



Electronic Band Structure Calculations

Stefan Hollos

Dec 1, 1998

Exstrom Laboratories LLC, 662 Nelson Park Dr, Longmont CO, 80503

www.exstrom.com

Contents

| | | |
|----------|---|-----------|
| 1 | Independent Electron Approximation | 2 |
| 2 | Bloch's Theorem | 2 |
| 3 | Wannier Functions | 3 |
| 4 | Wavefunction Expansion | 4 |
| 5 | The Energy Eigenvalue Equation | 5 |
| 6 | Integral Types | 5 |
| 7 | The Two Center Approximation | 6 |
| 8 | Appendix A: The Independent Electron Approximation | 8 |
| 9 | Appendix B: The Variational Method for Eigenvalue Problems | 10 |

1 Independent Electron Approximation

Most electronic band structure calculations begin with the assumption that the many body wavefunction of the electrons in a crystal can be expressed in terms of one electron wavefunctions that are solutions of the one electron Schrodinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + U(\vec{r})\right)\Psi_{\vec{k}}(\vec{r}) = E(\vec{k})\Psi_{\vec{k}}(\vec{r}) \quad (1)$$

The potential $U(\vec{r})$ must include the effect of the electron's interaction with the other electrons in the system as well as the potential due to the nuclei. This is known as the independent electron approximation. For a more detailed discussion and justification of the independent electron approximation see appendix A.

The potential will have the periodicity of the lattice: $U(\vec{r} + \vec{R}) = U(\vec{r})$, where \vec{R} is a lattice vector. The Hamiltonian will therefore also be periodic $H(\vec{r} + \vec{R}) = H(\vec{r})$. The solutions of the Schrodinger equation for a periodic Hamiltonian must obey Bloch's Theorem.

$$\Psi_{\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}}\Psi_{\vec{k}}(\vec{r}) \quad (2)$$

2 Bloch's Theorem

To prove Bloch's Theorem (for a detailed discussion of Bloch's Theorem see Ashcroft and Mermin [1] p. 133-141), define a translation operator $T_{\vec{R}}$ as follows:

$$T_{\vec{R}}f(\vec{r}) = f(\vec{r} + \vec{R}) \quad (3)$$

Then for Bloch's Theorem to be true $\Psi_{\vec{k}}(\vec{r})$ must be an eigenfunction of $T_{\vec{R}}$ with eigenvalue $e^{i\vec{k}\cdot\vec{R}}$. Since $\Psi_{\vec{k}}(\vec{r})$ is an eigenfunction of the Hamiltonian, it will also be an eigenfunction of $T_{\vec{R}}$ if $H(\vec{r})$ and $T_{\vec{R}}$ commute. The fact that they do commute can be seen from the following:

$$T_{\vec{R}}H(\vec{r})\Psi_{\vec{k}}(\vec{r}) = H(\vec{r} + \vec{R})\Psi_{\vec{k}}(\vec{r} + \vec{R}) = H(\vec{r})\Psi_{\vec{k}}(\vec{r} + \vec{R}) = H(\vec{r})T_{\vec{R}}\Psi_{\vec{k}}(\vec{r}) \quad (4)$$

To show that the eigenvalues of $T_{\vec{R}}$ have the form $e^{i\vec{k}\cdot\vec{R}}$ note that two successive translations are always equivalent to a single translation so that $T_{\vec{R}}T_{\vec{R}} = T_{\vec{R}+\vec{R}}$. This means that if $\lambda(\vec{R})$ is an eigenvalue of $T_{\vec{R}}$ then $\lambda(\vec{R})\lambda(\vec{R}') = \lambda(\vec{R} + \vec{R}')$. This identity is always satisfied by $\lambda(\vec{R}) = e^{i\vec{k}\cdot\vec{R}}$, proving Bloch's Theorem.

