

dislocation as they already lie on the glide plane. Jogs also will not impede the glide of an edge dislocation but on screw dislocations they will only be able to glide along the dislocation line; glide of the screw dislocation will require climb of the jog segment. Jogs then form pinning points on screw dislocations, preventing the movement of that section of the line due to the high energy cost associated with climb and requiring the absorption or emission of vacancies for each unit step.

2.2 Electronic Structure Methods

2.2.1 Tight Binding

In order to carry out any calculations for the wave function of a system one of the first approximations to be applied is that the motion of electrons and nuclei in a system may be separated, assuming that the electrons react instantly to any motion of the nuclei. Known as the Born-Oppenheimer approximation,[8] it works well because of the disparity in mass between electrons and nuclei; this means that electrons will remain in the lowest energy state, or ground state, as the atomic structure is changed. The ground state of the system may then be obtained from the time-independent Schrödinger equation below by minimisation of the energy with respect to the wave function.

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

Where \hat{H} is the Hamiltonian operator, made up of kinetic and potential energy operators, and Ψ is the wave function of the system, containing all information on its state. If the further assumption is made that the electrons may be modelled as independent particles, moving in an effective potential representing the other electrons and nuclei, then the Schrödinger equation for the non-interacting system of electrons separates

to give one-electron equations of the form:

$$\hat{H}\psi_n(\vec{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V_{eff}(\vec{r}) \right] \psi_n(\vec{r}) = \varepsilon_n \psi_n(\vec{r}) \quad (2.2)$$

\hat{H} now represents the single-particle Hamiltonian operator: the kinetic energy operator $-\frac{\hbar^2}{2m}\nabla^2$ reproduces the kinetic energy of a system of non-interacting electrons with the same density and the effective potential $V_{eff}(\vec{r})$ seen by each electron is independent of the others. This potential contains terms to approximate the many body effects, such that the ground-state charge density is the same as in the interacting system.[9] Each single-particle state, n , is described by its wave function, ψ_n , and its energy, ε_n , which are obtained as the eigenfunctions and eigenvalues of the single particle Hamiltonian respectively. Equation 2.2 may be simplified somewhat by assuming that the single-particle wave functions of the system are made up of a linear combination of localised atomic-orbital like basis states, ϕ_{RL} ; where R labels each atomic site by its position and $L = lm$ is a combined angular momentum index labelling each orbital type. Representing ψ_n with n and ϕ_{RL} with RL , this expansion may be written in Dirac bra-ket notation as:

$$|n\rangle = \sum_{RL} C_{RL}^n |RL\rangle \quad (2.3)$$

C_{RL}^n are the expansion coefficients of the basis set. The problem has now been simplified from one of finding unknown functions, ψ_n , to one of finding these unknown coefficients.

Generally equation 2.2 is treated as a matrix eigenvalue problem, and so the operators are rewritten in matrix form. the Hamiltonian matrix elements are conceived

as integrals of the basis states with the Hamiltonian operator.

$$H_{RLR'L'} = \int \phi_{RL}(\vec{r}) \hat{H} \phi_{R'L'}(\vec{r}) d^3\vec{r} = \langle RL | \hat{H} | R'L' \rangle \quad (2.4)$$

The basis states on neighbouring atoms are not necessarily orthogonal, and so may have an associated overlap matrix:

$$S_{RLR'L'} = \int \phi_{RL}(\vec{r}) \phi_{R'L'}(\vec{r}) d^3\vec{r} = \langle RL | R'L' \rangle \quad (2.5)$$

In the case of orthogonal orbitals, or those that have been orthogonalised by some procedure such as the Löwdin method,[10] each occurrence of the overlap matrix in the following equations may be replaced with the identity matrix.

The input density, ρ^0 , is commonly taken to be a superposition of spherically symmetric, atomic charge densities and so a related approximation is made in most empirical Tight Binding models that the effective potential may be represented as a sum of atom-centred contributions. These integrals are now made up of two orbitals and a potential centred on three sites: one-centre integrals have all three terms centred on the same site; two-centre, or hopping, integrals have the orbitals centred on two sites and the potential sharing one of these positions; and finally there are three-centre terms which have all three centred on different sites. The next approximation is to neglect the three-centre integrals, not just because their calculation is difficult and would obscure the physical clarity of the model but also due to the fact that this two centre approximation is implicit in the structure of the bond integrals described below and in the calculation of forces.[11] Other combinations, such as having both orbitals centred on one site and the potential on another, may or may not be included depending on the model. The remaining integrals are treated as adjustable parameters, fitted to reproduce previous results; the Hamiltonian will be made up of terms representing the orbital on-site energies and interactions between orbitals on neighbouring atoms.

In the interest of speed, further approximations are commonly made to restrict the interaction range to the first few shells of neighbours and the orbitals involved to those close to the energy range of interest, approximations which will anyway be more slight than discarding the three-centre terms.[12]

There would still be a large number of terms to fit but the number of independent elements may be reduced greatly using the convention laid down by *Slater and Koster*. [13] For a matrix element between two given orbitals on different atoms, spherical coordinate systems are constructed such that the vector between the atoms lies along the z -axis and the origin is centred on either atom; in integrating these functions over the spherical coordinates it is found that only those states that share the same component of angular momentum with respect to the bond axis will have a non-vanishing contribution to the integral. These components are labelled with σ , π and δ in analogy to the s, p and d orbitals; such that s orbitals may form σ bonds, p form σ and π bonds and d form σ , π and δ bonds, as illustrated in figure 2.3. The angular part of the bond integrals between any orbitals on neighbouring atoms is then a linear combination of these parametrised fundamental integrals depending on the direction cosines of the bond between the atoms, as laid out for several structures in *Slater and Koster's* paper.

