonian comprises short ranged two centre quantities only, the solution of the Schrödinger equation using this simple hamiltonian results in a density matrix that is possibly long ranged and includes many-atom interactions. Indeed the bond order potential exposes this many-atom expansion of the total energy explicitly.[8]

1.2 $\mathcal{O}(N^3)$ and $\mathcal{O}(N)$ implementations

The obvious way to tackle the tight binding electronic structure problem is the same as in density functional theory, namely by direct diagonalisation of the hamiltonian to obtain eigenvalues and eigenfunctions in the tight binding representation, section 2.1 below. This scales in the computer as the third power of the number of orbitals in the molecule or in the unit cell. In the solid state case one employs the Bloch theorem.[10] This means that one retains only the number of atoms in the primitive unit cell (rather than an infinite number) at the expense of having to diagonalise the hamiltonian at an infinite number of $\mathbf{k}$-points.

Luckily there is a well known and sophisticated number of ways to reduce this to a small number of points within the irreducible Brillouin zone.[11,12] The Bloch transform of a real space matrix $H_{\mathbf{R}L\,\mathbf{R}'L'}$ (in the notation described at equation (3) below) is

$$H_{\mathbf{R}L\,\mathbf{R}'L'}(\mathbf{k}) = \sum_{\mathbf{T}} H_{(\mathbf{R}+\mathbf{T})L\,\mathbf{R}'L'}\,\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{T}},$$

where $\mathbf{R}$ and $\mathbf{R}'$ run only over atoms in the primitive unit cell, while $\mathbf{T}$ are all the translation vectors of the lattice. As long as the matrix $H_{(\mathbf{R}+\mathbf{T})L\,\mathbf{R}'L'}$ is short ranged this can be done easily; for long ranged matrices such as the bare structure constants of (30) below, this must be done using the Ewald method. If you like you can *define* a two centre matrix as one for which the Bloch transformation can be reversed (using all $\mathcal{N}_{\mathbf{k}}$ points in the whole Brillouin zone)

$$H_{(\mathbf{R}+\mathbf{T})L\,\mathbf{R}'L'} = \frac{1}{\mathcal{N}_{\mathbf{k}}}\sum_{\mathbf{k}} H_{\mathbf{R}L\,\mathbf{R}'L'}(\mathbf{k})\,\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{T}}.$$

Indeed this is a way to extract a two centre tight binding hamiltonian from an LDA band-structure calculation;[13] an alternative approach is described in section 3.2 below.

In this lecture, I will concentrate solely on the method of direct diagonalisation, but an alternative and potentially much more powerful approach is to abandon $\mathbf{k}$-space, even for a periodic solid, and employ the recursion method to calculate not the eigenvalues and eigenfunctions of the hamiltonian H, but its greenian or Green function; formally for a complex variable z

$$\hat{G}(z) = (z - H)^{-1}.$$

Throwing away $\mathbf{k}$-space will lead to a huge computational benefit, namely that the calculation scales *linearly* with the number of orbitals, but there is a heavy price to pay—interatomic forces converge more slowly than the energy since they require off-diagonal greenian matrix elements and the sum rule derived in equation (16) below is not automatically guaranteed.[14,15] This can play havoc with a molecular dynamics simulation. The problem has been solved by the *bond order potential* which leads to a *convergent* expansion of the tight binding total energy in one-atom, two-atom, three-atom… terms—a many-atom expansion. This is the subject of the lecture by Ralf Drautz in this workshop.[8]

2 Traditional Non Self Consistent Tight Binding Theory

2.1 Density operator and density matrix

The traditional non self consistent tight binding theory, as described, say, by Harrison,[2] is explained here by following Horsfield *et al.*[16,17] We use H^0 to denote the hamiltonian to indicate that this is the non self consistent approximation to density functional theory as it appears in the Harris–Foulkes functional[5]—the first two lines in equation (37) below. (We follow the usual practice of suppressing the "hat" on the hamiltonian operator.) Hence, H^0 is the sum of non interacting kinetic energy and the effective potential generated by some *input*, superposition of atom centred, spherical charge densities.[5] The hamiltonian possesses a complete set of orthogonal eigenfunctions by virtue of the time independent Schrödinger equation,

$$H^0\psi_n = \varepsilon_n\psi_n,$$

which we will write using Dirac's bra-ket notation as

$$H^0 \left|n\right\rangle = \varepsilon_n \left|n\right\rangle . \tag{1}$$

ε_n are the eigenvalues of the hamiltonian and these are used to construct the *band energy*, $E_{\rm band}$, thus

$$E_{\rm band} = \sum_n f_n \, \varepsilon_n . \tag{2}$$

Here, f_n are *occupation numbers*. In an insulator or molecule assuming spin degeneracy these are either zero or two depending on whether ε_n is greater than or less than the Fermi energy. In a metal or molecule having a degenerate highest occupied level these are set equal to twice the Fermi function or some other smooth function having a similar shape.[12] As with any electronic structure scheme, if this is implemented as a *bandstructure* program and hence the hamiltonian is Bloch-transformed into $\mathbf{k}$-space, then the eigenstates are labelled by their band index and wave vector so that in what follows, the index n is to be replaced by a composite index $n\mathbf{k}$. (At the same time matrices become complex and you may assume that what follows until the end of this subsection applies separately at each $\mathbf{k}$-point.)

Central to the tight binding approximation is the expansion of the eigenstates of H^0 in a *linear combination of atomic(-like) orbitals* (LCAO). This means that we decorate each atomic site, which we denote $\mathbf{R}$ to label its position vector with respect to some origin, with orbitals having angular momentum $L = \ell m$. In this way, ℓ labels the orbitals as s, p or d character, while the L label runs as s, x, y, z, xy and so on. These orbitals may be written in bra-ket notation as

$$\left|\mathbf{R}L\right\rangle = \left|i\right\rangle \tag{3}$$

so that we can abbreviate the orbital site and quantum numbers into a single index i or j, k, l. In this way we have

$$\left|n\right\rangle = \sum_i c_i^n \left|i\right\rangle = c_i^n \left|i\right\rangle \tag{4}$$

and we use the famous Einstein summation convention, for brevity, whereby a summation over the indices i, j, k, l is understood if they appear repeated in a product. (Conversely we use n and m to label eigenstates of H^0 in equation (1) and these are not summed implicitly.) The expansion coefficients c_i^n are the eigenvectors of H^0 in the LCAO representation. The parameters of the tight binding model are the matrix elements of the hamiltonian in the LCAO basis which we write

$$H^0_{ij} = \left\langle i\right| H^0 \left|j\right\rangle .$$

We may *assume* that our chosen orbitals are orthogonal to each other, but to be more general there will a matrix of overlap integrals that may also comprise a part of our tight binding model. These are

$$S_{ij} = \left\langle i|j\right\rangle .$$

It then follows from (4) that (summing over j, remember)

$$\left\langle i|n\right\rangle = S_{ij} c_j^n \qquad \text{and} \qquad \left\langle n|i\right\rangle = \bar{c}_j^n S_{ji} \tag{5}$$

in which a "bar" indicates a complex conjugate. The Schrödinger equation (1) becomes a *linear eigenproblem*,

$$\left(H^0_{ij} - \varepsilon_n S_{ij}\right) c^n_i = 0. \tag{6}$$

In the case of an *orthogonal* tight binding model, we have $S_{ij} = \delta_{ij}$, otherwise we need to solve a generalised eigenproblem which is done by a Löwdin transformation. Denoting H^0_{ij} and S_{ij} in bold by matrices, we insert $\mathbf{S}^{-\frac{1}{2}}\mathbf{S}^{\frac{1}{2}}$ after the right parenthesis in (6) and multiply left and right by $\mathbf{S}^{-\frac{1}{2}}$:

$$0 = \left(\mathbf{S}^{-\frac{1}{2}}\mathbf{H}^0\mathbf{S}^{-\frac{1}{2}} - \varepsilon_n \mathbf{1}\right)\left(\mathbf{S}^{\frac{1}{2}}\mathbf{c}\mathbf{S}^{-\frac{1}{2}}\right) = \left(\tilde{\mathbf{H}} - \varepsilon_n \mathbf{1}\right)\mathbf{z},$$

which can be solved as an orthogonal eigenproblem, and recover $\mathbf{c}$ from $\mathbf{z}$ by back-substitution using the previously obtained Cholesky decomposition of $\mathbf{S}$. Now we have our eigenvectors c^n_i from which we construct a density matrix, which is central to the electronic structure problem. The density matrix provides us with the band energy, local "Mulliken" charges, bond charges (in the non orthogonal case)[5], bond orders,[4] interatomic forces, and in the case of time dependent tight binding the bond currents via its imaginary part.[18] The density operator $\hat{\rho}$ needs to have the following properties.

Property 1. Idempotency, meaning $\hat{\rho}^2 = \hat{\rho}$,

Property 2. Tr $\hat{\rho} = N$, the number of electrons,

Property 3. Tr $\hat{\rho}H^0 = \sum_n f_n\, \varepsilon_n = E_{\mathrm{band}}$, the band energy,

Property 4. Tr $\hat{\rho}\frac{\partial}{\partial\lambda}H^0 = \frac{\partial}{\partial\lambda}E_{\mathrm{band}}$, the Hellmann-Feynman theorem.

We know from quantum mechanics[19,20] that the one particle density operator is *defined* as

$$\hat{\rho} = \sum_n f_n\ |n\rangle\langle n|\,.$$

To find its representation in the LCAO basis, we first define a unit operator,

$$\hat{1} = |i\rangle\, S^{-1}_{ij}\, \langle j|\,. \tag{7}$$

To show that it *is* the unit operator, write

$$\begin{aligned}\langle n|n\rangle = 1 &= \langle n|i\rangle\, S^{-1}_{ij}\, \langle j|n\rangle \\ &= \bar{c}^n_k S_{ki} S^{-1}_{ij} S_{jl} c^n_l \\ &= \bar{c}^n_i S_{ij} c^n_j\end{aligned}$$

(after using (5) and swapping indices) which is consistent with (4). More generally we have

$$\langle n|m\rangle = \delta_{nm} = \bar{c}^n_i S_{ij} c^m_j. \tag{8}$$

Now using our unit vector, we write the density operator in our, possibly non orthogonal, LCAO basis,

$$\begin{aligned}\hat{\rho} = \sum_n f_n\ |n\rangle\langle n| &= \sum_n f_n\ \hat{1}\,|n\rangle\langle n|\,\hat{1} \\ &= \sum_n f_n\ |i\rangle\, c^n_i \bar{c}^n_j\, \langle j|\,.\end{aligned} \tag{9}$$

A matrix element of the density operator is

$$\begin{aligned}\rho_{kl} &= \sum_n f_n \ \langle k|i\rangle \, c_i^n \bar{c}_j^n \, \langle j|l\rangle \\ &= \sum_n f_n \, S_{ki} c_i^n \bar{c}_j^n S_{jl} \end{aligned} \tag{10}$$

and in an orthogonal basis this reduces to the familiar density matrix

$$\rho_{ij} = \sum_n f_n \, c_i^n \bar{c}_j^n .$$

If you are familiar with general relativity or non cubic crystallography then you may wish to view the matrix S_{ij} as the metric tensor that "raises" and "lowers" indices of covariant and contravariant vectors.[6,15,21] Finnis[5] makes this point by distinguishing between "expansion coefficients" and "matrix elements" of the density operator. In this way the expansion coefficients of the density operator in the LCAO basis are $\sum_n f_n \, c_i^n \bar{c}_j^n$, while to obtain density matrix elements their indices are "raised" by elements of the metric tensor as in (10); in the orthogonal case ($S_{ij} = \delta_{ij}$) this distinction vanishes.

Now we can demonstrate that $\hat{\rho}$ has the properties 1–4 above. The following is really included here for completeness as the student may not find it elsewhere in the literature. However, on a first reading you may skip to section 2.3 after looking at equations (11), (12), (13), (16) and (17).

Property 1. Idempotency follows immediately from (9).

