Chapter 4

Solving the Kohn-Sham equations

The LDA- or LSDA-DFT are well established methods that have been successfully applied to very different kinds of systems (molecules, clusters, solids, etc.). In the following, we will consider the application of the LDA-DFT to solids and discuss in successive chapters the description of strongly correlated systems.

Irrespective of whether one considers the Hartree, Hartree-Fock or density functional theory in the LDA approximation, one ends up with an infinite set of one-electron equations of the following type:

$$\underbrace{\left\{ -\frac{\hbar^2}{2m_e}\vec{\nabla}_i^2 + \frac{e^2}{4\pi\varepsilon_0}\int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,d\vec{r}' + V_\alpha + V_{\text{ext}} \right\}}_{\hat{H}_{\text{single-particle}}} \phi_i(\vec{r}) = \epsilon_i\phi_i(\vec{r}), \quad i = 1,\dots,N.$$

- ✓ For the Hartree approximation, V_α is equivalent to V_{SIC}.

- ✓ For the Hartree-Fock approximation, V_α is the exchange potential v_{exc}, ϕ_i are the *true* one-electron orbitals of the HF Slater determinant. *Exchange* is treated exactly, but *correlation* effects are not included at all.

- ✓ For the DFT in the LDA approximation (also LSDA, GGA), V_α is the exchange-correlation operator, ϕ_i are *mathematical* single-particle orbitals. *Exchange* and *correlation* are treated but both approximately.

Since formally the HF and KS equations are similar, one can use the same mathematical tools for both. Both sets of equations have to be solved self-consistently. In order to solve the equations, we need to find the expansion coefficients c_p^i in order to express the crystal wavefunction ϕ_i in a given basis set $\{\phi_p^{\text{basis}}\}$:

$$\phi_i = \sum_{p=1}^{P} c_p^i \phi_p^{\text{basis}}. \tag{4.1}$$

In principle, the wavefunction ϕ_i belongs to a function space that has infinite dimensions, i.e., P is *infinite*. In practice, though, we work with finite basis sets, which means that ϕ_i

cannot be exactly described. One's goal is to find a basis that generates a function close to ϕ_i.

Once a *basis* has been chosen, we come to an eigenvalue problem. Considering (4.1), the single-particle equation,

$$\hat{H}_{\text{KS}}\,\phi_i(\vec{r}) = \epsilon_i\,\phi_i(\vec{r}),$$

becomes

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \langle\phi_n^{\text{basis}}|\hat{H}_{\text{KS}}|\phi_m^{\text{basis}}\rangle - \epsilon_i\langle\phi_n^{\text{basis}}|\phi_m^{\text{basis}}\rangle & \dots \\ \dots & \dots & \dots \end{bmatrix} \begin{bmatrix} c_1^i \\ \vdots \\ c_p^i \end{bmatrix}. \tag{4.2}$$

If the basis set is orthonormal, the overlap matrix S is unity. Diagonalization of hamiltonian matrix (4.2) leads to P eigenvalues and P sets of expansion coefficients, which express each of the P eigenfunctions in the chosen basis. The larger P is, the better approximations for the eigenfunctions will be obtained. Index $\boxed{i}$ refers to relevant quantum numbers (n, $\vec{k}$, etc.), depending on the system to be described.

How to choose a good basis set? If the basis functions are similar to ϕ_i, then we only need a few of them to accurately describe ϕ_i (P is small). Such a basis set is called *efficient*. However, in order to be able to choose an efficient basis, one should know the solution before starting solving the model, which is rarely the case. In general, we would like to have a basis set that is simultaneously *efficient* and *unbiased*.

Let us consider a periodic solid. The lattice potential is translationally invariant,

$$V(\vec{r}) = V(\vec{r} + \vec{T}),$$

which implies restrictions on the wavefunctions of the crystal. These wavefunctions have to be basis functions of the irreducible representation of the translation group and are numerated with an index $\vec{k}$. They are eigenfunctions of the translation operator,

$$\mathcal{P}(\vec{T})\Psi_{\vec{k}}(\vec{r}) = \Psi_{\vec{k}}(\vec{r} + \vec{T}) = e^{i\vec{k}\vec{T}}\Psi_{\vec{k}}(\vec{r}),$$

and of the Schrödinger equation,

$$\left[-\nabla^2 + V(\vec{r}) \right]\Psi_{\vec{k}}(\vec{r}) = E(\vec{k})\Psi_{\vec{k}}(\vec{r}). \tag{4.3}$$

Wavefunctions defined in such a way are Bloch functions. With this definition, we can reduce the problem of determining the electronic structure of a crystal to considering a single unit cell with volume Ω.