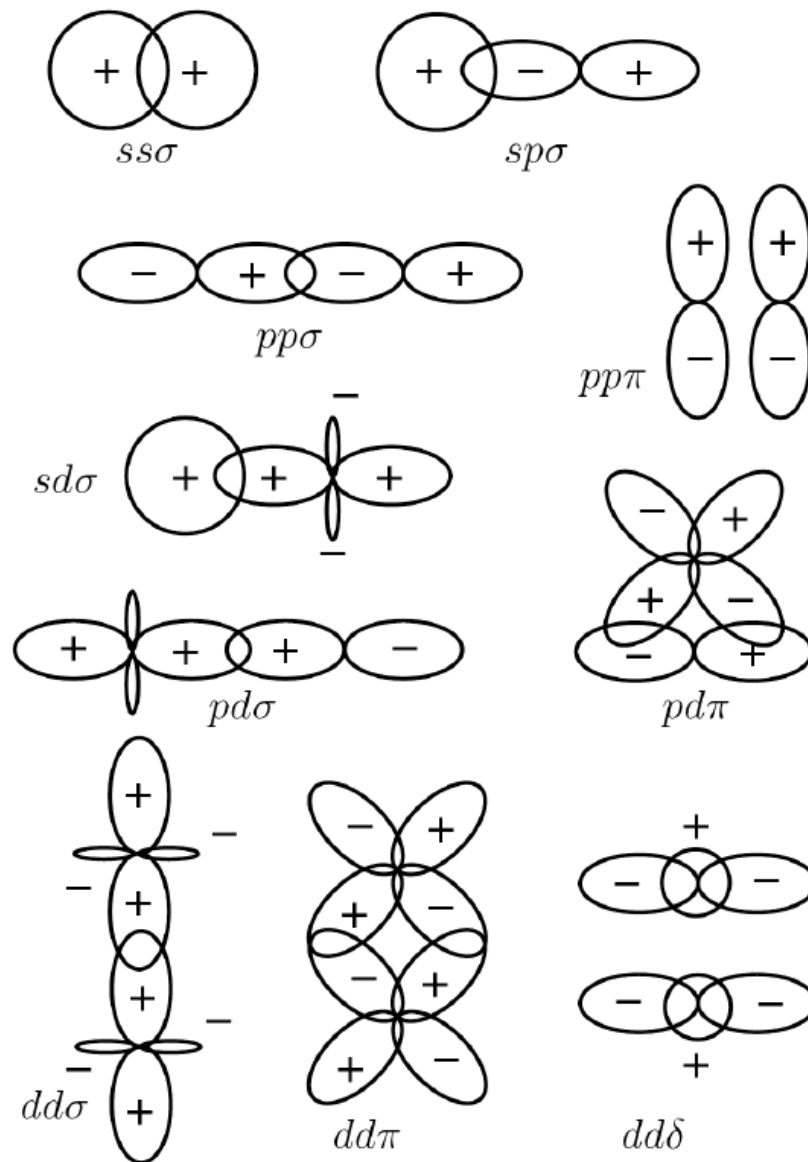


Fig. 2.3: Fundamental bond integrals that can be formed in a basis of s, p and d orbitals; the Latin letters label the orbitals involved and the Greek letter denotes the component of angular momentum they share with respect to the bond axis.[14]

The Hamiltonian and overlap matrices can then be written as:

$$H_{RLR'L'}^0 = \begin{cases} \epsilon_{RL}\delta_{LL'} & R = R' \\ f_{RLR'L'}(|R' - R|) & R \neq R' \end{cases} \quad (2.6)$$

$$S_{RLR'L'} = \begin{cases} \delta_{LL'} & R = R' \\ f_{RLR'L'}^{ovl}(|R' - R|) & R \neq R' \end{cases} \quad (2.7)$$

The superscript label on the Hamiltonian denotes that it is the non-self-consistent input Hamiltonian, before any of its elements have been altered by the self-consistency procedures that will be described below. The on-site terms (ϵ_{RL}) are parameters which must be chosen for each orbital type; the two-centre terms ($f_{RLR'L'}(|R' - R|)$ and $f_{RLR'L'}^{ovl}(|R' - R|)$) are the hopping integrals, which have an angular dependence from the linear combination of fundamental bond integrals, described above, multiplied by some function with a dependence on bond length. This function will constitute more parameters to fit, and is of great importance as the gradient of these integrals will partially determine the interatomic forces and so the stability of various structures. The functional form must be quite short ranged, going to zero beyond the first shells of neighbours without any discontinuity, but still flexible enough to fit to other results. These requirements may be met by with the use of two functions: one fitted to reproduce the previous results, and a cutoff function that matches the value, slope and curvature at some point and then smoothly reduces it to zero at a reduced range.[15] Another option is to use some function which has an increasing decay rate beyond a certain point, such as the Goodwin-Skinner-Pettifor function. [16]

The input Hamiltonian and overlap matrices have now been fully defined, and so the problem may be written as a linear eigenproblem and solved to obtain the band energies, ε_n , and expansion coefficients, C_{RL}^n , as the eigenvalues and eigenvectors

respectively.

$$\sum_{R'L'} H_{RLR'L'} C_{R'L'}^n = \varepsilon_n \sum_{R'L'} S_{RLR'L'} C_{R'L'}^n \quad (2.8)$$

The state of the system may be represented with the use of the density operator, defined by:

$$\hat{\rho} = \sum_n^\infty f_n |n\rangle \langle n| \quad (2.9)$$

Occupancy of the eigenstates is defined by f_n : at zero temperature this would be a step function, equal to two below the Fermi level and zero above in materials with degenerate spins. Finite electron temperature may be included with the Fermi-Dirac function below, although this will involve an additional entropic contribution to the energy (described at the end of section 2.2.2).

$$f_F(\varepsilon) = \frac{1}{1 + \exp(\frac{\varepsilon - \varepsilon_F}{k_B T})} \quad (2.10)$$

Where k_B is the Boltzmann constant, T the electron temperature and ε_F the Fermi level; the occupancy function is given by $f_n = 2f_F(\varepsilon_n)$. In materials with non-degenerate spins, i.e. magnetic materials, a separate Hamiltonian and so density matrix is produced for each spin, and so this occupancy function will vary between zero and one instead of zero and two (i.e. $f_n^\pm = f_F(\varepsilon_n)$); this will be dealt with in more detail when considering self-consistency.

The expansion coefficients of the Hamiltonian can then be used to calculate the expansion coefficients of the density operator for the electrons in each state.

$$\rho^{RLR'L'} = \sum_n^\infty f_n C_{RL}^n C_{R'L'}^{n*} \quad (2.11)$$

These density expansion coefficients may also be obtained through the use of the Green

function, a particularly useful relation as other methods exist to approximate it (such as the Bond Order Potential approximation). The Green function operator is defined by:

$$\hat{G}(\varepsilon) = \sum_n^\infty |n\rangle \frac{1}{\varepsilon - \varepsilon_n} \langle n| \quad (2.12)$$

This will be, by definition, equivalent to the inverse of the operator $(\varepsilon\hat{I} - \hat{H})$:

$$(\varepsilon\hat{I} - \hat{H}) \cdot \hat{G}(\varepsilon) = \hat{I} \quad (2.13)$$

The Green function operator is also related to the operator for the density of states, $\hat{D}(\varepsilon)$:

$$\hat{D}(\varepsilon) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} \frac{1}{\varepsilon + i\eta - \varepsilon_n} = -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} \hat{G}(\varepsilon + i\eta) \quad (2.14)$$