Property 2. Tr $\hat{\rho} = N$. We must take the trace in the eigenstate basis, hence

$$\begin{aligned}\text{Tr}\,\hat{\rho} &= \sum_m \sum_n f_n \ \langle m|i\rangle \, c_i^n \bar{c}_j^n \, \langle j|m\rangle \\ &= \sum_m \sum_n f_n \, \bar{c}_k^m S_{ki} c_i^n \bar{c}_j^n S_{jl} c_l^m \\ &= \sum_m \sum_n f_n \, \delta_{mn}\delta_{nm} = \sum_n f_n \ = N. \end{aligned}$$

After the second line we have used (8). We can make partial, "Mulliken" charges q_i which amount to the occupancy of orbital i,

$$N = \sum_i q_i = \sum_n f_n \, \bar{c}_i^n S_{ij} c_j^n ,$$

using (8). Because of its importance in tight binding, we will write the Mulliken charge associated with orbital i explicitly,

$$q_i = \sum_n f_n \ \sum_j \bar{c}_i^n S_{ij} c_j^n \tag{11}$$

in which the sum over i implied by the summation convention is, in this instance, suppressed. This is a *weighted decomposition of the norm.* Note that in this and the following you can easily extract the simpler expressions for the more usual orthogonal

tight binding by replacing S_{ij} with the Kronecker δ_{ij} in the implicit sums, in which case

$$q_i = \sum_n f_n \ |c_i^n|^2 \, .$$

It is worthwhile to note that in an orthogonal tight binding model the total charge can be decomposed into individual atom centred contributions; on the other hand non orthogonality introduces *bond charge*[4] so that as seen in (11) there is a summation over both atom centred and bond charges. You may prefer the latter picture: we all know that in a density functional picture the covalent bond arises from the accumulation of charge in between the atoms; in an orthogonal tight binding model one might ask how is this accumulation described? The answer is that it is captured in the *bond order*.[4,8]

Property 3. $\mathrm{Tr}\ \hat{\rho}H^0 = \sum_n f_n \, \varepsilon_n = E_{\mathrm{band}}$.

$$\begin{aligned}
\mathrm{Tr}\ \hat{\rho}H^0 &= \sum_m \sum_n f_n \ \langle m|i\rangle \, c_i^n \bar{c}_j^n \, \langle j| \, H^0 \, |m\rangle \\
&= \sum_m \sum_n f_n \, \bar{c}_k^m S_{ki} c_i^n \bar{c}_n^j S_{jl} c_l^m \varepsilon_m \\
&= \sum_m \sum_n f_n \, \delta_{mn} \delta_{nm} \varepsilon_m = \sum_n f_n \, \varepsilon_n = E_{\mathrm{band}}
\end{aligned}$$

using (1). One may wish to construct partial band energies, E_i, in an equivalent way as

$$E_{\mathrm{band}} = \sum_i E_i = \sum_n f_n \, \bar{c}_i^n H_{ij} c_j^n .$$

The corresponding decomposition of the *bond energy* (18) in section 2.3 is the starting point of the many-atom expansion in the bond order potential.[8]

Property 4. The Hellmann–Feynman theorem tells us that

$$\frac{\partial}{\partial \lambda} \left(\mathrm{Tr}\ \hat{\rho}H^0 \right) = \mathrm{Tr}\ \hat{\rho} \frac{\partial}{\partial \lambda} H^0$$

because solution of the eigenproblem (6), through the Rayleigh–Ritz procedure leads us to a density matrix that is variational with respect to any parameter λ which may be, for example, a component of the position vector of an atom $\mathbf{R}$. Hence to calculate the interatomic force we need to find

$$\begin{aligned}
\mathrm{Tr}\ \hat{\rho} \frac{\partial}{\partial \lambda} H^0 &= \sum_m \sum_n f_n \ \langle m|i\rangle \, c_i^n \bar{c}_j^n \, \langle j| \, \frac{\partial}{\partial \lambda} H^0 \, |m\rangle \\
&= \sum_n f_n \, \bar{c}_i^n c_j^n \, \langle i| \, \frac{\partial}{\partial \lambda} H^0 \, |j\rangle \, .
\end{aligned}$$

Now our tight binding model furnishes us with hopping integrals, H_{ij}^0, and by employing a suitable scaling law, for example equation (23) below, the two centre approximation and the Slater–Koster table we will know how these depend on bond

lengths and angles; so while we don't actually know $\langle i|\,\frac{\partial}{\partial\lambda}H^0\,|j\rangle$, the derivatives that we *do* know are

$$\frac{\partial}{\partial\lambda}H^0_{ij} = \frac{\partial}{\partial\lambda}\langle i|\,H^0\,|j\rangle = \langle\frac{\partial}{\partial\lambda}i|H^0\,|j\rangle + \langle i|\,H^0|\frac{\partial}{\partial\lambda}j\rangle + \langle i|\,\frac{\partial}{\partial\lambda}H^0\,|j\rangle\,.$$

So

$$\mathrm{Tr}\,\hat{\rho}\frac{\partial}{\partial\lambda}H^0 = \sum_n f_n\,\bar{c}^n_i c^n_j\left[\frac{\partial}{\partial\lambda}H^0_{ij} - \langle\frac{\partial}{\partial\lambda}i|H^0\,|j\rangle - \langle i|\,H^0|\frac{\partial}{\partial\lambda}j\rangle\right].$$

Now, to deal with the unknown last two terms, using (4)

$$\begin{aligned}\sum_n f_n\,\bar{c}^n_i c^n_j\left[\langle\frac{\partial}{\partial\lambda}i|H^0\,|j\rangle + \langle i|\,H^0|\frac{\partial}{\partial\lambda}j\rangle\right] &= \sum_n f_n\,\left[\bar{c}^n_i\langle\frac{\partial}{\partial\lambda}i|n\rangle\varepsilon_n + \varepsilon_n\langle n|\frac{\partial}{\partial\lambda}j\rangle c^n_j\right]\\ &= \sum_n f_n\,\varepsilon_n\left[\bar{c}^n_i c^n_j\langle\frac{\partial}{\partial\lambda}i|j\rangle + \bar{c}^n_i c^n_j\langle i|\frac{\partial}{\partial\lambda}j\rangle\right]\\ &= \sum_n f_n\,\varepsilon_n\bar{c}^n_i c^n_j\frac{\partial}{\partial\lambda}S_{ij}\end{aligned}$$

since

$$\frac{\partial}{\partial\lambda}S_{ij} = \frac{\partial}{\partial\lambda}\langle i|j\rangle = \langle\frac{\partial}{\partial\lambda}i|j\rangle + \langle i|\frac{\partial}{\partial\lambda}j\rangle.$$

Finally we arrive at

$$\mathrm{Tr}\,\hat{\rho}\frac{\partial}{\partial\lambda}H^0 = \sum_n f_n\,\bar{c}^n_i c^n_j\left[\frac{\partial}{\partial\lambda}H^0_{ij} - \varepsilon_n\frac{\partial}{\partial\lambda}S_{ij}\right]. \tag{12}$$

2.2 Density of states and bond order

The *density of states* is central to electronic structure theory and is defined to be[22]

$$n(\varepsilon) = \sum_n \delta(\varepsilon - \varepsilon_n). \tag{13}$$

We can define a partial or *local* density of states, $n_i(\varepsilon)$, which is the density of states projected onto the orbital i. We write

$$\begin{aligned}n(\varepsilon) &= \sum_n \langle n|\,\delta(\varepsilon - H^0)\,|n\rangle\\ &= \sum_n\sum_m \langle n|i\rangle\,S^{-1}_{ij}\,\langle j|m\rangle\,\langle m|\,\delta(\varepsilon - H^0)\,|n\rangle\\ &= \sum_n\sum_m \bar{c}^n_j S_{ji}S^{-1}_{ij}S_{jk}c^m_k\,\langle m|\,\delta(\varepsilon - H^0)\,|n\rangle\\ &= \sum_n \bar{c}^n_j S_{jk}c^n_k\,\langle n|\,\delta(\varepsilon - H^0)\,|n\rangle\,.\end{aligned}$$

The first line follows from the Schrödinger equation (1) and in the second line we have inserted our unit operator (7) and a further unit operator, $\sum_m |m\rangle\,\langle m|$. The fourth line follows because of the orthogonality of the eigenvectors, $|n\rangle$ which means we have

$\langle m|\,\delta(\varepsilon - H^0)\,|n\rangle = \langle n|\,\delta(\varepsilon - H^0)\,|n\rangle\,\delta_{mn}$. Remember that in the fourth line j and k are dummy orbital indices to be summed over. We can replace these with i and j for neatness and this leads to

$$n(\varepsilon) = \sum_n \bar{c}_i^n S_{ij} c_j^n \delta(\varepsilon - \varepsilon_n) = \sum_i n_i(\varepsilon). \tag{14}$$

Writing the summation over j explicitly we see that the local density of states is

$$n_i(\varepsilon) = \sum_n \sum_j \bar{c}_i^n S_{ij} c_j^n \delta(\varepsilon - \varepsilon_n), \tag{15}$$

with no summation over i, and that this is a *weighted* density of states;[17] the weight in an orthogonal basis is simply $|c_i^n|^2$—compare this with the Mulliken decomposition (11).

An example is shown in figure 1. This is a self consistent *magnetic* tight binding calculation of the electronic structure of a Cr impurity in Fe, modelled as a dilute, ordered $Fe_{15}Cr$ alloy.[23] Very briefly magnetic tight binding is achieved by including a spin index, $|i\rangle = |\mathbf{R}L\sigma\rangle$, (now the occupation numbers vary between zero and *one*, not two) and adding an exchange potential to the self consistent hamiltonian to allow these to split. In addition to the Hubbard-U (see section 4) one includes a "Stoner I" parameter. We cannot go into details here, but it's gratifying that the simple tight binding model *quantitatively* reproduces the LSDA result, even to the extent of predicting the "virtual bound state" on the Cr impurity.[24,25]

The density of states can be used to find the band energy, since by the properties of the Dirac delta function,

$$\sum_n f_n \int \delta(\varepsilon - \varepsilon_n)\,\varepsilon\,\mathrm{d}\varepsilon = \sum_n f_n\,\varepsilon_n = E_{\mathrm{band}}.$$

If we allow the occupation numbers to be represented by the spin degenerate Fermi–Dirac distribution, $2f(\varepsilon)$, then we find, using (13) and our property 3, above,

$$E_{\mathrm{band}} = 2\int f(\varepsilon)\,\varepsilon\,n(\varepsilon)\,\mathrm{d}\varepsilon = \mathrm{Tr}\,\hat{\rho}H^0 \tag{16}$$

which is an important identity in tight binding theory and one which bears heavily on the convergence of the many atom expansion in the bond order potential.[26]

Finally in this section we should mention that the *bond order* which is central to the bond order potential[8] is obtained directly from the density matrix elements. We define

$$\Theta_{ij} = \frac{1}{2}\left(\rho_{ij} + \rho_{ji}\right)$$

as the *partial order of the bond* as contributed by orbitals i and j, it being understood that these are on different atomic sites. The bond order between sites $\mathbf{R}$ and $\mathbf{R}'$ is obtained by summing the partial bond order over all the orbitals on each atom in question,

$$\Theta_{\mathbf{R}\mathbf{R}'} = \sum_{LL'} \Theta_{\mathbf{R}L\,\mathbf{R}'L'}. \tag{17}$$

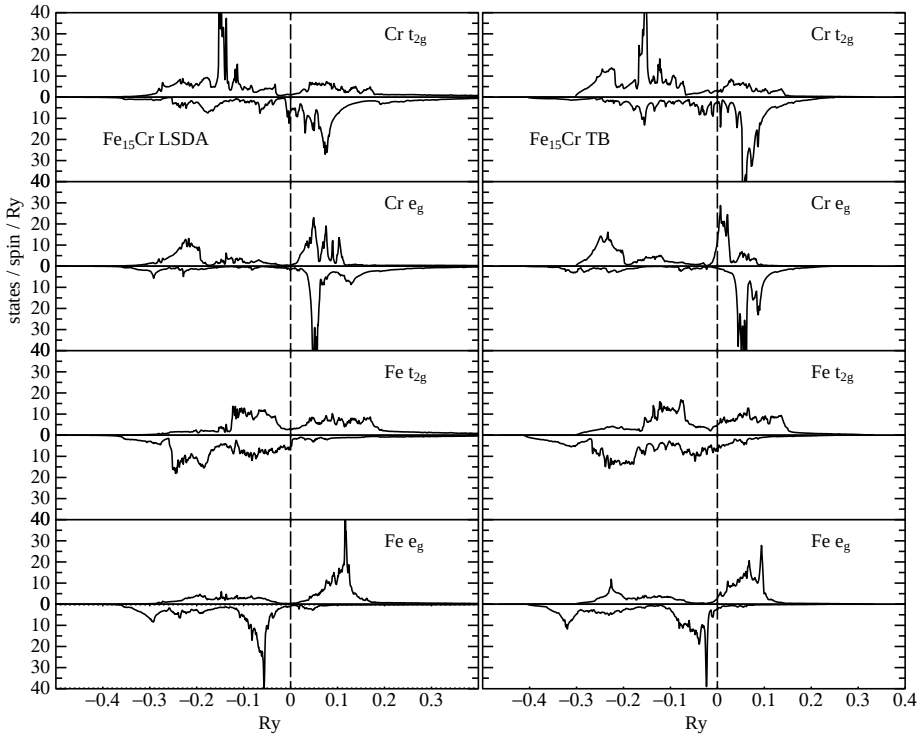

Figure 1. Example of a local density of states.[23] This is an ordered alloy, $Fe_{15}Cr$ on a body centred cubic (bcc) lattice. On the left is the local spin density functional result and on the right a simple, non orthogonal magnetic tight binding approximation. As is conventional, the spin up and down densities are shown as upright and upside down functions respectively. The Fe atom shown is the one closest to the Cr impurity and the density is projected onto the *d*-manifolds. Apart from the accurate description provided by the tight binding model, the most striking feature is the virtual bound state[24,25] seen as sharp peak in the local Cr density of states. It's notable that the *occupied*, spin up state has t_{2g} symmetry while its unoccupied partner belongs largely to the e_g manifold.