$E(\vec{k})$ in (4.3) describes the bandstructure of the solid. $\Psi_{\vec{k}}(\vec{r})$ and $E(\vec{k})$ possess, as functions of $\vec{k}$, the translation symmetry of the reciprocal lattice. Therefore, it is sufficient to only consider $\vec{k}$ vectors in the primitive cell of the reciprocal lattice, i.e., in the Brillouin zone (BZ) with volume Ω_{BZ},

$$\Omega_{\text{BZ}} = \frac{(2\pi)^3}{\Omega}.$$

In this case, the bands $E_\nu(\vec{k})$ and wavefunctions $\Psi_{\vec{k}\nu}(\vec{r})$ have to be additionally characterized by a band index ν.

With the knowledge of the bandstructure, we can calculate other quantities that are relevant for experimental comparison. Among these, the important quantities are the number of states for a given energy, i.e., the "density of states" (DOS), per spin (σ) $D_\sigma(E)$,

$$D_\sigma(E) = \frac{\Omega}{(2\pi)^3} \sum_\nu \int_{\text{BZ}} d\vec{k}\, \delta(E - E_{\nu\sigma}(\vec{k})),$$

and the integrated density of states $N_\sigma(E)$,

$$N_\sigma(E) = \int_{-\infty}^{E} D_\sigma(E')\, dE'.$$

With $D_\sigma(E)$ and $N_\sigma(E)$ known, we can determine the Fermi energy E_F, which separates occupied from unoccupied states, from the condition that

$$n_\sigma = \int_{-\infty}^{E_F} D_\sigma(E)\, dE,$$

where n_σ is the number of electrons per spin in the primitive unit cell.

The various existing methods, developed for the purpose of solving the bandstructure problem, differ from one another in two points:

(1) the basis set representation chosen for the crystal wavefunction and

(2) the kind of approximation used within this choice.

Let me enumerate and give a short description of these choices.

1 - **LCAO method** (local combination of atomic orbitals).

This method employs a fixed *energy-independent* basis set, which allows a very good *local description*. With the variation principle, one obtains an eigenvalue problem of the form:

$$(\hat{H} - E\hat{O}) \cdot \vec{a} = 0,$$

$\hat{H}$: hamilton operator matrix,

$\hat{O}$: overlap matrix.

In general, one needs a large number of LCAO basis functions in order to obtain a good accuracy. One difficulty associated with this choice is the calculation of integrals over potentials and orbitals centered at different sites.

2 - **Partial-wave method**.

With this method, one considers *energy-dependent* basis functions. The solution of the Schrödinger equation is given by functions that are continuous and differentiable at the cell border. An algebraic formulation of this condition leads to a homogeneous system of linear equations of the form:

$$A(E) \cdot \vec{x} = 0,$$

which have a complicated energy dependence. The advantage of using partial waves is that the number of basis functions to be considered is small (for instance, $(l_{\max} + 1)^2 = 9$ per atom, if one chooses spherical functions).

2.1 - A way of combining the advantages of the LCAO and partial-wave methods is given by the **linear methods**. In this formalism, one constructs fixed basis sets from partial waves, which are used in a variation process. One considers a Taylor expansion of the partial waves around a fixed energy E_μ. This method leads to a standard eigenvalue equation and provides a good description of the bandstructure around E_μ.

3 - A different kind of approximations consists in **modelling the cell-potential**.

Usually, the cell-potential in the vicinity of an atom can be very well approximated by a *central symmetric* potential, while between atoms (in the interstitial) it is relatively flat. Consequently, in the region near the atom, the wavefunction oscillates strongly, while in the interstitial it is quite smooth. In the context of the scattering theory, atoms are considered as scatterers, and the scattering process is defined in terms of *phase-shifts* δ. Based on these considerations, various formalisms have been developed, which we discuss below.

3.1 - **Pseudopotential method**.