Where η is a small, positive value that is taken to zero from above. The density of states is related to the density by integration:

$$\hat{\rho} = \int_{-\infty}^{\infty} f(\varepsilon) \hat{D}(\varepsilon) d\varepsilon \quad (2.15)$$

$f(\varepsilon)$ is simply the occupancy f_n taken as a function of the energy, and so it is given by $f(\varepsilon) = 2f_F(\varepsilon)$ for a spin-degenerate material. Finally this relation provides the link between the Green function and the density expansion coefficients:

$$\rho^{RLR'L'} = -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} \int_{-\infty}^{\infty} f(\varepsilon) G^{RLR'L'}(\varepsilon + i\eta) d\varepsilon \quad (2.16)$$

$G^{RLR'L'}(\varepsilon + i\eta)$ are the expansion coefficients of the Green function operator in the

local basis:

$$G^{RLR'L'}(\varepsilon) = \sum_n \frac{C_{RL}^n C_{R'L'}^{n*}}{\varepsilon - \varepsilon_n} \quad (2.17)$$

Additionally, the expansion coefficients of the Green function may be used to obtain the local density of states $D_{RL}(\varepsilon)$, or the projection of the total density of states on orbital RL :

$$D_{RL}(\varepsilon) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} \sum_{R'L'} G^{RLR'L'}(\varepsilon + i\eta) S_{R'L'RL} \quad (2.18)$$

Once the density expansion coefficients have been obtained the charge in the system may be apportioned between atoms using Mulliken population analysis;[17] in an orthogonal system this simply amounts to a summation over the diagonal elements of the density matrix but with a non-orthogonal basis there will be a "bond charge" between atoms which must be included by multiplying by the overlap matrix:

$$q_R = \sum_L q_{RL} = \sum_{LR'L'} \rho^{RLR'L'} S_{R'L'RL} \quad (2.19)$$

If the summation were also over R this would have an identical form to the trace of the density operator.

The band energy, or the sum of the energies of occupied states, may then be obtained from either the eigenvalues or the density coefficients:

$$E_{band} = \sum_n f_n \varepsilon_n = \text{Tr} \hat{\rho} \hat{H}^0 = \sum_{RLR'L'} \rho^{RLR'L'} H_{R'L'RL}^0 \quad (2.20)$$

In the TB Band Model this energy forms part of the total energy of the solid, which *Chadi*[18] expressed as:

$$E_{tot} = E_{ee} + E_{ei} + E_{ii} \quad (2.21)$$

Where E_{ee} , E_{ei} and E_{ii} are the electron-electron, electron-ion and ion-ion interaction energies respectively. Unfortunately the band structure energy includes the E_{ee} interaction more than once since, in summing over the one-electron states, the energy of each electronic interaction will be included from the perspective of each electron. The TB band model total energy was written as:

$$E_{tot} = E_{band} + E_{clas} \quad (2.22)$$

Where E_{clas} is a repulsive pairwise term, originally interpreted to represent the ion-ion interactions and compensating for the double counting of the electron electron interaction:

$$E_{clas} = \frac{1}{2} \sum_{R \neq R'} V_{RR'}(|R' - R|) \quad (2.23)$$

The Tight Binding Bond Model instead approximates the difference in energy between the condensed and free atom states.[19] The binding energy may then be expressed as:

$$E_{bind} = Tr(\rho - \rho^0)\hat{H}^0 + E_{clas} \quad (2.24)$$

As mentioned, ρ^0 generally represents a superposition of, non-interacting, free atomic charge densities; the first term then arises from the bonding and charge redistribution in forming the solid. It was demonstrated by *Sutton et al.*,[19] with the use of the 1st order density functional defined by *Harris*[20] and *Foulkes*[21], that the pair potential now represents the change in electrostatic energy (including all electron-electron, electron-ion and ion-ion electrostatic interactions) and the exchange and correlation energies as the atoms are condensed, and that these terms may be well approximated by a pair potential. Exchange represents the reduction in total energy due to the

antisymmetry of the electron wave functions, preventing electrons of the same spin from sharing spatial states and so reducing Coulomb repulsion; while correlation is a measure of how the motion of one electron is affected by the presence of others around it. These terms are not calculated explicitly but rather fitted to reproduce some structural properties, such as the bulk moduli or lattice parameter. The Bond Model has now superseded the Band model in most applications, and forms the starting point for the Bond Order Potential, and so it is what I will continue with here.

In a real orbital basis ρ^0 will have no off-diagonal terms, and so the first term in equation 2.24 may be picked apart into on-site and inter-site parts. The on-site part, or promotion energy, defines the energy cost for the change in orbital occupancy from the free atom values, $(\rho^0)^{RLRL}$ (or N_{RL}), to those in the condensed state, ρ_{RLRL} .

$$E_{prom} = \sum_{RL} (\rho_{RLRL} - N_{RL}) H_{RLRL}^0 \quad (2.25)$$

The inter-site part of the band energy is equivalent to the bond energy, or a sum over the energies of all bonds in the system; it represents the largest contribution to the cohesion of the solid.

$$E_{bond} = \sum_{\substack{RLR'L' \\ R \neq R'}} \rho_{RLR'L'} H_{R'L'RL}^0 \quad (2.26)$$

The sum of the bond and promotion energies is then clearly equivalent to the band energy minus some reference energy of free atoms, $E_{ref} = \sum_{RL} N_{RL} \epsilon_{RL}$.

As mentioned, in the case of non-orthogonal orbitals a bond charge will build up between atoms, and so the promotion energy in equation 2.25 will not include all of the charge in the system; this may be rectified by replacing the ρ_{RLRL} in this equation

with the q_{RL} from equation 2.19, so that the promotion energy is now:

$$E_{prom} = \sum_{RL} (q_{RL} - N_{RL}) H_{RLRL}^0 \quad (2.27)$$

This results in a new definition for the bond energy, the covalent energy:[22]

$$E_{cov} = \sum_{\substack{RLR'L' \\ R \neq R'}} \rho^{RLR'L'} (H_{R'L'RL}^0 - S_{R'L'RL} H_{RLRL}^0) \quad (2.28)$$

The overlap term is included so that the identity $E_{prom} + E_{cov} = E_{band} - E_{ref}$ still holds.