When \vec{R} is written in terms of the primitive lattice vectors, $\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3$, and \vec{k} is written in terms of the reciprocal lattice vectors, $\vec{k} = x_1\vec{b}_1 + x_2\vec{b}_2 + x_3\vec{b}_3$ where $\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{ij}$, then $e^{i\vec{k}\cdot\vec{R}} = e^{i2\pi(n_1x_1+n_2x_2+n_3x_3)} = \lambda(\vec{R}) = \lambda^{n_1}(\vec{a}_1)\lambda^{n_2}(\vec{a}_2)\lambda^{n_3}(\vec{a}_3)$ and $\lambda(\vec{a}_1) = e^{i2\pi x_1}$.

So far no restrictions have been placed on the vector \vec{k} . Imposing boundary conditions will limit the possible values of \vec{k} . To determine what kind of

boundary conditions to apply, note that in assuming a periodic potential we have actually assumed an infinite size crystal. Any real crystal obviously has a finite size and therefore surfaces that break the periodicity of the lattice. To remove the effect of the surfaces the crystal is periodically extended in the direction of each of the primitive lattice vectors. This leads to the following periodic boundary conditions:

$$\Psi_{\vec{k}}(\vec{r} + N_i \vec{a}_i) = \Psi_{\vec{k}}(\vec{r}) \quad i = 1, 2, 3 \quad (5)$$

where N_i is the number of lattice points in the \vec{a}_i direction. However, from Bloch's Theorem it must also be true that:

$$\Psi(\vec{r} + N_i \vec{a}_i) = e^{i2\pi x_i N_i} \Psi_{\vec{k}}(\vec{r}) \quad (6)$$

Comparing the last two equations: $2\pi x_i N_i = 2\pi m_i$ where m_i is an integer, so that $x_i = m_i/N_i$. Thus, for the stated periodic boundary conditions, \vec{k} must have the form:

$$\vec{k} = \frac{m_1}{N_1} \vec{b}_1 + \frac{m_2}{N_2} \vec{b}_2 + \frac{m_3}{N_3} \vec{b}_3 \quad (7)$$

It should also be noted that, as far as the wavefunction is concerned, two \vec{k} vectors that are related by a reciprocal lattice vector are entirely equivalent. To see this let $\vec{k}' = \vec{k} + \vec{G}$ where $\vec{G} = l_1 \vec{b}_1 + l_2 \vec{b}_2 + l_3 \vec{b}_3$, and l_i is an integer, then $e^{i\vec{k}' \cdot \vec{R}} e^{i\vec{G} \cdot \vec{R}} = e^{i\vec{k} \cdot \vec{R}}$ and $\Psi_{\vec{k} + \vec{G}}(\vec{r}) = \Psi_{\vec{k}}(\vec{r})$. The \vec{k} vectors can therefore be confined to the Wigner-Seitz primitive cell of the reciprocal lattice, also known as the first Brillouin zone, since any vector outside the zone differs from a vector inside the zone by a reciprocal lattice vector.

3 Wannier Functions

In developing the tight-binding formalism for band structure calculations it is convenient to write the wave function $\Psi_{\vec{k}}(\vec{r})$ in terms of a function $\phi(\vec{r})$ centered at each of the lattice sites:

$$\Psi_{\vec{k}}(\vec{r}) = \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \phi(\vec{r} - \vec{R}) \quad (8)$$

That this form satisfies Bloch's Theorem can be shown as follows:

$$\begin{aligned} \Psi_{\vec{k}}(\vec{r} + \vec{R}') &= \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \phi(\vec{r} + \vec{R}' - \vec{R}) \\ &= e^{i\vec{k} \cdot \vec{R}'} \sum_{\vec{R}} e^{i\vec{k} \cdot (\vec{R} - \vec{R}')} \phi(\vec{r} - (\vec{R} - \vec{R}')) \\ &= e^{i\vec{k} \cdot \vec{R}'} \sum_{\vec{R}} e^{i\vec{k} \cdot \vec{R}} \phi(\vec{r} - \vec{R}) \\ &= e^{i\vec{k} \cdot \vec{R}'} \Psi_{\vec{k}}(\vec{r}) \end{aligned}$$

The function $\phi(\vec{r})$ is called a Wannier function and it can always be chosen to be orthonormal in the sense that:

$$\int \phi^*(\vec{r})\phi(\vec{r}-\vec{R})d\vec{r} = \delta_{\vec{R},0} \quad (9)$$

In this case the properly normalized form of the wavefunction is

$$\Psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \phi(\vec{r}-\vec{R}) \quad (10)$$

where N is the number of lattice sites.