Instead of the true potential, one constructs an effective operator called a pseudopotential. The pseudopotential reproduces well the true potential of the valence wavefunction outside a chosen distance to the nucleus and is very weak inside the core region. In terms of the scattering theory, this approximation reproduces phase-shifts avoiding the complication of strong oscillating valence wavefunctions near the core. The quality of a given pseudopotential is defined by its ability to correctly reproduce electronic properties in different environments.

3.2 - **Korringa-Kohn-Rostoker** (KKR) scattering method.

The starting point of this method consists in splitting the unit cell into non-overlapping muffin-tin (MT) spheres around atoms and defining the remaining space as interstitial. Then, the cell-potential is replaced by a pseudopotential,

which is spherical symmetric inside the MT spheres and flat in the interstitial (IS):

$$V(r) = \begin{cases} V(r_R) & \text{for } r_R \leq S_R, \\ \\ V_0 & \text{for } r \in \text{IS} \end{cases}$$

(S_R is the radius of the MT sphere at $\vec{R}$). The wavefunction is expanded in terms of spherical Bessel/Neumann functions. The boundary conditions are set by reformulating the Schrödinger equation into an integral equation. The solution is obtained in the self-consistent multi-scattering between spheres. The KKR method is the starting point for the LMTO (linear muffin tin orbitals) method.

3.3 - <u>Augmented plane wave method</u> (APW).

In the APW method, the wavefunctions are solutions of the *radial Schrödinger equation* inside the MT spheres, while in the interstitial they are expanded in terms of plane waves. In this case, the boundary conditions are set by requiring continuity of the wavefunctions and their derivatives at the muffin tin borders.

Other approximations consider the full (non-spherical) cell-potential. Examples of the full-potential methods are FP-KKR, FLAPW and FP-LMTO.

4.1 Plane wave pseudopotential method

As was mentioned, there are two requirements for a basis set: it should be *unbiased* and *efficient*. Efficiency of a basis set means that the basis set size P,

$$\phi_i = \sum_{j=1}^{P} \phi_j^{\text{basis}},$$

should be small. From a technical point of view, it would also be good if the basis functions were mathematically *simple*.

The requirements of being unbiased and simple are realized in the *plane wave basis*. Any eigenfunction $\Psi_{\vec{k}}^n$ of a periodic hamiltonian can be expressed exactly in this basis set by means of an infinite set of coefficients $c_{\vec{K}}^{n,\vec{k}}$:

$$\Psi_{\vec{k}}^n(\vec{r}) = \sum_{\vec{K}} c_{\vec{K}}^{n,\vec{k}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}}, \tag{4.4}$$

where $\vec{K}$ is the reciprocal lattice vector. Expression (4.4) comes from the "*Bloch theorem*", which consists on the following.

Any eigenfunction $\Psi(\vec{r})$ of a periodic hamiltonian can be written as a product of a function that has the periodicity of the lattice $u_{\vec{g}}(\vec{r})$ and a planewave $e^{\vec{g}\cdot\vec{r}}$,

$$\Psi(\vec{r}) = u_{\vec{g}}(\vec{r})\, e^{i\vec{g}\cdot\vec{r}}, \tag{4.5}$$

where $\vec{g}$ is any vector of the reciprocal lattice:

$$\vec{g} = \vec{k} + \vec{K}, \tag{4.6}$$

$\vec{k}$: vector in the first BZ,
$\vec{K}$: reciprocal lattice vector.

Taking account of (4.6), expression (4.5) becomes

$$\Psi_{\vec{g}}(\vec{r}) = \left\{ u_{\vec{g}}(\vec{r})\, e^{i\vec{K}\cdot\vec{r}} \right\} e^{i\vec{k}\cdot\vec{r}},$$

where $e^{i\vec{K}\cdot\vec{r}}$ is commensurate with the lattice, i.e., $u_{\vec{g}}(\vec{r})\, e^{i\vec{K}\cdot\vec{r}}$ still has the periodicity of the lattice. One can rename $u_{\vec{g}}(\vec{r})$ as $u_{\vec{k}}^n(\vec{r})$, with n indicating the Brillouin zone where $\vec{g}$ was located. n and $\vec{k}$ contain the same information as $\vec{g}$. Then, given that

$$u_{\vec{k}}^n(\vec{r}) = \sum_{\vec{K}} c_{\vec{K}}^{n,\vec{k}} e^{i\vec{K}\cdot\vec{r}},$$

expression

$$\Psi_{\vec{k}}^n(\vec{r}) = u_{\vec{k}}^n(\vec{r}) e^{i\vec{k}\cdot\vec{r}}$$

can be written as

$$\boxed{\Psi_{\vec{k}}^n(\vec{r}) = \sum_{\vec{K}} c_{\vec{K}}^{n,\vec{k}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}}.}$$

