Chapter 2

Theoretical Background: Density Functional Theory and Linear Response Calculations

2.1 Density Functional Theory

Over the past few decades, density functional theory (DFT) has been the most successful, widely used method in condensed-matter physics, computational physics and quantum chemistry to describe properties of condensed matter systems, which include not only standard bulk materials but also complex materials such as molecules, proteins, interfaces and nanoparticles. The main idea of DFT is to describe a many-body interacting system via its particle density and not via its many-body wavefunction. Its significance is to reduce the 3N degrees of freedom of the N-body system to only three spatial coordinates through its particle density. Its basis is the well known Hohenberg-Kohn (HK) theorem[1], which claims that all properties of a system can

be considered to be unique functionals of its ground state density. Together with the Born-Oppenheimer (BO) approximation[2] and Kohn-Sham (KS) ansatz[3], practical accurate DFT calculations have been made possible via approximations for the so-called exchange-correlation (XC) potential, which describes the effects of the Pauli principle and the Coulomb potential beyond a pure electrostatic interaction of the electrons. Since it is impossible to calculate the exact XC potential (by solving the many-body problem exactly), a common approximation is the so-called local density approximation (LDA) which locally substitutes the XC energy density of an inhomogeneous system by that of a homogeneous electron gas evaluated at the local density.

In many cases the results of DFT calculations for condensed-matter systems agreed quite satisfactorily with experimental data, especially with better approximations for the XC energy functional since the 1990s. Also, the computational costs were relatively low compared to traditional ways which were based on the complicated many-electron wavefunction, such as Hartree-Fock theory[4, 5] and its descendants. Despite the improvements in DFT, there are still difficulties in using DFT to properly describe intermolecular interactions; charge transfer excitations; transition states, global potential energy surfaces and some other strongly correlated systems; and in calculations of the band gap of some semiconductors.

2.1.1 The Many-Body System and Born-Oppenheimer (BO) Approximation

The Hamiltonian of a many-body condensed-matter system consisting of nuclei and electrons can be written as:

$$H_{tot} = -\sum_{I} \frac{\hbar^2}{2M_I} \nabla^2_{\mathbf{R}_I} - \sum_{i} \frac{\hbar^2}{2m_e} \nabla^2_{\mathbf{r}_i} + \frac{1}{2} \sum_{\substack{I,J \\ I \neq J}} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}$$
$$+ \frac{1}{2} \sum_{\substack{i,j \\ i \neq j}} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,i} \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_i|}$$

(2.1.1)

where the indexes I, J run on nuclei, i and j on electrons, $\mathbf{R}_I$ and M_I are positions and masses of the nuclei, $\mathbf{r}_i$ and m_e of the electrons, Z_I the atomic number of nucleus I. The first term is the kinetic energy of the nuclei, the second term is the kinetic energy of the electrons, the third term is the potential energy of nucleus-nucleus Coulomb interaction, the fourth term is the potential energy of electron-electron Coulomb interaction and the last term is the potential energy of nucleus-electron Coulomb interaction. The time-independent Schrödinger equation for the system reads:

$$H_{tot}\Psi(\{\mathbf{R}_I\}, \{\mathbf{r}_i\}) = E\Psi(\{\mathbf{R}_I\}, \{\mathbf{r}_i\})$$

(2.1.2)

where $\Psi(\{\mathbf{R}_I\}, \{\mathbf{r}_i\})$ is the total wavefunction of the system. In principle, everything about the system is known if one can solve the above Schrödinger equation. However, it is impossible to solve it in practice. A so-called Born-Oppenheimer (BO) approximation was made by Born and Oppenheimer[2] in 1927. Since the nuclei are much heavier than electrons (the mass of a proton is about 1836 times the mass of an electron), the nuclei move much slower (about two order of magnitude slower) than the

electrons. Therefore we can separate the movement of nuclei and electrons. When we consider the movement of electrons, it is reasonable to consider the positions of nuclei are fixed, thus the total wavefunction can be written as:

$$\Psi(\{\mathbf{R}_I\}, \{\mathbf{r}_i\}) = \Theta(\{\mathbf{R}_I\})\phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) \tag{2.1.3}$$