In order to actually use the TB model for atomistic simulations some prescription is required to calculate the force, which is defined as the gradient of the binding energy with respect to atomic displacement. The derivative of the band energy part, or the trace of the product of the density matrix and Hamiltonian, may be obtained by making use of the Hellmann-Feynman theorem [23][24] to write it as the trace of the product of the density matrix and derivative of the Hamiltonian, or:

$$\frac{\partial}{\partial r_R} (Tr \hat{\rho} \hat{H}^0) = Tr \hat{\rho} \frac{\partial}{\partial r_R} \hat{H}^0 \quad (2.29)$$

Where r_R represents the coordinates of atom R . This derivative of \hat{H}^0 is not actually known, but it may be rewritten in terms of derivatives of the hopping integrals and overlap matrix elements:

$$Tr \hat{\rho} \frac{\partial}{\partial r_R} \hat{H}^0 = \sum_n f_n C_{RL}^n C_{R'L'}^{n*} \left[\frac{\partial H_{R'L'RL}^0}{\partial r_R} - \varepsilon_n \frac{\partial S_{R'L'RL}}{\partial r_R} \right] \quad (2.30)$$

Due to the two-center approximation applied previously these derivatives are not difficult to obtain. ρ^0 consists solely of on-site terms and so $Tr \hat{\rho}^0 \frac{\partial}{\partial r_R} \hat{H}^0$ will just be a summation over the diagonal. The derivative of the binding energy will then have the

form:

$$\frac{\partial E_{bind}}{\partial r_R} = Tr \hat{\rho} \frac{\partial}{\partial r_R} \hat{H}^0 - \sum_{RL} (\rho^0)^{RLRL} \frac{\partial H_{RLRL}^0}{\partial r_R} + \sum_{R' \neq R} \frac{\partial V_{RR'}}{\partial r_R} \quad (2.31)$$

The final term is the derivative of the pair potential, which is trivial to obtain due to its analytic form. The on-site terms of \hat{H}^0 do not vary with atomic displacement and so only gradients of the inter-site terms must be included in this calculation; the force on each atom R is then given by:

$$\begin{aligned} F_R = & - \sum_{\substack{LR'L' \\ R \neq R'}} \rho^{RLR'L'} \frac{\partial H_{R'L'RL}^0}{\partial r_R} \\ & + \sum_{\substack{LR'L' \\ R \neq R'}} E_{RLR'L'}^B \frac{\partial S_{R'L'RL}}{\partial r_R} \\ & - \sum_{R' \neq R} \frac{\partial V_{RR'}}{\partial r_R} \end{aligned} \quad (2.32)$$

The term $E_{RLR'L'}^B$ represents the band energy split into specific contributions:

$$\begin{aligned} E_{RLR'L'}^B &= \sum_n^{\infty} f_n C_{RL}^n C_{R'L'}^{n*} \varepsilon_n \\ &= -\frac{1}{\pi} \lim_{\eta \rightarrow 0} \int_{-\infty}^{\infty} f(\varepsilon) \varepsilon \text{Im} G^{RLR'L'}(\varepsilon + i\eta) d\varepsilon \end{aligned} \quad (2.33)$$

The formalism presented so far has been for finite systems, which may be sufficient for some calculations if the simulation cell is large enough but often the calculation could be greatly reduced with the use of periodic boundary conditions. For crystalline materials the atomic structure, and so the Hamiltonian and density matrix, has a repeating pattern for a given size of unit cell. The bulk properties of the crystal may then be calculated per unit cell by making the approximation that the crystal is infinite

and carrying the calculation out in k -space rather than real space. This is accomplished by replacing each local orbital with a Bloch sum that runs over the periodic images of that orbital. The Bloch-transformed Hamiltonian and overlap matrices are then:

$$H_{RLR'L'}(\vec{k}) = \sum_T e^{(i\vec{k}\cdot T)} H_{(R+T)LR'L'} \quad (2.34)$$

$$S_{RLR'L'}(\vec{k}) = \sum_T e^{(i\vec{k}\cdot T)} S_{(R+T)LR'L'} \quad (2.35)$$

Where T are the translation vectors of the unit cell or some integer combination of them; the summation runs over the infinite unit cells of the system but the number that actually have non-zero contributions will be determined by the range of the hopping-integrals. R and R' run over the atoms in the unit cell, so $T = 0$ would refer to matrix elements between atoms within the same unit cell, like in the finite system. The generalised eigenproblem (equation 2.8) can now be rewritten in terms of these Bloch-transformed matrices:

$$\sum_{R'L'} H_{RLR'L'}(\vec{k}) C_{R'L'}^n(\vec{k}) = \varepsilon_n(\vec{k}) \sum_{R'L'} S_{RLR'L'}(\vec{k}) C_{R'L'}^n(\vec{k}) \quad (2.36)$$

\vec{k} is restricted to values that lie within the first Brillouin zone, a primitive cell in k -space defined by the area that is closer to the central Γ -point ($\vec{k} = 0$) than any of the surrounding reciprocal lattice points. The eigenvalues of the Hamiltonian are only unique within the first Brillouin zone, and so it is unnecessary to consider k -points beyond it. The previous quantities obtained from the diagonalisation, such as the density matrix, must now be obtained as a summation over all of the allowed k -points; however, an infinite lattice will result in an infinite number of these points. This summation would then be equivalent to the integration of some continuous function of \vec{k} over the volume of the first Brillouin zone, which can be carried out by sampling at

carefully selected points and weighting them accordingly.[?] [12] This is an acceptable solution as the eigenvalues of the Hamiltonian vary smoothly with \vec{k} , so that only a finite number of points need to be treated; although a higher variation with \vec{k} will mean that a higher density sampling mesh is required for the convergence of results. The computational expense involved in the solution of the eigenproblem means that it is of interest to reduce the number of k -points that need be sampled and there are a number of sophisticated sampling methods to improve this convergence.[25][26]

2.2.2 Bond Order Potential

The BOP model is based on the same TB Hamiltonian but bypasses the computationally heavy matrix equation (equation 2.8) by generating the Green function matrix from the moments of the local density of states associated with each orbital. Each successive moment of the density of states will describe a further property of the distribution, such as its norm, its centre of gravity, its mean square width and skewness.[27][28] The p^{th} moment of the local density of states for an orbital RL may be expressed as:

$$\mu_{RL}^p = \int_{-\infty}^{\infty} \varepsilon^p D_{RL}(\varepsilon) d\varepsilon = \langle RL | \hat{H}^p | RL \rangle \quad (2.37)$$

This expression may be expanded through the insertion of the identity operator, defined by $\sum_{RL} |RL\rangle \langle RL| = \hat{I}$, between the \hat{H} terms $(p - 1)$ times.

$$\mu_{RL}^p = \sum_{R^{(1)}L^{(1)}R^{(2)}L^{(2)}\dots} \langle RL | \hat{H} | R^{(1)}L^{(1)} \rangle \langle R^{(1)}L^{(1)} | \hat{H} | R^{(2)}L^{(2)} \rangle \dots \langle R^{(p-1)}L^{(p-1)} | \hat{H} | RL \rangle \quad (2.38)$$

This relation then shows an explicit link between the p^{th} moment of the local density of states for an orbital RL and the sum of all hopping paths of length p through the

lattice starting and ending with orbital RL . The moments of the local density of states for each orbital may then be calculated independently, depending only on the value of p rather than the total atom number, and so this calculation will scale linearly with system size. Also, since the electronic properties are generated from the local atomic environment, periodic boundary conditions may be approximated by surrounding the cell with a sufficient width of images to saturate the hopping paths; the calculation may then be carried out entirely in real-space.