4 Wavefunction Expansion

The unknown function $\phi(\vec{r})$ can always be expanded in terms of some complete set of functions $\psi_n(\vec{r})$ as follows

$$\phi(\vec{r}) = \sum_n c_n \psi_n(\vec{r}) \quad (11)$$

In general, for an arbitrary set of basis functions $\psi_n(\vec{r})$, there will be an infinite number of terms in the expansion. In practice however, only a finite number of terms are used. The closer in form the $\psi_n(\vec{r})$ are to $\phi(\vec{r})$, the fewer the number of terms required to represent $\phi(\vec{r})$ with a given degree of accuracy. One perhaps obvious choice for the $\psi_n(\vec{r})$ are the atomic wavefunctions of the isolated atoms since these must evolve into $\phi(\vec{r})$ as the atoms are brought together to form a crystal (more will be said about this later). In any event, if the representation is not exact then the c_n coefficients in the expansion will have to be treated as variational parameters (for a review of the linear variational method see appendix B). Using the expansion for $\phi(\vec{r})$ the wavefunction becomes

$$\begin{aligned} \Psi_{\vec{k}}(\vec{r}) &= \frac{1}{\sqrt{N}} \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \sum_n c_n \psi_n(\vec{r}-\vec{R}) \\ &= \frac{1}{\sqrt{N}} \sum_n c_n \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \psi_n(\vec{r}-\vec{R}) \\ &= \sum_n c_n b_n(\vec{r}) \end{aligned} \quad (12)$$

where

$$b_n(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \psi_n(\vec{r}-\vec{R}) \quad (13)$$

5 The Energy Eigenvalue Equation

Evaluating the variation of the energy expectation value

$$E(\Psi_{\vec{k}}) = \frac{\langle \Psi_{\vec{k}} | H | \Psi_{\vec{k}} \rangle}{\langle \Psi_{\vec{k}} | \Psi_{\vec{k}} \rangle} \quad (14)$$

using the coefficients as variational parameters leads to the following equation (see appendix B)

$$\sum_n (H_{mn} - ES_{mn})c_n = 0 \quad (15)$$

This can be immediately recognized as a generalized eigenvalue problem involving the Hamiltonian matrix, H and an overlap matrix, S . The overlap matrix elements reflect the nonorthogonality of the basis functions and are given by:

$$\begin{aligned} S_{mn} &= \int b_m^*(\vec{r})b_n(\vec{r})d\vec{r} \\ &= \frac{1}{N} \sum_{\vec{R}\vec{R}'} e^{i\vec{k}\cdot(\vec{R}-\vec{R}')} \int \psi_m^*(\vec{r}-\vec{R}')\psi_n(\vec{r}-\vec{R})d\vec{r} \\ &= \sum_{\vec{R}} e^{i\vec{k}\cdot(\vec{R}-\vec{R}')} \int \psi_m^*(\vec{r}-\vec{R}')\psi_n(\vec{r}-\vec{R})d\vec{r} \\ &= \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \int \psi_m^*(\vec{r})\psi_n(\vec{r}-\vec{R})d\vec{r} \end{aligned} \quad (16)$$

If the $\psi_n(\vec{r})$ are orthonormal and there is no overlap between functions centered at different lattice sites then the overlap matrix should equal the identity matrix, $S = I$.