where $\Theta(\{\mathbf{R}_I\})$ describes the nuclei and $\phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\})$ the electrons (depending parametrically on the positions of the nuclei). With the BO approximation, Eq. (2.1.2) can be divided into two separate Schrödinger equations:

$$H_e\phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) = V(\{\mathbf{R}_I\})\phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) \tag{2.1.4}$$

where

$$H_e = -\sum_i \frac{\hbar^2}{2m_e}\nabla^2_{\mathbf{r}_i} + \frac{1}{2}\sum_{\substack{I,J \\ I \neq J}} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} + \frac{1}{2}\sum_{\substack{i,j \\ i \neq j}} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,i} \frac{Z_I e^2}{|\mathbf{R}_I - \mathbf{r}_i|} \tag{2.1.5}$$

and

$$[-\sum_I \frac{\hbar^2}{2M_I}\nabla^2_{\mathbf{R}_I} + V(\{\mathbf{R}_I\})]\Theta(\{\mathbf{R}_I\}) = E'\Theta(\{\mathbf{R}_i\}) \tag{2.1.6}$$

Eq. (2.1.4) is the equation for the electronic problem with the nuclei positions fixed. The eigenvalue of the energy $V(\{\mathbf{R}_I\})$ depends parametrically on the positions of the nuclei. After solving Eq. (2.1.4), $V(\{\mathbf{R}_I\})$ is known and by applying it to Eq. (2.1.6), which has no electronic degrees of freedom, the motion of the nuclei is obtained. Eq. (2.1.6) is sometimes replace by a Newton equation, i.e., to move the nuclei classically, using ∇V as the forces. Then the whole problem is solved.

The significance of the BO approximation is to separate the movement of electrons

and nuclei. Now we can consider that the electrons are moving in a static external potential $V_{ext}(\mathbf{r})$ formed by the nuclei, which is the starting point of DFT. The BO approximation was extended by Bohn and Huang known as Born-Huang (BH) approximation [6] to take into account more nonadiabatic effect in the electronic Hamiltonian than in the BO approximation.

2.1.2   Thomas-Fermi-Dirac Approximation

The predecessor to DFT was the Thomas-Fermi (TF) model proposed by Thomas[7] and Fermi[8] in 1927. In this method, they used the electron density $n(\mathbf{r})$ as the basic variable instead of the wavefunction. The total energy of a system in an external potential $V_{ext}(\mathbf{r})$ is written as a functional of the electron density $n(\mathbf{r})$ as:

$$E_{TF}[n(\mathbf{r})] = A_1 \int n(\mathbf{r})^{5/3} d\mathbf{r} + \int n(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (2.1.7)$$

where the first term is the kinetic energy of the non-interacting electrons in a homogeneous electron gas (HEG) with $A_1 = \frac{3}{10}(3\pi^2)^{2/3}$ in atomic units ($\hbar = m_e = e = 4\pi/\epsilon_0 = 1$). The kinetic energy density of a HEG is obtained by adding up all of the free-electron energy state $\varepsilon_k = k^2/2$ up to the Fermi wavevector $k_F = [3\pi^2 n(\mathbf{r})]^{1/3}$ as:

$$\begin{aligned} t_0[n(\mathbf{r})] &= \frac{2}{(2\pi)^3} \int_0^{k_F} \frac{k^2}{2} 4\pi k^2 dk \\ &= A_1 n(\mathbf{r})^{5/3} \end{aligned} \quad (2.1.8)$$

The second term is the classical electrostatic energy of the nucleus-electron Coulomb interaction. The third term is the classical electrostatic Hartree energy approximated by the classical Coulomb repulsion between electrons. In the original TF method, the exchange and correlation among electrons was neglected. In 1930, Dirac[9] extended

the Thomas-Fermi method by adding a local exchange term $A_2 \int n(\mathbf{r})^{4/3} d\mathbf{r}$ to Eq. (2.1.7) with $A_2 = -\frac{3}{4}(3/\pi)^{1/3}$, which leads Eq. (2.1.7) to

$$
\begin{aligned}
E_{TFD}[n(\mathbf{r})] = A_1 \int n(\mathbf{r})^{5/3} d\mathbf{r} + \int n(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} \\
+ \frac{1}{2} \int \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + A_2 \int n(\mathbf{r})^{4/3} d\mathbf{r}
\end{aligned}
\tag{2.1.9}
$$