2.3 The tight binding bond model

Just as in density functional theory, the sum of occupied eigenvalues of the one electron hamiltonian is not the total energy. In the traditional tight binding approximation, beginning probably with the papers of Jim Chadi,[27] one writes simply

$$E_{\rm tot} = E_{\rm band} + E_{\rm pair}$$

for the total energy in the *band model* and $E_{\rm pair}$ is a pairwise repulsive energy whose functional form and parameters constitute ingredients of the tight binding model; it is intended to represent the double counting and ion–ion contributions to the density functional total energy.[27] "Double counting" is a term given to the electron–electron interaction energy in density functional theory. Because the theory is cast into a one electron form through the Kohn–Sham equations, the band energy, by summing over the eigenvalues, counts the electron–electron interaction twice. The interaction between, say, electrons in occupied states 1 and 2 is counted first when eigenvalue 1 is added in and again when eigenvalue 2 is

added. One cannot simply divide by two because $E_{\rm band}$ also contains kinetic and electron–ion energies which are not double counted. Hence one recalculates the electron–electron interaction energy and subtracts it, calling this the "double counting" correction.

Pursuing an argument that goes back as far as the sixties,[28,29] Pettifor[30] formulates the total energy in terms of the *bond* energy, $E_{\rm bond}$, rather than the band energy. The tight binding bond model[6] (TBBM) is the starting point for both self consistent tight binding which is described below in section 4 and for the modern bond order potentials.[8] Therefore we will pursue only the bond model further here. The essential point is that one arrives at the *covalent bond energy*[3,6] by removing the diagonal elements of $E_{\rm band} = {\rm Tr}\,\hat{\rho}H^0$. We recall that orbital indices i and j are a composite of site labels and quantum numbers, and write

$$E_{\rm bond} = \frac{1}{2}\sum_{\substack{ij\\ \mathbf{R}'\neq\mathbf{R}}} 2\rho_{ij}H^0_{ji} = \frac{1}{2}\sum_{\substack{\mathbf{R}L\,\mathbf{R}'L'\\ \mathbf{R}'\neq\mathbf{R}}} 2\,\rho_{\mathbf{R}L\,\mathbf{R}'L'}\,H^0_{\mathbf{R}'L'\,\mathbf{R}L}. \tag{18}$$

Here all terms are excluded from the double sum if orbitals i and j are on the same site $\mathbf{R}$. Note how by dividing and multiplying by two we can expose this as a sum of bond energies which is then divided by two to prevent each bond being double counted in the same way as a pair potential is usually written.

In the TBBM, the remaining diagonal terms in $E_{\rm band}$ are grouped with the corresponding quantities in the free atom. In the non self consistent tight binding approximation, the on-site matrix elements of H^0 are simply the free atom orbital energies (eigenvalues of the atomic hamiltonian)

$$H^0_{\mathbf{R}L\,\mathbf{R}L'} = \varepsilon_{\mathbf{R}\ell}\,\delta_{LL'}$$

and in addition to the hopping integrals, these are parameters of the tight binding model, ε_s, ε_p and ε_d. Furthermore, we assume certain orbital occupancies in the free atom, say, $N_{\mathbf{R}\ell}$, whereas after diagonalisation of the tight binding hamiltonian one finds these orbitals have occupancy given by the diagonal matrix elements of the density matrix. Hence there is a change in energy in going from the free atom limit to the condensed matter which is

$$\begin{aligned} E_{\rm prom} &= \sum_{\mathbf{R}L}\left(\rho_{\mathbf{R}L\,\mathbf{R}L}\,H^0_{\mathbf{R}L\,\mathbf{R}L} - N_{\mathbf{R}\ell}\,\varepsilon_{\mathbf{R}\ell}\right)\\ &= \sum_{\mathbf{R}L}\left(\rho_{\mathbf{R}L\,\mathbf{R}L} - N_{\mathbf{R}\ell}\right)\varepsilon_{\mathbf{R}\ell}\\ &= \sum_{\mathbf{R}L}\Delta q_{\mathbf{R}L}\,\varepsilon_{\mathbf{R}\ell}. \end{aligned} \tag{19}$$

We have assumed for now that on-site elements of H^0 are strictly diagonal and we recognise the first term in the first line as the difference between $E_{\rm band}$ and $E_{\rm bond}$. $E_{\rm prom}$ is called the *promotion energy* since it is the energy cost in promoting electrons that is very familiar, say, in the s–p promotion in covalent semiconductors in "preparing" the atoms in readiness to form the sp^3 hybrids in the diamond structure or the sp^2 hybrids in graphite. Thus in the tight binding bond model the *binding energy* is written as the total energy take away the energy of the free atoms,

$$E_{\rm B} = E_{\rm bond} + E_{\rm prom} + E_{\rm pair}. \tag{20}$$

The *interatomic force* is minus the gradient of the pairwise $E_{\rm pair}$ which is trivial, minus Tr $\rho\boldsymbol{\nabla}H^0$ which can be computed using equation (12) assuming that *on-site* hamiltonian matrix elements remain constant; this is the fully *non* self consistent tight binding approximation. And in fact at this level of approximation the band and bond models are indistinguishable. The first order variation of $E_{\rm B}$ with respect to atom cartesian coordinate R_α is

$$\frac{\partial}{\partial R_\alpha}E_{\rm B} = \sum_{\substack{\mathbf{R}L\,\mathbf{R}'L' \\ \mathbf{R}'\neq\mathbf{R}}} 2\rho_{\mathbf{R}L\,\mathbf{R}'L'}\,\frac{\partial}{\partial R_\alpha}H^0_{\mathbf{R}'L'\,\mathbf{R}L} + \frac{\partial}{\partial R_\alpha}E_{\rm prom} + \frac{\partial}{\partial R_\alpha}E_{\rm pair}. \qquad (21)$$

This is written for the orthogonal case, since this approximation forms a tenet of the TBBM. However, it's easy enough to add in the term from (12) containing the overlap and of course the diagonal elements S_{ii} are constant and do not contribute to the force. Note that the half in front of (18) has vanished—in the calculation of the force one sums over all bonds emanating from the atom at $\mathbf{R}$, not just half of them!

Now comes a rather subtle point. Unlike the band model, the bond model is properly consistent with the force theorem.[31] This states that there is no contribution to the force from self consistent redistribution of charge as a result of the virtual displacement of an atom. If a self consistent electronic system is perturbed to first order then that change in the bandstructure energy due to electron–electron interaction is exactly cancelled by the change in the double counting. This remarkable result means that by making a first order perturbation one cannot distinguish between an interacting and a non interacting electron system.[32] Indeed to calculate the interatomic force it is sufficient to find the change in band energy while making the perturbation—in this case the virtual displacement of an atom—in the frozen potential of the unperturbed system. In the band model there will be a first order change in the band energy upon moving an atom which *ought* to be cancelled by an appropriate change in the double counting, but *is not* because this is represented by the pair potential. Now we can discuss $\partial E_{\rm prom}/\partial R_\alpha$. In the band model there is no contribution to the force from $E_{\rm prom}$ (19); because of the variational principle $\varepsilon_{\mathbf{R}L}\delta q_{\mathbf{R}L} = 0$, and $q_{\mathbf{R}L}\delta\varepsilon_{\mathbf{R}L} = 0$ because the $\varepsilon_{\mathbf{R}L}$ are constants. However the Mulliken charge transfers are not necessarily zero and the force theorem does require any electrostatic contributions due to charge transfer to be included in the interatomic force;[33,34] neglect of these leads to the inconsistency of the band model. In the TBBM the most limited self consistency is imposed, namely the *Ansatz* of local charge neutrality so that electrostatic charge transfer terms vanish. This requires that for each site the *total* Mulliken charge difference between free atoms and condensed phase summed over all orbitals is zero. This is achieved iteratively by adjusting the on-site orbital energies. Here is the simplest example of a self consistent tight binding theory. It only affects the diagonal, on-site hamiltonian matrix elements and hence only $E_{\rm prom}$ is changed. Suppose we now write the hamiltonian as

$$H = H^0 + H' \qquad (22)$$

where H' has only diagonal elements which we may call $\Delta\varepsilon_{\mathbf{R}L}$. Then

$$\begin{aligned} E^{\rm TBBM}_{\rm prom} &= \sum_{\mathbf{R}L}\left(\rho_{\mathbf{R}L\,\mathbf{R}L} - N_{\mathbf{R}\ell}\right)H_{\mathbf{R}L\,\mathbf{R}L} \\ &= \sum_{\mathbf{R}L}\Delta q_{\mathbf{R}L}\left(\varepsilon_{\mathbf{R}\ell} + \Delta\varepsilon_{\mathbf{R}\ell}\right). \end{aligned}$$

In a sense this isn't really "promotion energy" anymore because we have applied the on-site energy shift to the free atoms also, but it is consistent with the formulation of the TBBM.[6] There will now be a contribution to the force on atom $\mathbf{R}$ from the new term $\sum_L \Delta q_L \Delta\varepsilon_\ell$. If the self consistency is achieved in such a way that all orbital energies are shifted by the same amount at each site, then this contribution vanishes because $\Delta\varepsilon_\ell$ is independent of L, moves to the front of the summation sign and $\sum_L \Delta q_L = 0$ by the local charge neutrality condition. Further and complete discussion of the TBBM can be found in the original paper[6] and in Finnis' book.[5]

3 How to Find Parameters

Now we turn to the question that is probably the most controversial. Many people dislike the tight binding approximation because whereas on the one hand we claim it to be close to the *ab initio* local density approximation solution, on the other we are reduced to finding parameters empirically just as if this were another classical potential. My own view is that *if* the tight binding approximation contains enough of the physics of the system we are studying then any reasonably chosen set of parameters will provide us with a useful model. From this point of view we would also demand that only a very small number of parameters is actually employed in the model. Furthermore it should be possible to choose these by intelligent guesswork and refinement starting from some well established set of rules; for example Harrison's solid state table,[2] or the prescription of Spanjaard and Desjonquères for the transition metals.[35] For example, the latter prescription has furnished us with useful tight binding models[23,36] for Mo, Re, Nb and Fe each with some five to ten adjustable parameters. Alternatively a 53-parameter model for Mo was produced by very careful fitting to a huge database of properties.[37] There doesn't appear to exist a particular advantage of one approach over the other and both types of model have turned out to be predictive of electronic and structural properties of the transition metals.

We need to distinguish between hamiltonian parameters—on-site orbital energies $\varepsilon_{\mathbf{R}\ell}$ and hopping integrals $H^0_{\mathbf{R}L\mathbf{R}'L'}$—and the parameters of the pair potential. Additional complications arise as described later in section 3.3 in the case of *environmentally dependent* parameters.[37]

I wish to illustrate the problem by reference to some examples from the literature.

3.1 Parameters by "adjustment"—example of ZrO_2

The tight binding model for zirconia,[38] ZrO_2, was designed to provide a description of the structural properties of this industrially important ceramic material. ZrO_2 suffers a number of structural phase transitions as a function of temperature. This is exploited in an extraordinary phenomenon called transformation toughening.[39] Its low temperature phase is monoclinic, at intermediate temperatures it is tetragonal and the high temperature modification is cubic. An open question was whether the tetragonal to cubic transition is of first or second order thermodynamically, order–disorder or displacive. Additionally, it is known that the cubic structure is stabilised at low temperature by doping with aliovalent cations (Y, Ca, Mg *etc*) while the mechanism for this was unknown. The tight binding model turned out to be capable of addressing both these issues and the order of the transition was discovered[40] as well as the mechanism of stabilisation of the cubic phase.[41] The

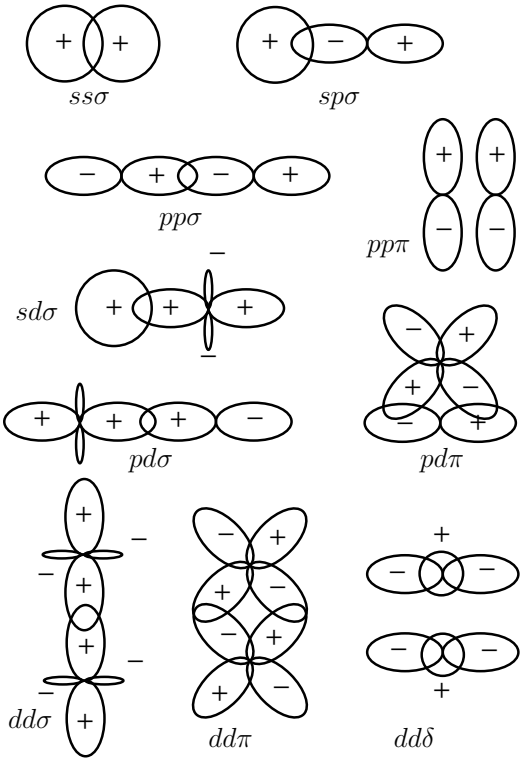

Figure 2. Bond integrals, after Majewski and Vogl.[42] This shows the well known atomic orbitals of the various s, p or d types joined along a bond. Radial symmetry along the bond is assumed leading to the designation of the bond as σ, π or δ. To construct a tight binding hamiltonian requires these fundamental bond integrals assembled through the Slater–Koster table using the direction cosines of the bond in a global cartesian system (these bond integrals are given with respect to a z-axis directed along the bond). This is illustrated in figure 6.5 in ref [3].

strategy of finding tight binding parameters was quite simple. Since the eigenvalues of the hamiltonian describe the energy bands it is sensible to adjust the on-site energies and hopping integrals to the LDA bandstructure, and then find a simple pair potential whose parameters are chosen to obtain, say, the equilibrium lattice constant and bulk modulus. In this case the smallest number of adjustable parameters was chosen to replicate the cubic phase in the hope that the model will then *predict* the ordering in energy of the competing phases. The steps are these.