However, recovering a function from its moments is difficult to accomplish; the recursion method[29][30] is commonly used to do this in a stable and convergent way. Starting with each state being considered, the Hamiltonian operator is recursively applied to generate a chain of orthogonal states, known as recursion orbitals.

$$|\widetilde{i+1}\rangle = \hat{H}|i\rangle - a_i|i\rangle - b_i|i-1\rangle \quad (2.39)$$

a_i and b_i are known as recursion coefficients, the tilde is used to mark that the new state has not yet been normalised. Starting from $i = 0$, this relation becomes:

$$|\widetilde{1}\rangle = \hat{H}|0\rangle - a_0|0\rangle \quad (2.40)$$

Since it is known that $b_0 = 0$. The state $|\widetilde{1}\rangle$ is required to be orthogonal to $|0\rangle$, and $|0\rangle$ is already normalised so operating from the left with it results in a definition for a_0 :

$$a_0 = \langle 0|\hat{H}|0\rangle \quad (2.41)$$

The state $|\widetilde{1}\rangle$ is then normalised by introducing the factor b_1 such that:

$$|1\rangle = \frac{1}{b_1}|\widetilde{1}\rangle \quad (2.42)$$

Which defines $b_1^2 = \langle \tilde{1} | \tilde{1} \rangle$. An alternate definition for b_1 may be given by operating on equation 2.40 from the left with $\langle 1 |$:

$$b_1 = \langle 1 | \tilde{1} \rangle = \langle 1 | \hat{H} | 0 \rangle = \langle 0 | \hat{H} | 1 \rangle \quad (2.43)$$

This procedure may then be repeated for the subsequent values of i and acts to tridiagonalise the Hamiltonian, it is known as the Lanczos algorithm.[31] The Hamiltonian elements are now given by:

$$\begin{aligned} H_{ii} &= a_i \\ H_{i+1,i} &= H_{i,i+1} = b_{i+1} \end{aligned} \quad (2.44)$$

The problem has now been mapped onto a linear chain of states, with on-site elements a_i and hopping integrals b_i between nearest neighbours only. The purpose of this transformation is that the first diagonal element of the Green function matrix for this system may be represented exactly by a continued fraction.

$$G_{00}(\varepsilon) = \frac{1}{\varepsilon - a_0 - \frac{b_1^2}{\varepsilon - a_1 - \frac{b_2^2}{\dots}}} \quad (2.45)$$

Each additional level of this continued fraction corresponds to adding a further 2 moments to the density of states distribution, where an infinite number of levels would be an exact solution of the TB model.[32]

If the density of states is made up of a single band with a finite width, such as that of a TB model with a single d -orbital, then the recursion coefficients will eventually converge to constant values (a_∞ and b_∞) and so may be used to terminate the continued fraction after a given level of accuracy has been obtained. This terminator must satisfy

the relation:

$$t(\varepsilon) = \frac{1}{\varepsilon - a_\infty - \frac{b_\infty^2}{\dots}} = \frac{1}{\varepsilon - a_\infty - b_\infty^2 t(\varepsilon)} \quad (2.46)$$

This has the form of a quadratic equation, that can be solved for the terminating function, known as the square root terminator.[29]

$$t(\varepsilon) = \frac{\varepsilon - a_\infty}{2b_\infty} \left(1 - \sqrt{1 - \left(\frac{2b_\infty}{\varepsilon - a_\infty} \right)^2} \right) \quad (2.47)$$

This terminator essentially couples the finite chain to an infinite effective medium.[12]

If the recursion operations are started on one of the basis states, ($|0\rangle = |RL\rangle$) then this continued fraction will correspond to the diagonal Green function matrix element for this orbital ($G_{RLRL}(\varepsilon)$), and so this constitutes a rapid method of obtaining the diagonal density matrix elements.

Obtaining the inter-site Green function matrix elements is more complex, and so only the key points will be summarised here; the full derivation may be found in the papers of *Horsfield et al.*[33][34] First, an auxiliary space is defined, spanned by vectors $|e_{RL}^\Lambda\rangle$, with the inner product:[35]

$$(e_{RL}^\Lambda | e_{R'L'}^\Lambda) = \Lambda_{RLR'L'} \quad (2.48)$$

Where $\sum_{RL} \Lambda_{RLRL} = 1$ and $\Lambda_{RLR'L'}$ has no physical meaning but serves to label bonds; each of these vectors in auxiliary space $|e_{RL}^\Lambda\rangle$ will have a unique corresponding vector in the Hilbert space of the TB basis set, $|RL\rangle$. A product space between the auxiliary space and the Hilbert space is also defined, in order to produce the composite vectors:

$$|W_0^\Lambda\rangle = \sum_{RL} |e_{RL}^\Lambda\rangle |RL\rangle \quad (2.49)$$

The expectation value of the Green function defined with respect to this vector is then:

$$\begin{aligned}
G_{00}^{\Lambda}(\varepsilon) &= \frac{\{W_0^{\Lambda}|\hat{G}(\varepsilon)|W_0^{\Lambda}\}}{\{W_0^{\Lambda}|W_0^{\Lambda}\}} \\
&= \frac{\sum_{RLR'L'} G_{RLR'L'}(\varepsilon) \Lambda_{RLR'L'}}{\sum_{RL} \Lambda_{RLRL}} \\
&= \sum_{RLR'L'} G_{RLR'L'}(\varepsilon) \Lambda_{RLR'L'} \tag{2.50}
\end{aligned}$$