Let us now go back to the discussion of the plane waves basis set. Since each basis function from this set is given as

$$\phi_{\vec{K}}^{\vec{k}}(\vec{r}) = |\vec{K}\rangle = e^{i(\vec{k}+\vec{K})\cdot\vec{r}},$$

this basis set is $\vec{k}$-dependent. This means that we need a whole basis set for each $\vec{k}$-point. In practical applications, it is not possible to work with an infinite set of plane waves. Therefore, one introduces a $\vec{K}$-vector cut-off,

$$K < K_{\max},$$

which corresponds to a sphere with radius $K_{\max}$ centered at the origin of the reciprocal lattice. All reciprocal lattice vectors inside this sphere are included into the basis set. The characteristic energy E_{cut}, defined as

$$E_{\text{cut}} = \frac{\hbar^2 K_{\max}^2}{2m_e},$$

is called the cut-off energy and corresponds to the energy of a free electron with wave-vector $K_{\max}$.

In order to write down the eigenvalue equation (4.2), the matrix elements of the KS hamiltonian are calculated in the plane wave basis. The overlap matrix in this basis is unity:

$$\langle \vec{K}_1 | \vec{K}_2 \rangle = \int e^{i(\vec{K}_2 - \vec{K}_1)\cdot\vec{r}} d^3\vec{r}$$
$$= \delta(\vec{K}_2 - \vec{K}_1).$$

For each eigenvalue $\epsilon^{n,\vec{k}}$, we get an eigenvector $[c_{\vec{K}}^{n,\vec{k}}]_P$, with P values for $c_{\vec{K}}^{n,\vec{k}}$. In fact, in one run, we get P different eigenfunctions with the same $\vec{k}$ but a different band index. Therefore, this process has to be repeated for each $\vec{k}$-point of a good sampling of the first Brillouin zone.

How many plane waves do we need for this procedure to expect an accurate result? The number of plane waves is determined by the smallest length scale to be described in real space. In order to get a feeling of typical length scales, let us consider the radial part of the $3s$ wavefunction in Ca (Fig. 4.1).
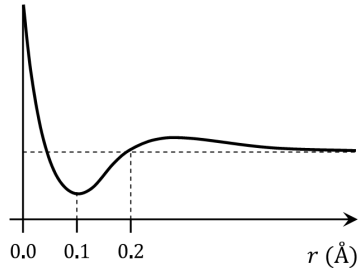

Figure 4.1: The radial part of the $3s$ wavefunction in Ca.

Near the nucleus, the wavefunction shows a steep behavior. In order to describe the wavefunction's behavior between 0.0 and 0.1 Å, we need plane waves with a period as small as roughly one order of magnitude less than this distance (0.01 Å or 10^{-12} m). Then,

$$K_{\max} = \frac{2\pi}{10^{-12}\,\mathrm{m}} = 6.3 \times 10^{12}\,\mathrm{m}^{-1}.$$

The volume of a sphere with radius $K_{\max}$ is

$$V \propto (K_{\max})^3 \sim 10^{39}\,\mathrm{m}^{-3}.$$

V divided by the volume of the first Brillouin zone V_{1BZ},

$$V_{1BZ} = \vec{c}^* \cdot \left(\vec{a}^* \times \vec{b}^* \right) \propto 9 \times 10^{30}\,\mathrm{m}^{-3} \quad \text{(assuming that } a, b, c \sim 3\,\text{Å}),$$

provides an estimate for the number of plane waves since the first Brillouin zone contains 1 $\vec{K}$:

$$\frac{V}{V_{1BZ}} \sim 10^8.$$

We need of the order of 10^8 plane waves. But diagonalization of a $10^8 \times 10^8$ matrix is beyond the computer capability! What shall we do?
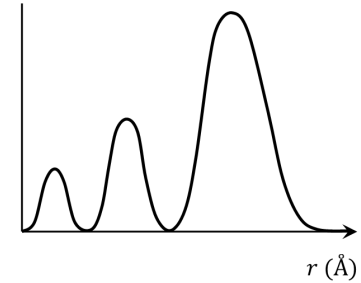

Figure 4.2: The radial probability distribution of a $3s$ electron in Ca.