1. Choose a *minimal* tight binding basis set. In this case d-orbitals were placed on the Zr atoms and s and p on the oxygen. We should mention that being an ionic crystal the TBBM is inadequate and this is in fact a *self consistent* tight binding model using polarisable ions. This is explained later in section 4. The hopping matrix elements are linear combinations of the fundamental bond integrals that are illustrated in figure 2. The particular linear combination depends on the bond angle geometry and is encapsulated in the Slater–Koster table.[1] This is illustrated in figure 6.5 in ref [3]. We only

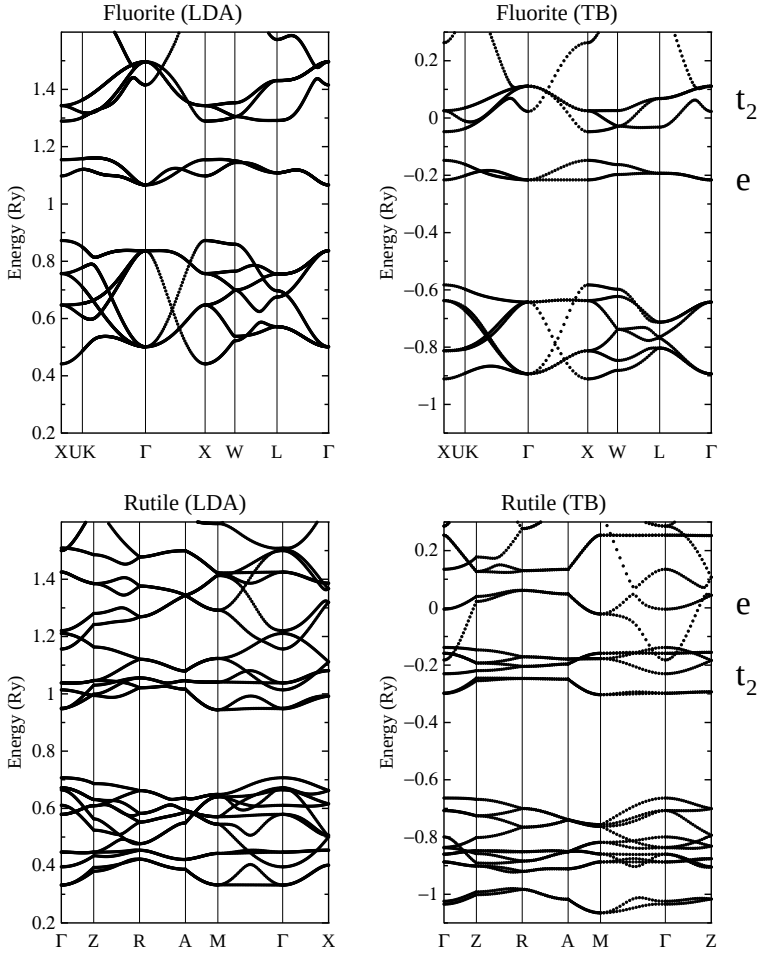

Figure 3. Energy bands of ZrO_2 using both LDA and a tight binding model in both fluorite and rutile crystal modifications. The model parameters were adjusted to the fluorite bands and the rutile bands are therefore a *prediction*. We should note that a number of features such as the splitting in the d-manifold into t_2 and e_g subbands and the crystal field widening of the p-derived ligand band in rutile are consequences of using the self consistent polarisable ion model, and this will be described later in section 4. But we can note in anticipation that it is the new Δ parameters that permit the ordering ($t_2 > e_g$) in the cubic crystal field and *vice versa* in the octahedral field to be reproduced automatically.

need to find the relevant fundamental bond integrals between neighbouring atoms. Zr–O first neighbour bonds require us to know $sd\sigma$, $pd\sigma$ and $pd\pi$ and we choose also to include second neighbour O–O bonds to be made by $ss\sigma$, $sp\sigma$, $pp\sigma$ and $pp\pi$ bond integrals. We have to choose both their value and the way in which they depend on bond length. There is a "canonical band theory," that is really appropriate for metals[9,43,44] but which *faux de mieux* we can apply more generally. This provides us with guidance on how the bond integrals decay with distance and also with certain ratios, namely $pp\sigma$:$pp\pi$ and $dd\sigma$:$dd\pi$:$dd\delta$, see equation (30) below. The required hopping

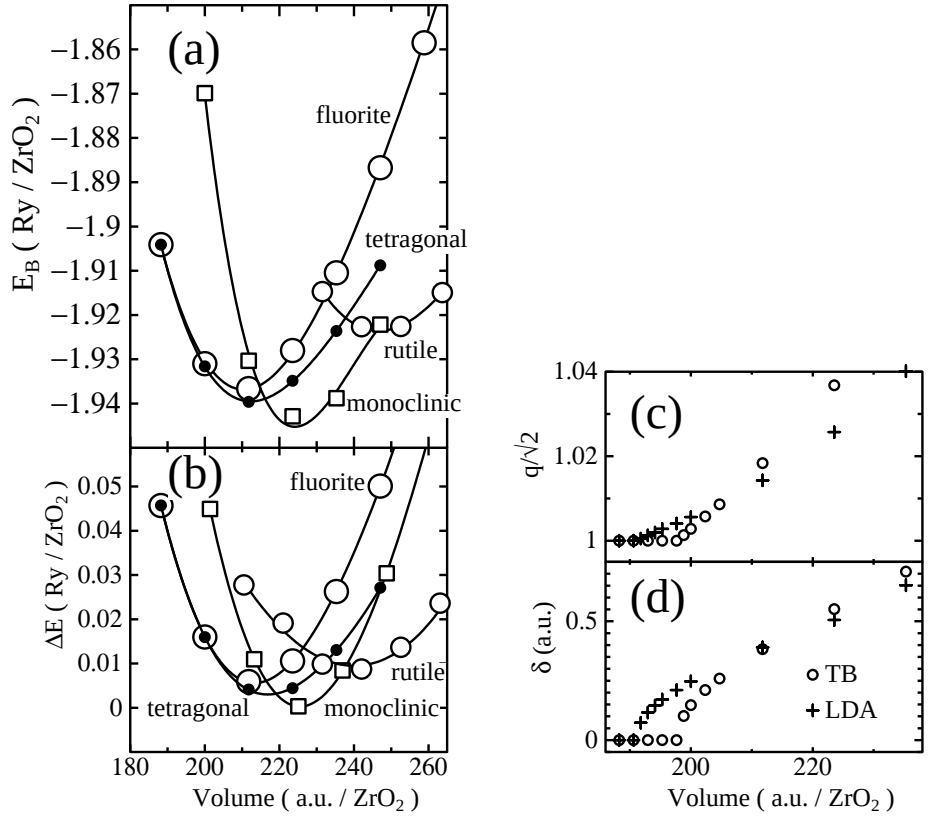

Figure 4. Total energy versus volume in four competing crystal structures of ZrO_2.[38] At each volume, the energy is minimised simultaneously with respect to all the remaining degrees of freedom. (a) LDA calculations of the absolute binding energy (energy with respect to spin polarised free atoms); (b) tight binding results referred to the equilibrium energy of the monoclinic phase. (c) and (d) show the axial ratio q and distortion parameter δ in the tetragonal modification as a function of volume.

integrals are initially taken from Harrison's solid state table and adjusted visually to obtain agreement with the shapes, and especially the widths of the LDA bands. One can also adjust to either the LDA or to experimental band gaps. Also the scaling of the bond integrals can be adjusted to the volume dependence of the LDA bandwidths.[a] The result is shown in figure 3.

We should give more detail of how the bond integrals depend on bond length, r. A very useful function is that of Goodwin, Skinner and Pettifor[46] (GSP)

$$(\ell\ell' m) = V_0 \left(\frac{d}{r}\right)^n \exp\left\{n\left[-\left(\frac{r}{r_c}\right)^{n_c} + \left(\frac{d}{r_c}\right)^{n_c}\right]\right\}. \qquad (23)$$

[a] It is very useful to have a computer program that can calculate energy bands, density of states, total energy using both LDA in some form and in the tight binding approximation, preferably all using the same input file. Luckily such a program exists.[45] Students may contact the author if they wish to learn how to use this.

Most important are the prefactor V_0 which is the value at the equilibrium bond length, d, and the exponent n which determines the slope of the function at equilibrium, since when $r = d$ the argument of the exponential vanishes. The role of n_c and r_c is to give a rapid decay to the function at around $r = r_c$.

2. A pair potential needs to be chosen. The GSP function can be used but in the ZrO_2 model a very simple Born–Mayer form was used between first neighbour Zr–O bonds only. The Born–Mayer function $\varphi(r) = A\,\mathrm{e}^{-br}$ has only two parameters which were fitted to the lattice constant and bulk modulus of cubic ZrO_2.

Figure 4 shows energy volume curves for the competing crystal structures comparing the tight binding model to LDA. Also shown are the order parameters that describe the tetragonal to cubic phase transition as functions of volume.

It is rather clear that the tight binding model for ZrO_2 gives a really excellent set of predictions, having been fitted (or adjusted, rather) only to the cubic structure. In particular the rutile structure is found to be much higher in energy than its competitors—a feature that cannot be reproduced in purely classical models. The vanishing of the order parameters with pressure is well reproduced qualitatively. This and the example shown in figure 1 where simple models, rather insensitive to the choice of parameters, reveal useful and predictive physics gives one confidence the tight binding approximation is indeed a valuable and reliable theory.

3.2 Parameters taken from first principles tight binding—example of Mo

Students who are not particularly interested in the details of an LMTO calculation, may skip this section after looking at figure 5 and subsequent comments. However section 3.3 *is* important. It makes sense to obtain the hamiltonian matrix elements from *ab initio* bandstructures. Probably the most transparent LDA bandstructure theory is the one provided by the linear muffin-tin orbitals (LMTO) method. In the atomic spheres approximation (ASA) the entire bandstructure problem is reduced to knowing four "potential parameters" in each $\mathbf{R}\ell$ site and angular momentum channel. Moreover these parameters have a clear interpretation in terms of the bandstructure. C is the centre of the band; Δ is the bandwidth parameter; γ is a distortion parameter describing the deviation from canonical bands and finally p is a small parameter allowing the eigenvalues to be correct up to third order in their deviation from some chosen energy called ε_ν. An LMTO is a composite orbital-like basis function. A sphere is inscribed about each atom with radius such that the sum of all sphere volumes equals the total volume; in a simple monatomic crystal this is the Wigner–Seitz radius. Within the sphere the radial Schrödinger equation is solved at the energy ε_ν in the current potential and this solution and its energy derivative are matched to solid Hankel and Bessel functions between the spheres. This matching condition is enough to provide the potential parameters which are functions of the logarithmic derivatives of the radial Schrödinger equation solutions $\phi_L(\mathbf{r}) = \phi_\ell(r)\,Y_L(\mathbf{r})$. Each LMTO envelope may be expanded about a given atomic site using the property that a Hankel function at one site may be written as a linear combination of Bessel functions at some other site. This property means that all the Hankel functions in the solid can be expressed as a "one centre" expansion about any one atomic sphere. The expansion coefficients are called "$\kappa = 0$ structure constants" and they transform under rotations according to the Slater–Koster table

and hence may be identified as $\ell\ell' m$ hopping integrals.[47,48] However conventional structure constants are very long ranged. To make contact with tight binding theory Andersen and Jepsen showed that one can make similarity transformations between sets of solid state LMTO's;[49] each basis set being equivalent to another since they give identical bandstructures. In particular Andersen demonstrated that one can defined a "most localised" and an "orthogonal" set of LMTOs. The transformation works like this. In the ASA an LMTO at site $\mathbf{R}$ is made up of a linear combination of a radial solution $\phi(\mathbf{r}-\mathbf{R})$ (the "head") and energy derivative functions $\dot{\phi}(\mathbf{r}-\mathbf{R}')$ ($\mathrm{d}\phi/\mathrm{d}\varepsilon$ evaluated at ε_ν) at all other sites (the "tails"). These are assembled into a one centre expansion using the structure constants. So an LMTO looks like this,

$$\chi_{\mathbf{R}L}(\mathbf{r}-\mathbf{R}) = \phi_{\mathbf{R}L}(\mathbf{r}-\mathbf{R}) + \sum_{\mathbf{R}'L'} \dot{\phi}_{\mathbf{R}'L'}(\mathbf{r}-\mathbf{R}')\, h_{\mathbf{R}'L'\,\mathbf{R}L}.$$