Where quantities in the auxiliary space are labelled with a superscript Λ and the fact that operators in the Hilbert space (such as $\hat{G}(\varepsilon)$) do not act upon the vectors in auxiliary space is used during this derivation. The off-diagonal elements of the Green function matrix can then be obtained by differentiating with respect to $\Lambda_{RLR'L'}$, resulting in the relation:

$$\begin{aligned}
G_{RLR'L'}(\varepsilon) &= \sum_{i=0}^{\infty} \frac{\partial G_{00}^{\Lambda}(\varepsilon)}{\partial a_i^{\Lambda}} \frac{\partial a_i^{\Lambda}}{\partial \Lambda_{RLR'L'}} \\
&\quad + \sum_{i=1}^{\infty} \frac{\partial G_{00}^{\Lambda}(\varepsilon)}{\partial b_i^{\Lambda}} \frac{\partial b_i^{\Lambda}}{\partial \Lambda_{RLR'L'}} \\
&\quad + G_{00}^{\Lambda}(\varepsilon) \delta_{RR'} \delta_{LL'} \tag{2.51}
\end{aligned}$$

The $G_{00}^{\Lambda}(\varepsilon)$, a_i^{Λ} and b_i^{Λ} can be found by applying the recursion method (equation 2.39) with a composite vector $|W_0^{\Lambda}\rangle$ as the starting state.

The Dyson equation is then employed to obtain relations for the derivatives of $G_{00}^{\Lambda}(\varepsilon)$, which relates the change in the Green function to that in the Hamiltonian:

$$\begin{aligned}
\hat{G}(\varepsilon) &= \hat{G}_0(\varepsilon) + \hat{G}_0(\varepsilon) \delta \hat{H} \hat{G}(\varepsilon) \\
&= \hat{G}_0(\varepsilon) + \hat{G}_0(\varepsilon) \delta \hat{H} \hat{G}_0(\varepsilon) + \dots \tag{2.52}
\end{aligned}$$

Taking the variation in the Green function to first order in $\delta\hat{H}$ and expressing it in matrix form gives:

$$\delta G_{ij}(\varepsilon) = \sum_{kl} G_{ik}(\varepsilon) \delta H_{kl} G_{lj}(\varepsilon) \quad (2.53)$$

Therefore the derivatives of $G_{00}^\Lambda(\varepsilon)$ with respect to the tridiagonalised Hamiltonian elements are given by:

$$\frac{\partial G_{00}(\varepsilon)}{\partial a_i} = G_{0i}(\varepsilon) G_{i0}(\varepsilon) \quad (2.54)$$

$$\frac{\partial G_{00}(\varepsilon)}{\partial b_i} = G_{0i}(\varepsilon) G_{i-1,0}(\varepsilon) + G_{0,i-1}(\varepsilon) G_{i,0}(\varepsilon) \quad (2.55)$$

The $G_{0i}(\varepsilon)$ (which are equal to $G_{i0}(\varepsilon)$) elements may be obtained by taking the definition of the Green function $(Z - \hat{H})\hat{G} = \hat{I}$ and inserting the elements of the tridiagonal Hamiltonian to obtain another recursion relation:

$$(\varepsilon - a_i)G_{ij} - b_i G_{i-1,j}(\varepsilon) - b_{i+1} G_{i+1,j}(\varepsilon) = \delta_{ij} \quad (2.56)$$

Which can be rearranged to obtain:

$$G_{i+1,j}(\varepsilon) = \frac{\varepsilon - a_i}{b_{i+1}} G_{ij}(\varepsilon) - \frac{b_i}{b_{i+1}} G_{i-1,j}(\varepsilon) - \frac{\delta_{ij}}{b_{i+1}} \quad (2.57)$$

Each new element of $G_{ij}(\varepsilon)$ can then be obtained from the diagonal element $G_{00}(\varepsilon)$.

Finally equation 2.58 may be rewritten as:

$$\begin{aligned}
G_{RLR'L'}(\varepsilon) &= \sum_{i=0}^{\infty} G_{0i}^{\Lambda}(\varepsilon) G_{i0}^{\Lambda}(\varepsilon) \frac{\partial a_i^{\Lambda}}{\partial \Lambda_{RLR'L'}} \\
&\quad + 2 \sum_{i=1}^{\infty} G_{0,i-1}^{\Lambda}(\varepsilon) G_{i0}^{\Lambda}(\varepsilon) \partial b_i^{\Lambda} \frac{\partial b_i^{\Lambda}}{\partial \Lambda_{RLR'L'}} \\
&\quad + G_{00}^{\Lambda}(\varepsilon) \delta_{RR'} \delta_{LL'}
\end{aligned} \tag{2.58}$$

The derivation for the derivatives of the recursion coefficients with respect to $\Lambda_{RLR'L'}$ is quite involved and so will not be included here, but again may be found in the papers of *Horsfield et al.*[33][34]

There are several points in this method that require the evaluation of an integral with the Fermi-Dirac function (such as equations 2.16 and 2.68); this may be carried out in the complex plane by summing up an infinite series over the Matsubara poles, but the convergence of this series is very slow.[33] This convergence may be greatly improved by using a modified Matsubara scheme,[36] beginning with an approximation to the exponential function:

$$\exp(Z) \approx \left(1 + \frac{Z}{n}\right)^n \tag{2.59}$$

Which becomes exact as n tends to infinity. This approximation leads to an alternate representation of the Fermi-Dirac function (equation 2.10):

$$f_F(\varepsilon) = \frac{1}{1 + \exp(\beta(\varepsilon - \varepsilon_f))} \approx \frac{1}{\left(1 + \frac{\beta(\varepsilon - \varepsilon_f)}{2M}\right)^{2M} + 1} \tag{2.60}$$

Where $\beta = \frac{1}{k_B T}$. This approximation to the Fermi-Dirac function has $2M$ poles in a circle in the complex plane off the real axis, these poles are given by:

$$E_p = \varepsilon_F + \frac{2M}{\beta}(z_p - 1) \tag{2.61}$$

$$z_p = \exp\left(\frac{i\pi(2p+1)}{2M}\right) \quad p = 0, 1, \dots, 2M-1 \quad (2.62)$$

The integral of some function $A(\varepsilon)$, that is defined in the complex plane, with the Fermi-Dirac function may now be given as a summation over the poles of the approximated Fermi function:

$$\lim_{\eta \rightarrow 0} \text{Im} \int_{-\infty}^{\infty} f_F(\varepsilon) A(\varepsilon + i\eta) d\varepsilon = -\frac{2\pi}{\beta} \text{Re} \left[\lim_{M \rightarrow \infty} \sum_{p=0}^{M-1} z_p A(E_p) \right] \quad (2.63)$$

The relation between the density matrix and the Green matrix (equation 2.16) can then be replaced with:

$$\rho^{RLR'L'} = \frac{2}{\beta} \text{Re} \left[\sum_{p=0}^{M-1} z_p G^{RLR'L'}(E_p) \right] \quad (2.64)$$

For non-magnetic systems with degenerate spins this density matrix must be doubled, while in magnetic systems it is calculated with the Green matrix of each spin individually and summed.