The Hamiltonian matrix elements likewise will be given by:

$$H_{mn} = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \int \psi_m^*(\vec{r})H\psi_n(\vec{r}-\vec{R})d\vec{r} \quad (17)$$

6 Integral Types

The periodic nature of the potential means that it can be written as a sum of potential functions centered on the lattice sites as follows:

$$U(\vec{r}) = \sum_{\vec{R}} u(\vec{r}-\vec{R}) \quad (18)$$

This means that the integrals in the expression for the Hamiltonian matrix elements will actually be a sum of integrals that can be classified according to how the two wavefunctions and the site-potential functions, $u(\vec{r})$, are centered on the atoms. When the two wavefunctions and the site-potential function are all

centered on the same atom, the integral will be called an **on-site integral**. An integral with the two wavefunctions centered on one atom and the site-potential function centered on a different atom, will also be called an on-site integral (when the integrals are later taken to be parameters the effect of these integrals will be considered to have been included in the regular on-site integrals and they will not be considered separately). When the site-potential function and one of the wavefunctions are centered on one atom and the other wavefunction is centered on a different atom, the integral will be called a **two center integral**. When the two wavefunctions and the site-potential function are all centered on different atoms, the integral will be called a **three center integral**.

7 The Two Center Approximation

The three center integrals are smaller and contribute less to the Hamiltonian matrix elements than the two center integrals. It is therefore tempting to ignore the three center terms altogether, since this greatly simplifies the calculation. This is known as the two center approximation. Using the two center approximation the integral in Eq. 17 will then only include a sum of site-potential functions centered on the atoms on which the wavefunctions are centered.

Evaluation of the integrals can be further simplified by assuming that the site-potential functions are spherically symmetric (this assumption holds better for ionic solids such as NaCl than it does for metals and covalent solids such as the diamond lattice semiconductors). In this case the functions must also be angular momentum eigenfunctions and they can therefore be written as the product of a radial function and a spherical harmonic.

$$\psi_n(\vec{r}) = R_L(r)Y_M^L(\theta, \phi) \quad (19)$$

The angular momentum quantum numbers can then be used to characterize the $\psi_n(\vec{r})$ so that they can be described as s-type ($L = 0$), p-type ($L = 1$), d-type ($L = 2$), and so on. The angular dependence of the $\psi_n(\vec{r})$ makes them analogous, in a way, to atomic wave functions. However, the radial part of $\psi_n(\vec{r})$ will in general be quite different than the wavefunction of the isolated atom. The $\psi_n(\vec{r})$ will be referred to as orbitals from here on.

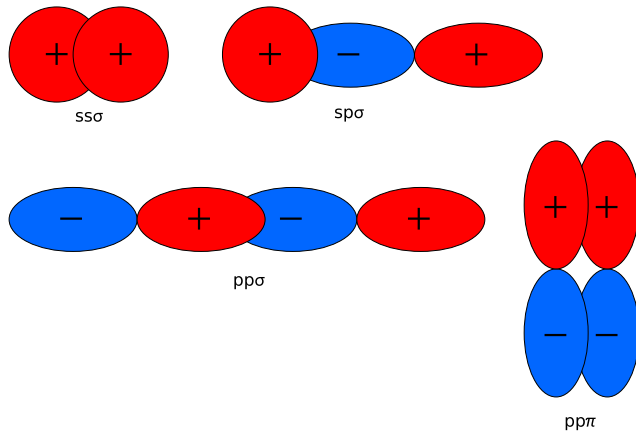


Figure 1: s and p type orbitals, in the two center approximation.

Using only s and p type orbitals, and the two center approximation, the types of integrals appearing in Eq. 17 are represented schematically in Fig. (1) for the simplest case in which the centers of the two atoms lie on the same Cartesian axis. The integrals are labeled using the Slater-Koster notation ($ll'm$) where l and l' give the angular momentum state of the two orbitals, and m gives the component of the angular momentum about the axis between the two atoms. This component is labeled as follows: s for $m = 0$, p for $m = 1$, and d for $m = 2$. The integral will be zero unless m is the same for both orbitals.