By a choice of normalisation, one can choose the $\dot{\phi}(\mathbf{r}-\mathbf{R}')$ to be those that are *orthogonal* to the radial solutions in each sphere. This particular set of energy derivative functions is given a superscript γ and one is said to be using the "γ-representation." More generally one can vary the normalisation by mixing in some radial solutions with the $\dot{\phi}(\mathbf{r}-\mathbf{R}')$ to make up the tails of the LMTO. To do this we write

$$\dot{\phi}_{\mathbf{R}L}(\mathbf{r}-\mathbf{R}) = \dot{\phi}^{\gamma}_{\mathbf{R}L}(\mathbf{r}-\mathbf{R}) + \phi_{\mathbf{R}L}(\mathbf{r}-\mathbf{R})\, o_{\mathbf{R}L}, \tag{24}$$

so that in the γ-representation, the potential parameter $o_{\mathbf{R}L}$ is zero. It's called o for overlap but has units of energy^{-1}. To construct the overlap matrix in the ASA one has to expand out $\langle\chi|\chi\rangle$; and similarly $\langle\chi\,|-\nabla^2+V_{\text{eff}}|\,\chi\rangle$ for the hamiltonian. If we write that $h_{\mathbf{R}'L'\,\mathbf{R}L}$ is an element of a matrix $\mathbf{h}$ and $o_{\mathbf{R}L}$ and $p_{\mathbf{R}L}$ are elements of diagonal potential parameter matrices, o and p, then Andersen finds for the overlap matrix[48]

$$\mathbf{S} = \mathbf{1} + o\mathbf{h} + \mathbf{h}o + \mathbf{h}p\mathbf{h}. \tag{25}$$

As we mentioned p is a small potential parameter. So in the γ-representation $o=0$ and to second order the overlap is unity and we have an orthogonal basis. The hamiltonian matrix turns out to be[48]

$$\mathbf{H} = \varepsilon_\nu + \mathbf{h} + \mathbf{h}o\varepsilon_\nu + \varepsilon_\nu o\mathbf{h} + \mathbf{h}\,(o+p\varepsilon_\nu)\,\mathbf{h}. \tag{26}$$

Again, in the γ-representation, neglecting third order terms the hamiltonian is just $\mathbf{H} = \varepsilon_\nu + \mathbf{h}$. So if one calculates structure constants and self consistent potential parameters using an LMTO code then one can build an orthogonal tight binding model by explicitly building $\mathbf{H}$ in the γ-representation. By construction, to second order it will reproduce the LDA energy bands.

Unfortunately there is no guarantee that this hamiltonian is short ranged. Andersen made a particular choice of the potential parameter $o_{\mathbf{R}L}$ by defining "screening constants" $\alpha_{\mathbf{R}L}$ in this way: ref [9], eq (91),

$$\frac{1}{o_{\mathbf{R}L}} = C_{\mathbf{R}L} - \varepsilon_{\nu,\mathbf{R}L} - \frac{\Delta_{\mathbf{R}L}}{\gamma_{\mathbf{R}L}-\alpha_{\mathbf{R}L}}. \tag{27}$$

They are called screening constants because the effect of adding radial solutions to the $\dot{\phi}^\gamma$ in (24) is to match the Schrödinger equation solutions in the sphere to Hankel functions $K_L(\mathbf{r}-\mathbf{R})$ that have been screened by additional Hankel functions at surrounding atomic

sites. There is an electrostatic analogy. The solid Hankel function represents the electrostatic potential due to a 2^ℓ multipole. This can be screened by surrounding the sphere with further grounded metal spheres, whose contribution to the potential is then provided by these further Hankel functions at the surrounding spheres. If one chooses the screening constants $\alpha_{\mathbf{R}L}$ to equal the band distortion parameters $\gamma_{\mathbf{R}L}$ then one arrives at the γ-representation since we get $o_{\mathbf{R}L} = 0$ in (27). All other representations are specified by choices of screening constants. The choice $\alpha_{\mathbf{R}L} = 0$ corresponds to the so called "first generation" LMTO which employs the standard $\kappa = 0$ KKR structure constants[b]

$$B_{\mathbf{R}'L'\,\mathbf{R}L} = -8\pi \sum_{L''} (-1)^\ell \frac{(2\ell''-1)!!}{(2\ell-1)!!(2\ell'-1)!!}\, C_{L'LL''}\, K_{L''}(\mathbf{R}-\mathbf{R}') \qquad (28)$$

where

$$K_L(\mathbf{r}) = r^{-\ell-1}\, Y_L(\mathbf{r}),$$

is the solid Hankel function,

$$C_{L''L'L} = \iint \mathrm{d}\Omega\, Y_{L''}\, Y_{L'}\, Y_L \qquad (29)$$

are Gaunt coefficients and Y_L are real spherical harmonics (see Appendix). The whole Slater–Koster table is encapsulated in this formula; the Gaunt coefficients provide selection rules that pick out certain powers of r and angular dependencies. By pointing a bond along the z-axis one can see how the canonical scaling and ratios come about since these structure constants are simply,[48]

$$\begin{aligned}
B_{ss\sigma} &= -2/d \\
B_{sp\sigma} &= 2\sqrt{3}/d^2 \\
B_{pp\{\sigma,\pi\}} &= 6\{2,-1\}/d^3 \\
B_{sd\sigma} &= -2\sqrt{5}/d^3 \\
B_{pd\{\sigma,\pi\}} &= 6\sqrt{5}\{-\sqrt{3},1\}/d^4 \\
B_{dd\{\sigma,\pi,\delta\}} &= 10\{-6,4,-1\}/d^5
\end{aligned} \qquad (30)$$

in which d is a dimensionless bond length r/s, where s is conventionally chosen to be the Wigner–Seitz radius of the lattice. These can be compared with the cartoons in figure 2 in which the overlapping of two positive lobes leads to a negative bond integral and *vice versa*. This is because the orbitals are interacting with an attractive, negative, potential (section 1.1). Note how the factor $(-1)^\ell$ in (28) neatly takes care of the cases like $ps\sigma = -sp\sigma$. You have to be careful of these if you program the Slater–Koster table by hand.[5]

Transformations from the "first generation" to "second generation" LMTO basis sets are quite easily done. Having chosen screening constants one transforms the structure constants thus,[c]

$$\mathbf{B}^\alpha = \mathbf{B} + \mathbf{B}\alpha\mathbf{B}^\alpha \qquad (31)$$

[b] Andersen uses the symbol S for structure constants but we've already used it for the overlap, which is standard tight binding usage. Here we use B for Andersen's which differ by a prefactor $2/[(2\ell-1)!!(2\ell'-1)!!]$ and a minus sign from the KKR structure constants.[50]

[c] Clearly $B_{\mathbf{R}'L'\,\mathbf{R}L}$ has two centre form, section 1.1, as it depends only on the connecting vector $\mathbf{R}-\mathbf{R}'$ (28). It's less obvious that $\mathbf{B}^\alpha$ is a two centre matrix because of the three centre terms introduced by the second term

which is a Dyson equation, and α is a diagonal matrix. Then one transforms the potential parameters by defining a vector (we suppress the $\mathbf{R}L$ subscripts)

$$\xi = 1 + (C - \varepsilon_\nu)\,\frac{\alpha}{\Delta}$$

after which (ref [9], p. 88)

$$c = \varepsilon_\nu + (C - \varepsilon_\nu)\,\xi \;\; ; \qquad d = \xi^2 \Delta$$

where c and d are the band parameters C and Δ in the new representation. The overlap parameter o is transformed according to (27).

Andersen and Jepsen[49] determined empirically a set of screening constants, namely[d]

$$\alpha_s = 0.3485 \qquad \alpha_p = 0.05304 \qquad \alpha_d = 0.010714, \tag{32}$$

which lead to the "most localised" or most tight binding LMTO basis. Now one can construct hamiltonians and overlaps according to (26) and (25) by noting that the *first order* hamiltonian is constructed from potential parameters and structure constants[9,48]

$$h_{\mathbf{R}L\,\mathbf{R}'L'} = (c_{\mathbf{R}L} - \varepsilon_{\nu,\mathbf{R}L})\,\delta_{\mathbf{R}L\,\mathbf{R}'L'} + \sqrt{d_{\mathbf{R}L}}\,B^{\alpha}_{\mathbf{R}L\,\mathbf{R}'L'}\,\sqrt{d_{\mathbf{R}'L'}}.$$

Now we want our tight binding hamiltonian to have two centre form and it is easy to identify which are the three centre terms in the LMTO hamiltonian and overlap matrices—they are contained in the terms bilinear in $\mathbf{h}$, the last terms in (26) and (25). These terms (as do the linear terms) also contain two and one centre terms, of course, arising from the diagonal terms of $\mathbf{h}$. We can dispose of three centre terms in two ways.

1. We can work to *first order*, in which case, in both α- and γ-representations

$$\mathbf{H}^{(1)} = \varepsilon_\nu + \mathbf{h} \tag{33}$$

   and since $o\mathbf{h}$ terms are of second order, both these are orthogonal models with overlap being unity.

2. We can work to second order by retaining $o\mathbf{h}$ terms but neglecting the small potential parameter p^γ in the γ-representation. In this representation ($o = 0$) this is no different from the first order hamiltonian, and the overlap is unity. In the α-representation this introduces some additional two centre contributions to the matrix elements of the hamiltonian and overlap, and we are careful to extract one and two centre contributions from the last term in (26).

All this is illustrated in figure 5 for the bcc transition metal Mo. The screening constants from (32) are used. Here are some noteworthy points.

1. Clearly the two representations deliver different sets of hopping integrals. *You cannot expect density functional theory to furnish you with THE tight binding model.* On the other hand they show a proper decay with increasing bond length. The decay is

in (31). Nonetheless because the transformation is done in real space it is also a two centre matrix by virtue again of its dependence only upon $\mathbf{R} - \mathbf{R}'$. On the other hand it possesses additional "environmental dependence," see section 3.3.

[d] An alternative is to define $\alpha_{\mathbf{R}\ell} = (2\ell+1)(r_{\mathbf{R}\ell}/s)^{2\ell+1}$ by choosing site and ℓ-dependent "hard core radii" $r_{\mathbf{R}\ell}$.[51] This is consistent with "third generation LMTO."[52]

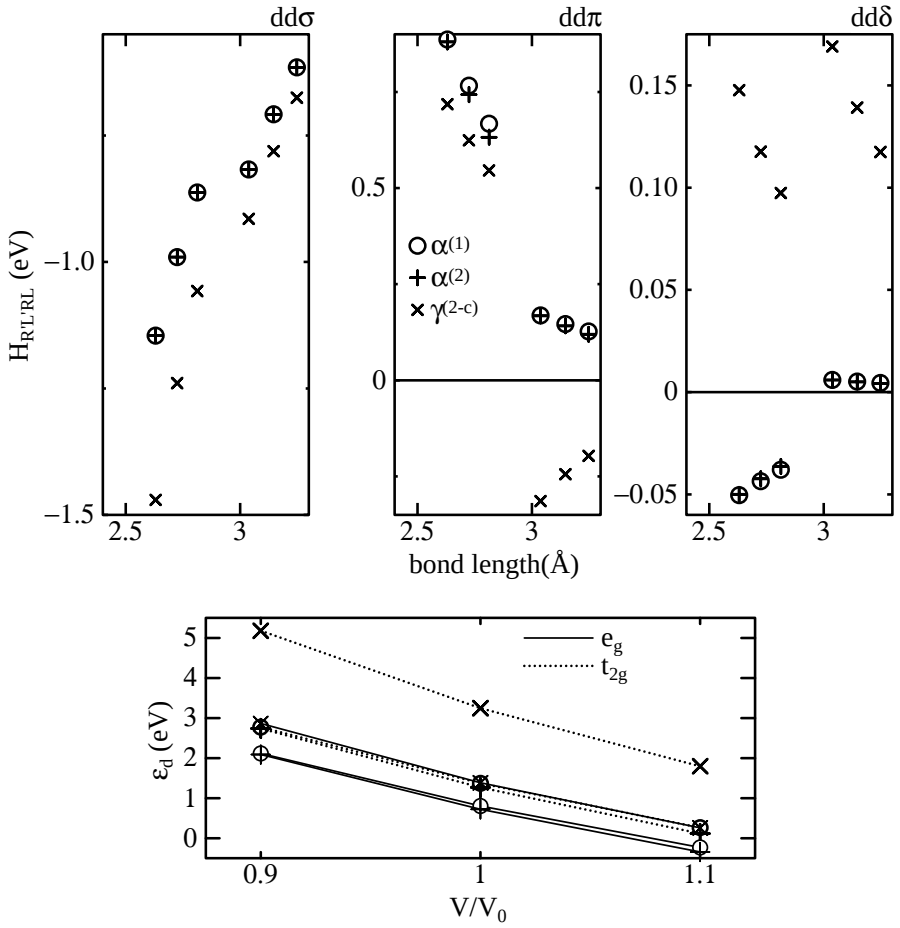

Figure 5. Hopping integrals $\ell\ell' m$ in the body centred cubic transition metal Mo, calculated using LMTO theory. The three integrals $dd\sigma$, $dd\pi$, and $dd\delta$ are found by rotating the z axis to first and then to second neighbour bonds and doing this at three different atomic volumes;[53] hence for each integral six values of $\ell\ell' m$ are shown as a function of bond length. Three model LMTO hamiltonians are used. The crosses refer to the two centre γ-representation; the circles to the *first order* α-representation and the pluses to the *second order*, two centre $\mathbf{H}^{\alpha}$. In the lower panel are shown the diagonal matrix elements and their, rather strong, volume dependence.

more rapid in the tight binding, α-representation as expected, furthermore the first order tight binding representation is strictly orthogonal; not shown in figure 5 are the overlap matrix elements in the second order tight binding representation, but indeed these are very small—no greater than 0.025 in magnitude. Note that the tight binding bond integrals respect the signs and roughly the canonical ratios of the bare structure constants (30) while in the γ-representation $dd\delta$ and the second neighbour $dd\pi$ have the "wrong" signs. Furthermore we would find that while the tight binding bond integrals shown reproduce the LDA bands using just first and second neighbour matrix elements, this is not the case for the γ-representation. Note that the first and second

order tight binding matrix elements are essentially the same; the additional second order terms may be safely neglected and the first order orthogonal hamiltonian (33) is clearly the proper one to use for this case.