As mentioned, finite electron temperature may be included with the Fermi-Dirac distribution; note that this is unrelated to the temperature of the solid as the motion of the electrons and nuclei are uncoupled. Increasing electron temperature will decrease the range of the density matrix and dampen long-range Friedel oscillations in metals, and so fewer moments will be required to accurately evaluate the local density of states.[37] Also, it has the effect of reducing the number of poles required for convergence in the modified Matsubara scheme. Using a finite temperature means that, instead of just the band energy $E_{band}(T)$, the band free energy $F_{band}(T)$ must be calculated by including the contribution from the entropy:

$$F_{band}(T) = E_{band}(T) - TS(T) \quad (2.65)$$

This entropic term accounts for the variation in states of the electrons near the Fermi energy, it is given by:

$$S(T) = \sum_{RL} \int D_{RL}(\varepsilon) \sigma(\varepsilon) d\varepsilon \quad (2.66)$$

Again, this contribution must be doubled for spin degenerate systems or calculated for each spin and summed otherwise. σ represents the entropic density function, which is:

$$\sigma(\varepsilon) = -k_B \{ f_F(\varepsilon) \log[f_F(\varepsilon)] + (1 - f_F(\varepsilon)) \log[1 - f_F(\varepsilon)] \} \quad (2.67)$$

This contribution has no effect on the evaluation of the forces provided that they are evaluated at a constant number of electrons.[33]

2.2.3 Self-Consistency

Neglect of the self consistency can cause inaccuracy in the energy and forces due to the lack of charge conservation and an improper treatment of the potential variation from charge flow.[19]

Local Charge Neutrality is implemented in many TB models as the simplest form of self-consistency. It is based on the insight that each atom will remain neutral, or that the charge associated with it is fixed; this is a good approximation for metallic materials as excess charge associated with atoms would be neutralised due to the screening of the electron gas.[38] This scheme may be effected in practice by iteratively updating the on-site elements of the Hamiltonian: atoms with an excess of charge will have their energy levels increased by an amount proportional to the excess and vice versa for those with a deficit. This continues until the charge transfer between atoms has been minimised to a certain degree, although charge transfers may still occur

between orbitals on the same atom.

A good estimate for these energy shifts can come from the susceptibilities, which may be calculated for each orbital from the diagonal elements of the Green function:

$$\begin{aligned}\chi_{00,00} &= \frac{1}{\pi} \lim_{\eta \rightarrow 0} \text{Im} \int f(\varepsilon) G_{00}(\varepsilon + i\eta) G_{00}(\varepsilon + i\eta) d\varepsilon \\ &= -\frac{1}{2} \frac{\partial q_{RL}}{\partial a_0}\end{aligned}\tag{2.68}$$

a_0 will be equal to the first on-site element in the linear recursion chain produced from each orbital, and so this may be rewritten as:

$$\chi_{00,00(RL)} = -\frac{1}{2} \frac{\partial q_{RL}}{\partial \epsilon_{RL}}\tag{2.69}$$

The shifts to the Hamiltonian for each orbital will then have the form:

$$H'_{RLRL} = \frac{\delta q_R}{2 \sum_L \chi_{00,00(RL)}}\tag{2.70}$$

These shifts in the on-site potentials must be included in the energy; the promotion energy with LCN self-consistency is then modified so that it includes the updated Hamiltonian rather than the input:

$$E_{prom}^{LCN} = \sum_{RL} (q_{RL} H_{RLRL} - N_{RL} H_{RLRL}^0)\tag{2.71}$$

This would now mean that the calculation for the atomic forces, or $\frac{\partial E_{bind}}{\partial r_R}$, would change: the on-site terms of the Hamiltonian in the promotion energy now vary with atomic separation and so their gradients would have to be included. Following *Sutton et al.*, [19] this may be avoided by adding and subtracting $\sum_{RL} N_{RL} H_{RLRL}$ so that the

promotion energy becomes:

$$E_{prom}^{LCN} = \sum_{RL} (q_{RL} - N_{RL}) H_{RLRL} \quad (2.72)$$

The energy splitting between orbitals on the same atom is maintained during the LCN cycle, and so this term could be restructured as:

$$E_{prom}^{LCN} = \sum_R H'_{RLRL} \sum_L (q_{RL} - N_{RL}) + \sum_{RL} (q_{RL} - N_{RL}) H_{RLRL}^0 \quad (2.73)$$

Due to the LCN requirements $\sum_L (q_{RL} - N_{RL}) = 0$, and the second term is constant with atomic displacement; therefore there is no contribution to the atomic force from this new promotion energy. The additional term that arises from this reshuffle, $\sum_{RL} N_{RL} (H_{RLRL} - H_{RLRL}^0)$, is assumed to be absorbed in the fitting of the pairwise potential, although this assumption would only be justified for potentials that were originally fit for LCN self-consistency.

A description of magnetism is essential for the simulation of magnetic iron phases, with an important contribution to the relative stability of each phase.[39] Magnetism may be included with the Stoner theory for itinerant ferromagnetism,[40][41] which includes the presence of local exchange fields as a contribution to the band energy. The magnetic moments are determined by the difference in occupation between states with spins parallel or anti-parallel to the resultant field; these electrons will see a different on-site potential depending on this orientation, giving Hamiltonian updates of the form:

$$H'_{RLRL} = \frac{\delta q_R}{2 \sum_L \chi_{00,00}(RL)} \pm \frac{1}{2} I_R m_R \quad (2.74)$$

I_R here represents the Stoner parameter, another ingredient of the model which must be fitted to reproduce the correct magnetic behaviour; the \pm sign gives the value of

these on-site terms for the spin-polarised Hamiltonians, depending on whether the spin is parallel or anti-parallel to the local magnetic moment. This means the each spin now has a different Hamiltonian and so density matrix, which must be summed, the occupancy function is now given by $f(\varepsilon) = f_F(\varepsilon)$. The magnetic moments must now also be made self-consistent like the charges; this procedure may be carried concurrently with the charge self-consistency procedure, although it is generally found that the magnetic moments converge more slowly than the charges and so there may be advantages in separating them. The magnetic contribution to the band energy will then have the form:

$$E_{mag}^{LCN} = \sum_R -\frac{1}{4} I_R m_R^2 \quad (2.75)$$

Polarisable Ion Tight Binding enables the description of charge transfer between atoms and the formation of point-multipoles on each atom. [42] The point multipoles are described through the inclusion of off-diagonal, on-site terms in the Hamiltonian, which act to couple orbitals of different angular momentum on the same site.[43] The Hamiltonian is adjusted iteratively with the potential generated by this charge distribution, until the charge difference between iterations is minimised beyond a certain degree and so the potential matches the induced charges.