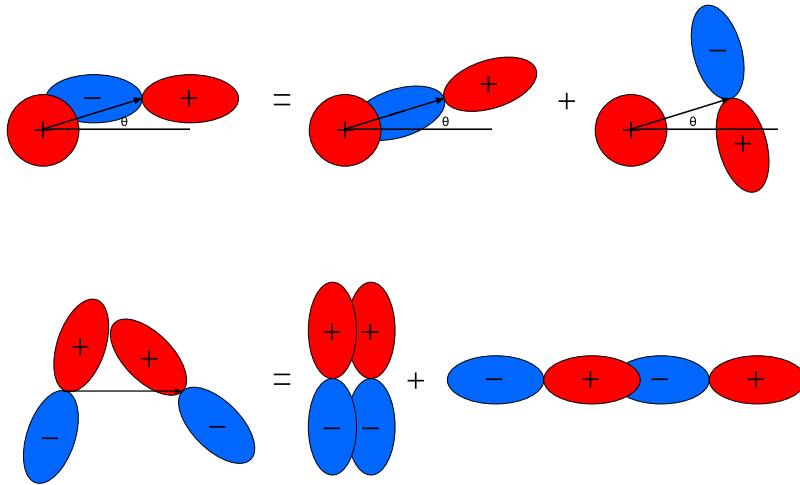


Figure 2: s and p type orbitals, in the general case.

For the general case, in which the two atoms do not lie on the same Cartesian axis, the integrals will depend on both the magnitude and direction of the vector

between the two centers as illustrated in Fig. (2) for s and p orbitals. For an sp integral the p orbital can be decomposed into a component along the axis between the two atoms and a component perpendicular to the axis. The component perpendicular to the axis will not contribute to the integral and the integral will just equal $(sps) \cos q$. In general let $\hat{\mathbf{a}}$ be a unit vector along the symmetry axis of the p orbital (pointing from the negative to the positive lobe) and let $\hat{\mathbf{d}}$ be a unit vector along the axis between the two atoms, then the sp integral will be given by: $\hat{\mathbf{a}} \cdot \hat{\mathbf{d}}(sps)$. For a pp integral the orbitals can be decomposed into components perpendicular and parallel to the axis between the two atoms. Let $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{a}}_2$ be unit vectors along the symmetry axes of the p orbitals and let $\hat{\mathbf{d}}$ be a unit vector along the axis between the two atoms, then the pp integral will be given by: $(\hat{\mathbf{a}}_1 \cdot \hat{\mathbf{d}})(\hat{\mathbf{a}}_2 \cdot \hat{\mathbf{d}})(pp\sigma) + (\hat{\mathbf{a}}_1 - (\hat{\mathbf{a}}_1 \cdot \hat{\mathbf{d}})\hat{\mathbf{d}}) \cdot (\hat{\mathbf{a}}_2 - (\hat{\mathbf{a}}_2 \cdot \hat{\mathbf{d}})\hat{\mathbf{d}})(pp\pi)$. Tables of two center integrals in terms of the $(ll'm)$ integrals can be found in Table 1 of Slater and Koster [8] and in Harrison [3] p. 481.

At this point, to proceed with the calculation of the electronic band structure, we need to obtain the values of the $(ll'm)$ integrals. An actual evaluation of the integrals would require knowledge of the site-potential function and the radial part of the expansion orbitals. A priori knowledge of these functions is usually not available. Several methods for calculating band structures in a self-consistent manner have been developed in this case. For an excellent review of these methods see reference [2].

Another tack to take is to regard the $(ll'm)$ integrals as parameters chosen so as to fit known energies at certain k -points. This method was first proposed by Slater and Koster and it is essentially an interpolation method. Nevertheless, the accuracy of the results can, in many cases, be quite good. When using this method the $(ll'm)$ integrals are usually referred to as SK parameters. SK parameters have been determined for many elements and compounds. An extensive source of SK parameters is the book by Papaconstantopoulos [6].

Once the $(ll'm)$ integrals are known, the Hamiltonian and overlap matrix elements can be calculated for a given vector and energy eigenvalues found. For a monatomic lattice using s, p, and d orbitals the Hamiltonian and overlap matrices will be 9x9 and we will get 9 eigenvalues at each vector. To generate an energy band we choose a path in the Brillouin zone of the lattice and calculate the eigenvalues at a given number of k -points along the path.