2. If you have the patience, then you can do this exercise by hand in the case of the first order hamiltonian.[48] However the scheme has been automated and is implemented in our LMTO suite.[45]

3. Unfortunately the on-site energies shown in the lower panel of figure 5 are far from independent of volume. This is a remaining unsolved question for the construction of tight binding models in which the on-site energies are invariably constant (except of course for the adjustments in self consistent models to account for electrostatic shifts due to charge transfer, see (44) below). Andersen[48] points out that the Dyson equation (31) provides guidance on how to account for this volume dependence in terms of the local neighbour environment. Whereas on-site matrix elements of the bare structure constants, $\mathbf{B}$ are zero, we have from (31)

$$B^{\alpha}_{\mathbf{R}L\,\mathbf{R}L} = \sum_{\mathbf{R}''\neq\mathbf{R}} \sum_{L''} B_{\mathbf{R}L\,\mathbf{R}''L''}\; \alpha_{\mathbf{R}''\ell''}\; B^{\alpha}_{\mathbf{R}''L''\,\mathbf{R}L}$$

and the on-site matrix element of (33) is[51]

$$\varepsilon_{\mathbf{R}L} = c_{\mathbf{R}L} + d_{\mathbf{R}L} B^{\alpha}_{\mathbf{R}L\,\mathbf{R}L}.$$

However the band centre parameter c and bandwidth parameter d are also strongly volume dependent.[9,44] An important contrast with the ASA is that in tight binding, the on-site parameters are constant—the scaling law has to take care of both the bond length dependence at constant volume *and* the volume dependence itself.[54]

3.3 Environmentally dependent tight binding matrix elements

Possibly the most striking feature displayed in figure 5 is a discontinuity, most notably in the $dd\pi$ and $dd\delta$ bond integrals, between first and second neighbours. This is of particular importance to structures like bcc which have first and second neighbours rather similar in bond length. It means that one *cannot* find a simple scaling law, such as the GSP (23) that can connect all the points in the graph. This effect was noticed in the case of the $ss\sigma$ bond integral in Mo by Haas *et al.*[37] and they proposed a very significant development in tight binding theory, namely the use of *environmentally dependent* bond integrals.[55] The discontinuities in the dd bond integrals were noticed by Nguyen-Manh *et al.*[53] who offered the physical explanation in terms of "screening." The basic idea is that the bond between two atoms is *weakened* by the presence of a third atom. Therefore the scaling of a bond integral, say by the GSP function (23) is modified by multiplying it by $(1-\mathcal{S}_{\ell\ell' m})$ where the "screening function," $\mathcal{S}_{\ell\ell' m}$, is the hyperbolic tangent of a function[37]

$$\xi^{\mathbf{R}\mathbf{R}'}_{\ell\ell' m} = A_{\ell\ell' m} \sum_{\substack{\mathbf{R}'' \\ \mathbf{R}''\neq\mathbf{R},\mathbf{R}'}} \exp\left[-\lambda_{\ell\ell' m}\left(\frac{|\mathbf{R}-\mathbf{R}''|+|\mathbf{R}'-\mathbf{R}''|}{|\mathbf{R}-\mathbf{R}'|}\right)^{\eta_{\ell\ell' m}}\right], \qquad (34)$$

in which A, λ and η are parameters to be fitted. This complicated expression can be simply explained.[37,53] As a third atom, $\mathbf{R}''$ approaches the $\mathbf{R}-\mathbf{R}'$ bond the term in parentheses becomes small, and approaches one in the limit that atom $\mathbf{R}''$ sits inside the $\mathbf{R}-\mathbf{R}'$

bond. This increases the value of the exponential and the tanh function smoothly reduces the $\mathbf{R} - \mathbf{R}'$ bond integral. Whereas Tang *et al.*[55] introduced this function empirically, Nguyen-Manh *et al.*[53] were able to derive its form using the theory of bond order potentials, and explain *why* $dd\sigma$ is not strongly screened while $dd\pi$ and $dd\delta$ are. Modern tight binding models[51,56,57] for transition metals are now fitted to curves such as those in figure 5 using (34). Indeed in these new schemes a repulsive energy is also fitted to an environmentally dependent function similar to (34). This is intended to make a better description of the valence–core overlap[44,58] between atoms which is short ranged but not pairwise and is otherwise not properly captured in the tight binding bond model. So nowadays one finds instead of (20)

$$E_{\mathrm{B}} = E_{\mathrm{bond}} + E_{\mathrm{prom}} + E_{\mathrm{env}} + E_{\mathrm{pair}} \tag{35}$$

in the TBBM, and E_{env} is the new environmentally dependent repulsive energy; it being understood that E_{bond} may be constructed using environmentally dependent hopping integrals too. E_{prom} is sometimes omitted,[56,57] in the instance that only one orbital angular momentum is included in the hamiltonian, for example if one employs a d-band model for transition metals.

4 Self Consistent Tight Binding

We described a tight binding model for ZrO_2 in section 3.1. The local charge neutrality of the TBBM is clearly inadequate to describe an ionic crystal for which a dominant part of the total energy is the Madelung sum of electrostatic pair terms.[10] A way to deal with this in tight binding was proposed by Majewski and Vogl[42,59] based on a Hubbard-like hamiltonian of Kittler and Falicov.[60] In this scheme the total charge transfer at each site, $\Delta q_{\mathbf{R}}$, from (11) and (19) are taken as point charges. The hamiltonian is again

$$H = H^0 + H' \tag{36}$$

as in (22). Two terms make up H', the Madelung energy of the lattice of point charges and a positive energy that is quadratic in $\Delta q_{\mathbf{R}}$, namely $U_{\mathbf{R}} \Delta q_{\mathbf{R}}^2$; employing the well-known "Hubbard U" that acts to resist the accumulation of charge. This problem is solved self consistently. An extension of this scheme to allow the charge to be expressed as multipoles, not just monopoles, was proposed independently by Schelling *et al.*[61] and Finnis *et al.*[38] In the latter paper, the connection was made to density functional theory and the TBBM, so we will pursue the same argument here. As noticed by Elstner *et al.*[62] the Hohenberg–Kohn total energy in DFT can be expanded about some reference electron density, $\rho^0(\mathbf{r})$. If H^0 is the hamiltonian with effective potential generated by the reference density, and just as in section 2.1 its eigenfunctions are $|n\rangle$ then the total energy correct to second order is[63] (e is the electron charge)

$$
\begin{aligned}
E^{(2)} &= \sum_n f_n \left\langle n \left| H^0 \right| n \right\rangle \\
&- \int \rho^0(\mathbf{r}) V^0_{\text{xc}}(\mathbf{r}) \mathrm{d}\mathbf{r} - E^0_{\text{H}} + E^0_{\text{xc}} + E_{\text{ZZ}} \\
&+ \frac{1}{2} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \left\{ e^2 \, \frac{\delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right. \\
&+ \left. \delta\rho(\mathbf{r}) \frac{\delta^2 E_{\text{xc}}}{\delta\rho(\mathbf{r}) \delta\rho(\mathbf{r}')} \delta\rho(\mathbf{r}') \right\}. \qquad (37)
\end{aligned}
$$

E^0_{H} is the Hartree energy and E^0_{xc} and V^0_{xc} the exchange–correlation energy and potential belonging to the reference density, $\rho^0(\mathbf{r})$. The first two lines make up the Harris–Foulkes *first order* functional; we recognise the first line as the band energy, in fact the sum of occupied eigenvalues of the non self consistent *input* hamiltonian, and the second as the interaction term (double counting) plus the ion–ion pair potential, E_{ZZ}. In the *self consistent polarisable ion tight binding model*[38] (SCTB) we approximate the last two lines by a generalised Madelung energy and a Hubbard energy, which adds a *second order* energy[5] to (35)

$$
E_2 = \frac{1}{2} \, e^2 \sum_{\mathbf{R}L \, \mathbf{R}'L'} Q_{\mathbf{R}'L'} \, \tilde{B}_{\mathbf{R}'L' \, \mathbf{R}L} \, Q_{\mathbf{R}L} + \frac{1}{2} \sum_R U_{\mathbf{R}} \Delta q^2_{\mathbf{R}}. \qquad (38)
$$

These two terms represent the electron–electron interactions. All the exchange and correlation complexities are rolled into a single parameter, the Hubbard U. The first term in (38) is a classical interaction energy between point multipoles. The monopole term is just a straight forward sum of Coulomb energies, $\frac{1}{2} e^2 \, \Delta q_{\mathbf{R}} \Delta q_{\mathbf{R}'} / \, |\mathbf{R} - \mathbf{R}'|$, while the generalised Madelung matrix is just the LMTO bare structure constant matrix (28), or to be precise $B_{\mathbf{R}'L' \, \mathbf{R}L} = -(1/2\pi)(2\ell + 1)(2\ell' + 1) \tilde{B}_{\mathbf{R}'L' \, \mathbf{R}L}$. In general $Q_{\mathbf{R}L}$ is the multipole moment of angular momentum L at site $\mathbf{R}$. If we knew the charge density, which we don't in tight binding, then we could define the moment

$$
Q_{\mathbf{R}L} = \int \mathrm{d}\mathbf{r} \; \rho(\mathbf{r}) \; r^\ell \, Y_L(\mathbf{r}) \qquad (39)
$$

for $\ell > 0$; while for $\ell = 0$ we'll have

$$
Q_{\mathbf{R}0} = \Delta q_{\mathbf{R}} Y_0 = \sqrt{\frac{1}{4\pi}} \, \Delta q_{\mathbf{R}}.
$$

Although we don't know the charge density in tight binding, we know the eigenvectors of the hamiltonian and we can construct multipole moments from these. The monopole is of course proportional to the Mulliken charge transfer. Although in tight binding we don't even specify what the basis functions (3) are, we can take it that they comprise a radial part times an angular, spherical harmonic part, that is

$$
\langle \mathbf{r} | \mathbf{R}L \rangle = f_{\mathbf{R}\ell} \left(|\mathbf{r} - \mathbf{R}| \right) Y_L(\mathbf{r} - \mathbf{R}). \qquad (40)
$$

Then in terms of the eigenvector expansion coefficients (4), for $\ell > 0$ we may define

$$
Q_{\mathbf{R}L} = \sum_{L'L''} \sum_n f_n \, \bar{c}^n_{\mathbf{R}L'} c^n_{\mathbf{R}L''} \left\langle \mathbf{R}L' \left| \hat{Q}_{\mathbf{R}L} \right| \mathbf{R}L'' \right\rangle \qquad (41)
$$

in which the multipole moment operator is[64]

$$\hat{Q}_{\mathbf{R}L} = \hat{r}^\ell \, Y_L(\hat{\mathbf{r}}), \tag{42}$$

which follows as a consequence of (39). If we expand out the matrix element of $\hat{Q}_{\mathbf{R}L}$ using (40) and (42) we have

$$\begin{aligned}\left\langle \mathbf{R}L' \left| \hat{Q}_{\mathbf{R}L} \right| \mathbf{R}L'' \right\rangle &= \int r^2 \mathrm{d}r \; f_{R\ell'} \, f_{R\ell''} \, r^\ell \iint \mathrm{d}\Omega \, Y_{L''} \, Y_{L'} \, Y_L \\ &= \Delta_{\ell'\ell''\ell} \, C_{L'L''L},\end{aligned}$$

which introduces new tight binding parameters, $\Delta_{\ell'\ell''\ell}$. Selection rules which are policed by the Gaunt coefficients (29) demand that there are only seven new parameters, or two if one has a basis of only s and p orbitals. These parameters are

$$\begin{aligned}\Delta_{011} = \Delta_{101} &= \Delta_{spp} \\ \Delta_{112} &= \Delta_{ppd} \\ \Delta_{022} = \Delta_{202} &= \Delta_{sdd} \\ \Delta_{121} = \Delta_{211} &= \Delta_{pdp} \\ \Delta_{222} &= \Delta_{ddd} \\ \Delta_{123} = \Delta_{213} &= \Delta_{pdf} \\ \Delta_{224} &= \Delta_{ddg}.\end{aligned}$$