At each step of the self-consistency cycle the charge multipoles are calculated.

$$Q_{RL} = \begin{cases} \sqrt{\frac{1}{4\pi}} \delta q_R & L = 0 \\ \sum_{L'L''} \rho_{RL'RL''} \Delta_{l'l''l} C_{L'L''L} & L > 0 \end{cases} \quad (2.76)$$

The parameters $\Delta_{l'l''l}$ must be set for the model, defining the strength of the l -multipole from the coupling of the orbitals with angular momentum l' and l'' . $C_{L'L''L}$ are the

Gaunt coefficients, integrals over three real spherical harmonics Y_L , which may be written as:

$$C_{L'L''L} = \int \int Y_L Y_{L'} Y_{L''} d\Omega \quad (2.77)$$

These may be calculated in the beginning and reused; the Gaunt coefficients will also decide the selection rules, limiting the number of coupling "strength" parameters that are required.[11]

If the assumption is made that the charges are sufficiently localised around the atomic sites, then the angular components of the electrostatic potential at an atomic site R may be related to the multipole moments of the charges on all atoms in the system:

$$V_{RL}^M = \sum_{\substack{R'L' \\ R \neq R'}} \tilde{B}_{RLR'L'}(|R - R'|) Q_{R'L'} \quad (2.78)$$

The $\tilde{B}_{RLR'L'}(|R - R'|)$ represents a matrix of structure constants, defined by:

$$\tilde{B}_{RLR'L'}(|R - R'|) = 16\pi^2 \sum_{L''} (-1)^l \frac{(2l'' - 1)!!}{(2l + 1)!!(2l' + 1)!!} C_{L'LL''} K_{L''}(|R - R'|) \quad (2.79)$$

Where the summation is restricted to values for which $L'' = L + L'$ and the $K_{L''}(r) = r^{-l-1} Y_L(r)$ are solid Hankel functions.

There is also a Hubbard potential, which represents the on-site self-interaction of the atomic charge.

$$V_R^U = \delta q_R U_R \quad (2.80)$$

The Hubbard U_R parameter will be another parameter to be fitted for each species:

a greater U will incur a greater energy cost for the accumulation of charge on any one site, and indeed the LCN approximation could be taken as the limit of having an infinite U .

The updates to the Hamiltonian due to these potentials will then have the form:

$$H'_{RL'RL''} = e^2 \sum_L V_{RL}^M \Delta_{l'l''l} C_{L'L''L} + V_R^U \delta_{L'L''} \quad (2.81)$$

This will result in a 2^{nd} order contribution to the energy, representing all of the electron-electron interactions:

$$E_2 = \frac{1}{2} \sum_R (e^2 \sum_L Q_{RL} V_{RL}^M + \delta q_R V_R^U) \quad (2.82)$$

Finally the force on an atom R due to these inter-site electrostatic interactions will be given by:

$$F_R^M = -\frac{e^2}{2} \sum_L Q_{RL} \nabla_R V_{RL}^M \quad (2.83)$$

In the case when only charge transfer is considered (without the formation of point multipoles) the Madelung potential is simply restricted to the monopole moment:

$$V_{R0}^M = \sum_{\substack{R' \\ R' \neq R}} \sqrt{4\pi} \frac{\delta q_{R'}}{|R - R'|} \quad (2.84)$$

The Hamiltonian updates and energy contribution then become:

$$H'_{RLRL}^{(CT)} = e^2 \sum_{R'} \frac{\delta q_{R'}}{|R - R'|} + V_R^U \quad (2.85)$$

$$E_2^{(CT)} = \frac{1}{2} \sum_R (e^2 \sum_{R'} \frac{\delta q_R \delta q_{R'}}{|R - R'|} + \delta q_R V_R^U) \quad (2.86)$$

The updates to the Hamiltonian are now strictly diagonal: the off-diagonal, on-site terms that coupled orbitals of different angular momenta become zero.

Again the Stoner theory of itinerant ferromagnetism[40] is employed to include magnetic effects, resulting in two spin-polarised Hamiltonians and a additional contribution to the band energy. The Hamiltonian increments and 2^{nd} order contributions to the energy for a magnetic PITB model would then have the form:

$$H'_{RLRL} = e^2 \sum_L V_{RL}^M \Delta_{l'l''l} C_{L'L''L} + ((U_R - \frac{1}{2}I_R)\delta q_R \pm \frac{1}{2}I_R m_R)\delta_{L'L''} \quad (2.87)$$

$$E_2 = \frac{1}{2} \sum_R (e^2 \sum_L Q_{RL} V_{RL} + (U_R - \frac{1}{2}I_R)\delta q_R^2 - \frac{1}{2}I_R m_R^2) \quad (2.88)$$

Note that in this case there is an additional contribution to the potential and energy from the Stoner terms that acts to reduce the Hubbard on-site interactions; this represents the degree to which the on-site Coulomb repulsion is reduced for electrons with like spins due to exchange.[44]

In the case of non-orthogonal orbitals there will also be off-site contributions to the Hamiltonian due to the presence of bond charge.

$$H'^M_{RLR'L'} = \frac{1}{2} O_{RLR'L'} (V_{R0}^M + V_{R'0}^M) \quad (2.89)$$

$$H''^U_{RLR'L'} = \frac{1}{2} O_{RLR'L'} (V_R^U + V_{R'}^U) \quad (2.90)$$

This will result in additional contributions to the forces:

$$\partial \rho_{RR'}^S = \sum_{LL'} \rho^{RLR'L'} \frac{\partial S_{RLR'L'}}{\partial r} \quad (2.91)$$

$$F_R^M = - \sum_{R'} (V_{R0}^M + V_{R'0}^M) \partial \rho_{RR'}^S \quad (2.92)$$

$$F_R^U = - \sum_{R'} (V_R^U + V_{R'}^U) \partial \rho_{RR'}^S \quad (2.93)$$

The origin of these forces can be understood by considering a pair of atoms moving relative to each other in a system: as the separation of the atoms changes so will their overlap, and so the bond charge that can be attributed to each atom. It is this change in charge, and so Hubbard potential, that gives rise to the Hubbard contribution to the force. The changing bond charge will also change the monopole moment on each atom, and so the Madelung contribution arises from the change in electrostatic interaction between these two sites and all other sites.[44]