8 Appendix A: The Independent Electron Approximation

Solving for the eigenstates of a crystal is a many-body problem of extraordinary complexity. The electrons and the nuclei all interact with one another, making the problem impossible to solve analytically.

To get anywhere at all, some simplifying approximations have to be made. The first such approximation is called the Born-Oppenheimer approximation. The approximation states that, because the electrons are much lighter and move

much faster than the nuclei, the positions of the nuclei can be considered fixed and stationary when solving for the motion of the electrons. The motion of the electrons and the nuclei can then be solved for separately.

Within Born-Oppenheimer approximation the electrons experience a static periodic potential due to the nuclei and coulomb interactions with each other, so that the Hamiltonian for the electrons can then be written as:

$$H = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 - Ze^2 \sum_{i=1}^N \sum_{\vec{R}} \frac{1}{|\vec{r}_i - \vec{R}|} + \frac{1}{2} e^2 \sum_{i \neq j} \frac{1}{|\vec{r}_i - \vec{r}_j|} \quad (20)$$

where the vectors \vec{r}_i are the electron coordinates and the \vec{R} vectors are the nuclei coordinates. The first term is the kinetic energy operator for the electrons. The second term is the attractive electrostatic potential of the nuclei, and the third term is the repulsive electrostatic potential between the electrons.

This Hamiltonian is still too complicated to allow for an exact solution of its eigenstates. The major complicating factor is the electron-electron interaction term. One way to simplify this term is to assume that each electron moves in a static average potential due to all the other electrons. This is the Hartree approximation (HA). Energies calculated using the HA are usually much higher than the actual values. This is because the HA takes electron-electron interaction into account only in an average sense. It ignores electron correlation effects which tend to lower the energy by keeping the electrons further apart. A better approximation is the Hartree-Fock approximation (HFA). The HFA takes into account one type of correlation effect called the exchange interaction. The exchange interaction is due to the fact that electrons with identical spins tend to keep apart, thus lowering their energy. This is a physical manifestation of the required antisymmetry of a many-electron wavefunction under the interchange of two electrons.

A unifying framework for these approximations is the variational form of the Schrodinger equation. Let $\Psi(\vec{\mathbf{r}}_1 s_1, \vec{\mathbf{r}}_2 s_2, \dots, \vec{\mathbf{r}}_N s_N)$ be an N -electron wavefunction, then the energy expectation value for this state is

$$E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (21)$$

The expectation value is minimized when Ψ is the ground state eigenfunction of the Hamiltonian. In the Hartree approximation, the assumption is that Ψ can be expressed as a product of one-electron wavefunctions.

$$\Psi = \psi_1(\vec{\mathbf{r}}_1 s_1) \psi_2(\vec{\mathbf{r}}_2 s_2) \cdots \psi_N(\vec{\mathbf{r}}_N s_N) \quad (22)$$

The particular $\psi_i(\vec{\mathbf{r}}_i s_i)$ that minimize Eq. 21 will then be given by the solution of the Hartree equations:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi_i(\vec{\mathbf{r}}) - Ze^2 \sum_{\vec{\mathbf{R}}} \frac{1}{|\vec{\mathbf{r}} - \vec{\mathbf{R}}|} \psi_i(\vec{\mathbf{r}}) + e^2 \sum_j \int \frac{|\psi_j(\vec{\mathbf{r}}')|^2}{|\vec{\mathbf{r}} - \vec{\mathbf{r}}'|} d\vec{\mathbf{r}}' \psi_i(\vec{\mathbf{r}}) = E_i \psi_i(\vec{\mathbf{r}}) \quad (23)$$

This is a one-electron Schrodinger equation with a potential due to the ions and a charge density given by

$$\rho(\vec{\mathbf{r}}) = -e \sum_j |\psi_j(\vec{\mathbf{r}})|^2 \quad (24)$$

The problem with Eq. 22 however, is that Y is not antisymmetric. Maintaining Y as a function of one-electron wavefunctions and making it antisymmetric can be accomplished by expressing it as a determinant.