In fact these parameters are not entirely new, but are recognisable as the elements of crystal field theory—in the case $\ell' = \ell''$ they are the quantities $\langle r^\ell \rangle$.[65,66] So it's perhaps not surprising that these new parameters introduce *crystal field* terms into the hamiltonian. These are off-diagonal, on-site terms that we have up to now taken to be zero. However they are crucial in describing the bands of, for example, the transition metal oxides as in figure 3. The generalised Madelung energy in (38) implies that the electrons are seeing an electrostatic potential due to the multipole moments at all the atomic sites. Indeed, if the electrostatic potential in the neighbourhood of the atom at site $\mathbf{R}$ is expanded into spherical waves, we could write,

$$V_{\mathbf{R}}(\mathbf{r}) = \sum_L V_{\mathbf{R}L} \, r^\ell \, Y_L(\mathbf{r}) \tag{43}$$

and using standard electrostatics the $\mathbf{R}L$ coefficient in this expansion is

$$V_{\mathbf{R}L} = \sum_{\mathbf{R}'L'} \tilde{B}_{\mathbf{R}L\,\mathbf{R}'L'} \, Q_{\mathbf{R}'L'}.$$

Now in the same way that we arrived at (41), using (43) we can find the matrix elements of H', namely

$$H'_{\mathbf{R}L'\,\mathbf{R}L''} = U_{\mathbf{R}} \, \Delta q_{\mathbf{R}} \, \delta_{L'L''} + e^2 \sum_L V_{\mathbf{R}L} \, \Delta_{\ell'\ell''\ell} \, C_{L'L''L}. \tag{44}$$

Now all the ingredients of the self consistent tight binding scheme are assembled. H^0 is given by its matrix elements, determined as in non self consistent tight binding, described in section 3. After solving the orthogonal, or non orthogonal eigenproblem and finding

the eigenvector expansion coefficients, you build the multipole moments and using structure constants find the components, $V_{\mathbf{R}L}$, of the potential. Having also chosen the Δ and Hubbard U parameters, elements of H' are assembled and the eigenproblem is solved for $H^0 + H'$. This continues until self consistency.

One or two extensions have been omitted here.

1. Only *on-site* matrix elements of H' are non zero in this self consistent scheme. In fact in the case of a non orthogonal basis, due to the explicit appearance of bond charge (see equation (11) and subsequent remarks) also intersite matrix elements of H' are introduced. This is important because it allows the hopping integrals themselves to be affected by the redistribution of charge, as might be intuitively expected.[6,67] Details are to be found elsewhere.[5,23]

2. This scheme can be extended to admit spin polarisation in imitation of the local spin density approximation. This *magnetic tight binding* (figure 1) has also been described elsewhere and is omitted from these notes for brevity.[23]

Finally we should remark that the interatomic force is easily obtained in self consistent tight binding. Only the *first* and *third* terms in the TBBM (21) survive; in particular one still requires the derivatives of the matrix elements of H^0. The only additional contribution to the force comes from the *first term* in (38); there is no contribution from the second term (or from the Stoner term in magnetic tight binding[68]) because of the variational principle. Hence one requires only the classical electrostatic force on atom $\mathbf{R}$,

$$\mathbf{F}^{\text{es}}_{\mathbf{R}} = -\sum_{L} Q_{\mathbf{R}L} \boldsymbol{\nabla} V_{\mathbf{R}L}$$

which is consistent with the force theorem,[31–34] and repairs the inconsistency of the band model mentioned in section 2.3.

We illustrated the self consistent polarisable ion tight binding model (SCTB) in the study of phase transitions in ZrO_2 in section 3.1. It turns out that the extension of the point charge model to include polarisability introduces new physics that is essential in describing these phenomena. In particular the dipole polarisation of the anions drives the cubic to tetragonal transition. Furthermore, as seen in figure 3 the crystal field splitting of the cation d-bands is achieved naturally and the correct ordering is reproduced in cubic and octahedral crystal fields. Crystal field splitting is also largely responsible for the ligand bandwidth in the low symmetry rutile structure.

4.1 Application to small molecules

Now we will turn to a second example, the application to small molecules. The self consistent point charge model in this context and in the study of biological molecules has enjoyed enormous success thanks in particular to the work of Frauenheim, Elstner and colleagues.[69]

Here we demonstrate the SCTB model applied to the question of the polarisability of two small molecules, azulene and para-nitroaniline (pNA). Hopping parameters were taken from Horsfield *et al.*[70] and Hubbard U and Δ parameters chosen to to reproduce the ground state dipole moments predicted by the local density approximation. For azulene it is found that the self consistent point charge model is sufficient, but pNA cannot be described

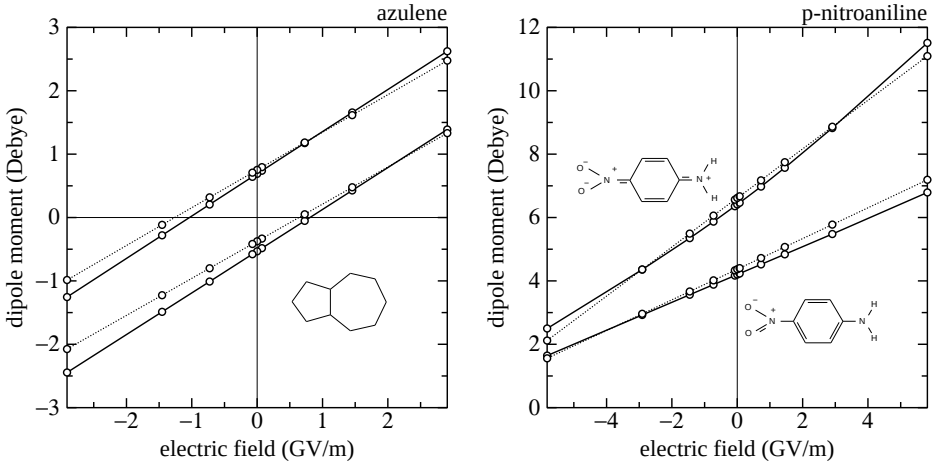

Figure 6. Dipole moment as a function of applied electric field calculated using LSDA, solid lines, and SCTB, dotted lines.[71] LSDA calculations were made using a molecule LMTO program.[72,73] The left hand figure shows the molecule azulene and the upper set of lines refer to the ground state and lower set to the so called S_1 excited state. The right hand figure shows p-nitroaniline; the lower set are the ground state and the upper set the "zwitterionic" first excited state

properly without dipole polarisability.[71] Figure 6 shows that the SCTB model provides a very accurate rendering of the dipole response to an applied electric field compared to LSDA calculations. We discuss now the two molecules in turn.

1. Azulene is a very interesting molecule having the same chemical formula as napthalene but comprising a five and seven membered ring instead of two six membered rings. According to Hückel's "$4n + 2$ rule," a ring molecule is especially stable if it has N π-electrons and $N = 4n + 2$, where n is an integer. This is because this leads to a closed shell of π-electrons.[74] Hence benzene is stable, having $n = 1$. By a similar argument a seven membered ring has an unpaired electron which can be used to occupy an unpaired hole in a five membered ring. Hence the ground state of azulene possesses a large dipole moment. An excited state is created if the electron is returned to the seven membered ring. As shown to the left of figure 6 the ground state dipole moment is positive (the positive axis pointing to the right) while its sign is reversed in the first excited state. Here we use a device which is not quite legitimate, namely in both LSDA and SCTB an electron–hole pair is created and self consistency arrived at under this constraint. While a very crude approximation to an excited state[75] (given that LSDA is a ground state theory) this does provide a useful test of the validity of the SCTB model. Indeed it is quite remarkable how the SCTB faithfully reproduces the LSDA even to the extent of accurately reproducing the polarisability of both ground and excited states. (The polarisability is the linear response of the dipole moment to an applied electric field, namely the slope in these figures.)

2. pNA is the archetypal "push–pull" chromophore.[76] In the ground state the dipole moment is small, but the first excited state is thought to be "zwitterionic," meaning

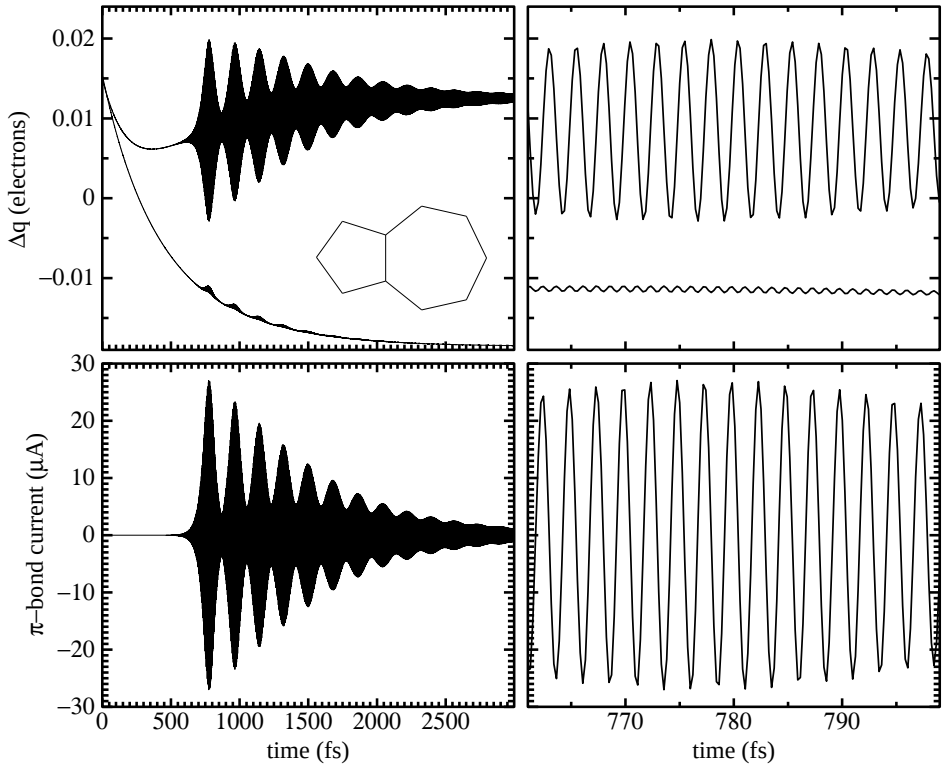

Figure 7. Charge transfer and bond current as a function of time in the relaxation of the S_2 excited state in azulene. The upper panels show the excess charge on a "bridge" atom and on the rightmost atom in the seven membered ring (lower curve). The lower panels show the π–bond current in the "bridge" bond.

that an electron transfers from the amine group on the right to the NO_2 group at the left increasing the dipole moment as shown on the right hand side of figure 6. Transfer of the electron through the π-system is called a push–pull process. Again the SCTB faithfully reproduces the LSDA with quantitative accuracy. We should mention again that it did not seem possible to obtain this result using a point charge self consistent tight binding model.

4.2 Ring currents in azulene

The SCTB model provides a simple scheme for the study of electron transfer as in the push–pull process. This is done by solving the time dependent Schrödinger equation using the hamiltonian H including electron–electron interactions. Indeed this is probably the simplest quantum mechanical model that goes beyond non interacting electrons. We have applied this approach to the relaxation of the S_2 excited state in azulene with some quite spectacular results.[77]

In terms of the density operator, the time dependent Schrödinger equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{\rho} = (\mathrm{i}\hbar)^{-1}\,[H,\hat{\rho}] - \Gamma\,(\hat{\rho} - \hat{\rho}_0)\,.$$

We have added a damping term with time constant Γ^{-1}. This allows us to prepare the molecule in an excited state and relax it into the ground state whose density operator is $\hat{\rho}_0$. The equation of motion is solved numerically using a simple leapfrog algorithm. While at the outset, the density matrix is real, during the dynamics it acquires complex matrix elements whose imaginary parts describe *bond currents*,[18]

$$j_{\mathbf{R}\mathbf{R}'} = \frac{2e}{\hbar}\sum_{LL'} H_{\mathbf{R}'L'\,\mathbf{R}L}\ \mathrm{Im}\,\rho_{\mathbf{R}L\,\mathbf{R}'L'}$$

which is the total current flowing from atom $\mathbf{R}$ to atom $\mathbf{R}'$. By selecting certain L-channels we can extract orbital contributions to j; in the present case of push–pull transfer we are interested in the current carried by the π-system of electrons.