The most oscillating part of the wavefunction is the tail close to the nucleus (Fig. 4.2). Since this inner region of the atom is shielded from the more outer regions, which are the regions we are most interested in, one can replace the potential in the inner region to yield very smooth tails of wavefunctions inside the atom such that only few plane waves are needed. As the outer regions of the atom are approached, the pseudopotential continuously evolves into the true potential (Fig. 4.3). A pseudopotential is called *soft* when few plane waves are needed. A pseudopotential is called *transferable* when it can be used for a considered element in various environments (molecule, cluster, solid, etc.).

Let us see mathematically how the pseudopotential is constructed. The valence wavefunctions can be expanded using a basis consisting of plane waves $\phi_{OPW}(\vec{k} + \vec{K})$ that have been *orthogonalized* to the lower lying core state $\{\varphi_c\}$.

$$\phi_{OPW}(\vec{k} + \vec{K}) = \phi_{PW}(\vec{k} + \vec{K}) - \sum_c \langle \varphi_c | \phi_{PW}(\vec{k} + \vec{K}) \rangle \varphi_c; \qquad (4.7)$$

$$\langle \varphi_{c_i} | \phi_{OPW} \rangle = \langle \varphi_{c_i} | \phi_{PW} \rangle - \sum_{c_j} \langle \varphi_{c_j} | \phi_{PW} \rangle \underbrace{\langle \varphi_{c_i} | \varphi_{c_j} \rangle}_{\delta_{ij}} = 0.$$
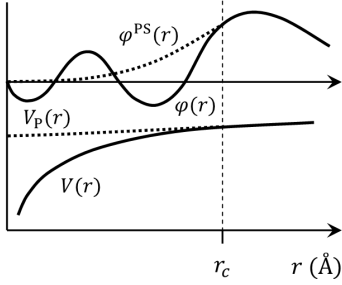
Figure 4.3: Illustration of the replacement of the all-electron wavefunction and core potential by a pseudo-wavefunction and pseudopotential.

A pseudopotential related to construction (4.7) can be defined as follows. Let H be the original Hamiltonian, with core and valence wavefunctions. We consider the pseudostates φ_v^{PS},

$$\varphi_v^{\mathrm{PS}} = \varphi_v + \sum_c a_{vc} \varphi_c, \qquad a_{vc} = \langle \varphi_c | \varphi_v^{\mathrm{PS}} \rangle. \tag{4.8}$$

Applying H to φ_v^{PS} gives

$$\begin{aligned} H | \varphi_v^{\mathrm{PS}} \rangle &= \epsilon_v | \varphi_v \rangle + \sum_c a_{vc} \epsilon_c | \varphi_c \rangle \\ &= \epsilon_v | \varphi_v^{\mathrm{PS}} \rangle + \sum_c a_{vc} (\epsilon_c - \epsilon_v) | \varphi_c \rangle, \end{aligned}$$

where ϵ_c and ϵ_v are the core and valence eigenvalues. Using the definition of a_{vc} [eq. (4.8)], we obtain

$$\left[H + \sum_c (\epsilon_v - \epsilon_c) | \varphi_c \rangle \langle \varphi_c | \right] \varphi_v^{\mathrm{PS}} = \epsilon_v^{\mathrm{PS}} \varphi_v^{\mathrm{PS}}.$$

The pseudostates satisfy a Schrödinger-like equation with an additional contribution V^R to the Hamiltonian:

$$V^R = \sum_c (\epsilon_v - \epsilon_c) | \varphi_c \rangle \langle \varphi_c |.$$

V^R is energy dependent. The *Phillips-Kleinman pseudopotential* V^{PK}, defined as

$$V^{\mathrm{PK}} = V + V^R,$$

is equal V outside the core region since there the core wavefunctions vanish.