$$\Psi = \begin{vmatrix} \psi_1(\vec{\mathbf{r}}_1 s_1) & \psi_1(\vec{\mathbf{r}}_2 s_2) & \cdots & \psi_1(\vec{\mathbf{r}}_N s_N) \\ \psi_2(\vec{\mathbf{r}}_1 s_1) & \psi_2(\vec{\mathbf{r}}_2 s_2) & \cdots & \psi_2(\vec{\mathbf{r}}_N s_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\vec{\mathbf{r}}_1 s_1) & \psi_N(\vec{\mathbf{r}}_2 s_2) & \cdots & \psi_N(\vec{\mathbf{r}}_N s_N) \end{vmatrix}$$

In this case the particular $\psi_i(\vec{\mathbf{r}}_i s_i)$ that minimize Eq. 21 will be given by the solution of the Hartree-Fock equations. These equations are the same as Eq. 23 except for the appearance of a new potential term that gives the effect of the exchange interaction.

$$-e^2 \sum_j \int \frac{\psi_j^*(\vec{\mathbf{r}}') \psi_i(\vec{\mathbf{r}}') \psi_j(\vec{\mathbf{r}})}{|\vec{\mathbf{r}} - \vec{\mathbf{r}}'|} d\vec{\mathbf{r}}' \delta_{s_i s_j} \quad (25)$$

Both the Hartree and Hartree-Fock approximations therefore, lead to single particle Schrodinger equations with a periodic potential. Kohn and Sham [5, 7, 4] have furthermore shown that it is possible, in principle, to suitably modify the potential so as to include all electron correlation effects. This leads to density functional theory which we will not discuss here.

The conclusion then, is that in most, if not all cases, the interacting electron problem can be cast into an equivalent problem involving independent electrons moving in a static potential that takes into account the effect of the other electrons.

9 Appendix B: The Variational Method for Eigenvalue Problems

If $|\psi\rangle$ is an eigenvector of the Hamiltonian, then $H|\psi\rangle = E|\psi\rangle$ or

$$E = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} \quad (26)$$

Let $|\phi\rangle$ be an approximation to the eigenvector $|\psi\rangle$ such that $|\phi\rangle = |\psi\rangle + \alpha|\chi\rangle$ where α is a constant (possibly complex). We can define the energy associated with $|\phi\rangle$ as

$$E(\phi) = \frac{\langle \phi | H | \phi \rangle}{\langle \phi | \phi \rangle} \quad (27)$$

so that as $\alpha \rightarrow 0$, $E(\phi) \rightarrow E$. The error in the energy will then be

$$E(\phi) - E = \frac{\langle \phi | H | \phi \rangle \langle \psi | \psi \rangle - \langle \psi | H | \psi \rangle \langle \phi | \phi \rangle}{\langle \psi | \psi \rangle \langle \phi | \phi \rangle} \quad (28)$$

Using $|\phi\rangle = |\psi\rangle + \alpha|\chi\rangle$ this simplifies to

$$E(\phi) - E = |\alpha|^2 \frac{\langle \chi | H - E | \chi \rangle}{\langle \phi | \phi \rangle} \quad (29)$$

The error is second order in α , so that as long as α is relatively small, the eigenvalue approximation should be reasonably good.

If the eigenvalue problem is defined in a finite dimensional space of dimension N and the set of vectors $|u_i\rangle$ where $i = 1, \dots, N$ is a complete basis for the space then an eigenvector can always be represented exactly as a linear combination of the basis vectors $|u_i\rangle$.

$$|\psi\rangle = \sum_{i=1}^N c_i |u_i\rangle \quad (30)$$

Substituting this into the eigenvalue equation $H|\psi\rangle = E|\psi\rangle$ we get

$$\sum_{j=1}^N c_j H |u_j\rangle = E \sum_{j=1}^N c_j |u_j\rangle \quad (31)$$

Multiply both sides by $\langle u_i |$ to get

$$\sum_{j=1}^N \langle u_i | H | u_j \rangle c_j = E \sum_{j=1}^N c_j \langle u_i | u_j \rangle \quad (32)$$