Figure 7 shows results of such a simulation in azulene, using a time constant $\Gamma^{-1} = 500$ fs. Examine first the lower curve in the upper left panel. This is the excess total charge on the rightmost atom in the seven membered ring (see the inset in the top left panel). In the excited state, the dipole moment points to the left, that is, there is excess charge on this atom which transfers through the π-system to the left as the molecule relaxes into the ground state for which the dipole moment has opposite sign. The curve clearly show a smooth transfer of charge away from this site. However superimposed upon this is a series of oscillatory excursions in charge transfer, shown in a narrow time window by the lower curve in the upper right panel. Accompanying these oscillations are much larger fluctuations in the charge on the upper atom belonging to the "bridge" bond which is shared by both the five and seven membered rings. This excess charge is plotted in the upper curves of the upper left and right hand panels. As the upper and lower left hand panels show these oscillations die away, but analysis shows a quite characteristic frequency as seen in the right hand panels. The lower two panels show the π-bond current in the "bridge" bond. What is happening here is the setting up of ring currents in both rings whose directions are alternating with a period of a few femtoseconds. The ring currents at any one time are travelling in opposite senses in the two rings. This phenomena is a consequence of the electron–electron interaction, as we can verify by repeating the calculations using the non interacting hamiltonian, H^0. Because two bonds enter each bridge atom but only one leaves, the opposing sense of the currents means that charge will accumulate on one of these atoms to the point at which the Coulomb repulsion (described by the Hubbard U) resists further current flow and indeed reverses its direction. Note that each current reversal (lower right panel) is mirrored by the alternating charge transfer on the bridge atoms (upper right panel). It is not yet understood what fixes the frequency at which the reversal happens or what it is that makes the molecule particularly susceptible to this instability. We note that these ring currents require a long lead-in time, on the order of the time constant, to become established and this is probably because the symmetry breaking comes about through numerical round-off in the computer. In a more detailed simulation coupling the electrons to the molecular vibrations,[78] this symmetry breaking will derive from the coupling. We can confirm that the great majority of the current is indeed carried by the π-electron system.

5 Last Word

The intention here has been to provide a practical introduction to the tight binding method and to motivate students to try it for themselves. While this is a long article it is mostly conspicuous for what is missing, rather than what is included. This is not surprising in view of the vast literature and considerable age of the tight binding approximation, but I've tried to bring out issues that are less widely discussed elsewhere. Regrettably no connection has been made to the semi empirical approaches in quantum chemistry that bear a close resemblance. This reflects the fact that physicists and chemists frequently discover the same science independently and often without much awareness of each other's work. I hope that some of the most glaring omissions will be covered by other authors in this volume.[8,79]

Appendix

Real spherical harmonics are described in ref [64]. One takes the conventional, complex spherical harmonics[80] and makes linear combinations to get the real and imaginary parts.[81] Instead of m running from $-\ell$ to ℓ, m now runs from 0 to ℓ but for each $m > 0$, there are two real functions: $Y^c_{\ell m}$ which is $(-1)^m\sqrt{2}$ times the real part of $Y_{\ell m}$; and $Y^s_{\ell m}$ which is $(-1)^m\sqrt{2}$ times the imaginary part of $Y_{\ell m}$. For $m = 0$, $Y_{\ell m}$ is anyway real, so we throw away $Y^s_{\ell 0}$. We end up with the same number of functions, properly orthonormal. Specifically,

$$Y^c_{\ell m} = (-1)^m \frac{1}{\sqrt{2}} \left(Y_{\ell m} + \bar{Y}_{\ell m}\right)$$
$$Y^s_{\ell m} = (-1)^m \frac{1}{\mathrm{i}\sqrt{2}} \left(Y_{\ell m} - \bar{Y}_{\ell m}\right).$$

References

1. J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498, 1954
2. W. A. Harrison, *Electronic structure and the properties of solids: the physics of the chemical bond*, (W. H .Freeman, San Fransisco, 1980)
3. A. P. Sutton, *Electronic structure of materials*, (Clarendon Press, Oxford, 1993). The Slater–Koster table is reproduced in table 9.3
4. D. G. Pettifor, *Bonding and structure in molecules and solids*, (Clarendon Press, Oxford, 1995)
5. M. W. Finnis, *Interatomic forces in condensed matter*, (Oxford University Press, 2003)
6. A. P. Sutton, M. W. Finnis, D. G. Pettifor and Y. Ohta, J. Phys. C: Solid State Phys. **21**, 35, 1988
7. W. M. C. Foulkes, Phys. Rev. B **48**, 14216, 1993
8. R. Drautz, in this volume
9. O. K. Andersen, O. Jepsen and D. Glötzel, *Canonical description of the band structures of metals*, in *Proc. Intl. Sch. Phys., LXXXIX Corso, Varenna*, edited by F. Bassani, F. Fumi and M. P. Tosi (Soc. Ital. di Fisica, Bologna, 1984) p. 59

10. N. W. Ashcroft and N. D. Mermin, *Solid state physics*, (Holt-Saunders, 1976)
11. O. Jepsen and O. K. Andersen, Solid State Commun. **9**, 1763, 1971
12. G. Kresse and J. Furthmüller, Comp. Mater. Sci. **6**, 15, 1996
13. M. Reese, M. Mrovec and C. Elsässer, to be published
14. S. Glanville, A. T. Paxton and M. W. Finnis, J. Phys. F: Met. Phys. **18**, 693, 1988
15. A. T. Paxton, Phil. Mag. B **58**, 603, 1988
16. A. P. Horsfield, private communications
17. A. P. Horsfield, A. M. Bratkovsky, M. Fearn, D. G. Pettifor and M. Aoki, Phys. Rev. B **53**, 12694, 1996
18. T. N. Todorov, J. Phys.: Condens. Matter **14**, 3049, 2002
19. L. I. Schiff, *Quantum mechanics, 3*$^{\mathrm{rd}}$ *ed.*, (McGraw-Hill, 1968) p. 379
20. C. Kittel, *Quantum theory of solids, 2*$^{\mathrm{nd}}$ *ed.*, (John Wiley, 1987) p. 101
21. L. E. Ballentine and M. Kolář, J. Phys. C: Solid State Phys. **19**, 981, 1986
22. S. Elliott, *The physics and chemistry of solids*, (Wiley, 1998)
23. A. T. Paxton and M. W. Finnis, Phys. Rev. B **77**, 024428, 2008
24. J. Friedel, Suppl. Nuovo Cimento, Serie X **VII**, 287, 1958
25. P. W. Anderson, Phys. Rev. **124**, 41, 1961
26. M. Aoki, Phys. Rev. Letters **71**, 3842, 1993
27. D. J. Chadi, Phys. Rev. Letters **41**, 1062, 1978
28. J. Friedel, Trans. Metall. Soc. AIME **230**, 616, 1964
29. F. Ducastelle, J. de Physique **31**, 1055, 1970
30. D. G. Pettifor, in *Physical Metallurgy, 3*$^{\mathrm{rd}}$ *ed.*, edited by R. W. Cahn and P. Hassen, (Elsevier, 1983)
31. D. G. Pettifor, J. Chem. Phys. **69**, 2930, 1978
32. M. Methfessel and J. Kübler, J. Phys. F: Met. Phys. **12**, 141, 1982
33. A. R. Mackintosh and O. K. Andersen, in *Electrons at the Fermi surface*, edited by M. Springford (Cambridge University Press, 1980) p. 149
34. V. Heine, Solid State Physics **35**, edited by H. Ehrenreich, F. Seitz and D. Turnbull, (Academic Press, 1980) p. 1
35. D. Spanjaard and M. C. Desjonquères, Phys. Rev. B **30**, 4822, 1984
36. A. T. Paxton, J. Phys. D: Appl. Phys. **29**, 1689, 1996
37. H. Haas, C. Z. Wang, M. Fähnle, C. Elsässer and K. M. Ho, Phys. Rev. B **57**, 1461, 1998
38. M. W. Finnis, A. T. Paxton, M. Methfessel and M. van Schilfgaarde, Phys. Rev. Letters **81**, 5149, 1998
39. D. W. Richerson, *Modern ceramic engineering*, (Marcel Dekker, New York, 1992)
40. S. Fabris, A. T. Paxton and M. W. Finnis, Phys. Rev. B **63**, 94101, 2001
41. S. Fabris, A. T. Paxton and M. W. Finnis, Acta Materialia **50**, 5171, 2002
42. J. A. Majewski and P. Vogl, in *The structures of binary compounds*, edited by F. R. de Boer and D. G. Pettifor, (Elsevier Science Publishers, 1989) p. 287
43. O. K. Andersen, Solid State Commun. **13**, 133, 1973
44. D. G. Pettifor, J. Phys. F: Met. Phys. **7**, 613, 1977
45. `http://titus.phy.qub.ac.uk/Programs`
46. L. Goodwin, A. J. Skinner and D. G. Pettifor, Europhys. Letters **9**, 701, 1989
47. D. G. Pettifor, J. Phys. C: Solid State Phys. **5**, 97, 1972
48. O. K. Andersen, O. Jepsen and M. Šob, in Lecture Notes in Physics **282**, *Electronic*

band structure and its applications, edited by M. Yusouff, (Springer, Berlin, 1987)
49. O. K. Andersen and O. Jepsen, Phys. Rev. Letters **53**, 2571, 1984
50. A. R. Williams, J. Kübler and C. D. Gelatt, Jr., Phys. Rev. B **19**, 6094, 1979
51. H. Nakamura, D. Nguyen-Manh and D. G. Pettifor, J. Alloys Compounds **306**, 113, 2000
52. R. W. Tank and C. Arcangeli, phys. stat. sol. (b) **217**, 131, 2000
53. D. Nguyen-Manh, D. G. Pettifor and V. Vitek, Phys. Rev. Letters **85**, 4136, 2000
54. A. T. Paxton, in *Atomistic simulation of materials: beyond pair potentials*, edited by V. Vitek and D. J. Srolovitz, (Plenum, 1989) p. 327
55. M. S. Tang, C. Z. Wang, C. T. Chan and K. M. Ho, Phys. Rev. B **53**, 979, 1996
56. M. Mrovec, D. Nguyen-Manh, D. G. Pettifor and V. Vitek, Phys. Rev. B **69**, 094115, 2004
57. M. Mrovec, R. Gröger, A. G. Bailey, D. Nguyen-Manh, C. Elsässer and V. Vitek, Phys. Rev. B **75**, 104119, 2007
58. A. J. Skinner and D. G. Pettifor, J. Phys.: Condens. Matter **3**, 2029, 1991
59. J. A. Majewski and P. Vogl, Phys. Rev. Letters **57**, 1366, 1986
60. R. C. Kittler and L. M. Falicov, Phys. Rev. B **18**, 2506, 1978
61. P. K. Schelling, N. Yu and J. W. Halley, Phys. Rev. B **58**, 1279, 1998
62. M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, Th. Frauenheim, S. Suhai and G. Seifert, Phys. Rev. B **58**, 7260, 1998
63. W. M. C. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520, 1989
64. A. J. Stone, *The theory of intermolecular forces*, (Oxford University Press, 1996)
65. D. S. McClure, in *Phonons*, edited by R. W. H. Stevenson, (Oliver and Boyd, London, 1966) p. 314
66. A. M. Stoneham, *Theory of defects in solids*, (Oxford University Press, 1975)
67. M. Lannoo and J. Bourgoin, *Point defects in semiconductors I*, Springer Series in Solid State Sciences, **22**, (Springer, Berlin, 1981)
68. G. Liu, D. Nguyen-Manh, B.-G. Liu and D. G. Pettifor, Phys. Rev. B **71**, 174115, 2005
69. M. Elstner, Th. Frauenheim, J. McKelvey and G. Seifert, in *Special Section: DFTB symposium*, J. Phys. Chem. A **111**, 5607–5944, 2007
70. A. P. Horsfield, P. D. Godwin, D. G. Pettifor and A. P. Sutton, Phys. Rev. B **54**, 15773, 1996
71. A. T. Paxton, unpublished
72. M. Methfessel and M. van Schilfgaarde, Phys. Rev. B **48**, 4937, 1993
73. A. T. Paxton and J. B. Harper, Mol. Phys. **102**, 953, 2004
74. R. McWeeny, *Coulson's valence*, (Oxford University Press, 1979) p. 243
75. A. Hinchliffe and H. J. Soscún, Chem. Phys. Letters **412**, 365, 2005
76. A. M. Moran and A. Myers Kelly, J. Chem. Phys. **115**, 912, 2001
77. A. M. Elena, A. T. Paxton and T. N. Todorov, to be published
78. A. P. Horsfield, D. R. Bowler, H. Ness, C. G. Sánchez, T. N. Todorov and A. J. Fisher, Rep. Prog. Phys. **69**, 1195, 2006
79. M. Elstner, in this volume
80. J. D. Jackson, *Classical electrodynamics, 2*$^{\rm nd}$ *ed.*, (John Wiley, 1975) p. 99
81. M. Methfessel, *Multipole Green functions for electronic structure calculations*, (Katholieke Universiteit te Nijmegen, 1986)