Let $\langle u_i | u_j \rangle = S_{ij}$ and $\langle u_i | H | u_j \rangle = H_{ij}$ then we can write

$$\sum_{j=1}^N (H_{ij} - E S_{ij}) c_j \quad (33)$$

This is the usual finite dimensional generalized eigenvalue problem. The solution of this equation gives the exact eigenvalues of the Hamiltonian operator H . Note that for an orthonormal basis $\langle u_i | u_j \rangle = \delta_{ij}$ and $S = I$.

If the space is infinite dimensional or of dimension greater than N , then the linear combination of vectors $\sum_{i=1}^N c_i |u_i\rangle$ will not in general be able to exactly represent an eigenvector and will be an approximation to the actual eigenvector:

$$|\phi\rangle = \sum_{i=1}^N c_i |u_i\rangle \quad (34)$$

A familiar example of this is the expansion of a function as a fourier series:

$$f(x) = \sum_k c_k e^{ikx} \quad (35)$$

In general there will be an infinite number of terms in the series. However for some functions - smoothly varying functions that look similar to sines and cosines for example - a small finite number of terms will reproduce the function almost exactly. Indeed there are some functions that can be represented exactly with a finite fourier series. The key point is that the closer in form the basis functions are to the function you are trying to expand, the fewer the number of terms required to reproduce the function. Therefore taking Eq. 34 as an approximation to the true eigenvector and substituting into Eq. 27 we get an expression for the approximate eigenvalue

$$E(\phi) = \frac{\sum_{ij}^N c_i^* c_j \langle u_i | H | u_j \rangle}{\sum_{ij}^N c_i^* c_j \langle u_i | u_j \rangle} \quad (36)$$

Let $S_{ij} = \langle u_i | u_j \rangle$ and $H_{ij} = \langle u_i | H | u_j \rangle$ then we can write

$$E(\phi) \sum_{ij}^N c_i^* c_j S_{ij} = \sum_{ij}^N c_i^* c_j H_{ij} \quad (37)$$

$$\sum_{ij}^N (H_{ij} - E(\phi) S_{ij}) c_i^* c_j = 0 \quad (38)$$

We find the best approximation by minimizing $E(\phi)$ with respect to the coefficients c_i . This involves taking the derivative of the above equation with respect to c_i and setting $dE(\phi)/dc_i = 0$ which gives

$$\sum_j^N (H_{ij} - E(\phi) S_{ij}) c_j = 0 \quad (39)$$

The difference between Eq. 39 and Eq. 33 is that in Eq. 33 E is an exact eigenvalue since the eigenvector had an exact representation in the basis, but here we do not know whether the eigenfunction can be exactly represented in the basis so that $E(\phi)$ will in general be an approximation to the actual eigenvalue.

It can be shown that the $E(\phi)$'s will always be an upper bound to the true eigenvalues, i.e. the true eigenvalues will always be less than or equal to the $E(\phi)$'s.

References

- [1] Neil W. Ashcroft and N. David Mermin. *Solid State Physics*. Saunders College, 1976.
- [2] W. Y. Ching. Theoretical studies of the electronic properties of ceramic materials. *J. Am. Ceram. Soc.*, 73(11):3135, nov 1990.
- [3] W. A. Harrison. *Electronic Structure and the Properties of Solids*. Dover, 1989.

- [4] R. O. Jones and O. Gunnarsson. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.*, 61(3):689–746, july 1989.
- [5] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133–A1138, nov 1965.
- [6] D. A. Papaconstantopoulos. *Handbook of the Band Structure of Elemental Solids*. Plenum, 1986.
- [7] M. Schluter and L. J. Sham. Density functional theory. *Phys. Today*, 35(2):36–43, feb 1982.
- [8] J. C. Slater and G. F. Koster. Simplified lcao method for the periodic potential problem. *Phys. Rev.*, 94(6):1498–1524, june 1954.