The ground state density and energy can be obtained by minimizing the Thomas-Fermi-Dirac equation (2.1.9) subject to conservation of the total number (N) of electrons. By using the technique of Lagrange multipliers, the solution can be found in the stationary condition:

$$
\delta\{E_{TFD}[n(\mathbf{r})] - \mu(\int n(\mathbf{r}) d\mathbf{r} - N)\} = 0
\tag{2.1.10}
$$

where μ is a constant known as a Lagrange multiplier, whose physical meaning is the chemical potential (or Fermi energy at T=0 K). Eq. (2.1.10) leads to the Thomas-Fermi-Dirac equation,

$$
\frac{5}{3} A_1 n(\mathbf{r})^{2/3} + V_{ext}(r) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{4}{3} A_2 n(\mathbf{r})^{1/3} - \mu = 0
\tag{2.1.11}
$$

which can be solved directly to obtain the ground state density.

The approximations used in Thomas-Fermi-type approach are so crude that the theory suffers from many problems. The most serious one is that the theory fails to describe bonding between atoms, thus molecules and solids cannot form in this theory.[10] Although it is not good enough to describe electrons in matter, its concept to use electron density as the basic variable illustrates the way DFT works.

2.1.3   The Hohenberg-Kohn (HK) Theorems

DFT was proven to be an exact theory of many-body systems by Hohenberg and Kohn[1] in 1964. It applies not only to condensed-matter systems of electrons with fixed nuclei, but also more generally to any system of interacting particles in an external potential $V_{ext}(\mathbf{r})$. The theory is based upon two theorems.

The HK theorem I:

The ground state particle density $n(\mathbf{r})$ of a system of interacting particles in an external potential $V_{ext}(\mathbf{r})$ uniquely determines the external potential $V_{ext}(\mathbf{r})$, except for a constant. Thus the ground state particle density determines the full Hamiltonian, except for a constant shift of the energy. In principle, all the states including ground and excited states of the many-body wavefunctions can be calculated. This means that **the ground state particle density uniquely determines all properties of the system completely**.

Proof of the HK theorem I:

For simplicity, here I only consider the case that the ground state of the system is nondegenerate. It can be proven that the theorem is also valid for systems with degenerate ground states.[11] The proof is based on minimum energy principle. Suppose there are two different external potentials $V_{ext}(\mathbf{r})$ and $V'_{ext}(\mathbf{r})$ which differ by more than a constant and lead to the same ground state density $n_0(\mathbf{r})$. The two external potentials would give two different Hamiltonians, $\hat{H}$ and $\hat{H}'$, which have the same ground state density $n_0(\mathbf{r})$ but would have different ground state wavefunctions, Ψ and Ψ', with $\hat{H}\Psi = E_0\Psi$ and $\hat{H}'\Psi' = E'_0\Psi'$. Since Ψ' is not the ground state of $\hat{H}$,

it follows that

$$E_0 < \langle \Psi' \mid \hat{H} \mid \Psi' \rangle$$

$$< \langle \Psi' \mid \hat{H}' \mid \Psi' \rangle + \langle \Psi' \mid \hat{H} - \hat{H}' \mid \Psi' \rangle \qquad (2.1.12)$$

$$< E_0' + \int n_0(\mathbf{r})[V_{ext}(\mathbf{r}) - V_{ext}'(\mathbf{r})]d\mathbf{r}$$

Similarly

$$E_0' < \langle \Psi \mid \hat{H}' \mid \Psi \rangle$$

$$< \langle \Psi \mid \hat{H} \mid \Psi \rangle + \langle \Psi \mid \hat{H}' - \hat{H} \mid \Psi \rangle \qquad (2.1.13)$$

$$< E_0 + \int n_0(\mathbf{r})[V_{ext}'(\mathbf{r}) - V_{ext}(\mathbf{r})]d\mathbf{r}$$

Adding Eq. (2.1.12) and (2.1.13) lead to the contradiction

$$E_0 + E_0' < E_0 + E_0' \qquad (2.1.14)$$

Hence, no two different external potentials $V_{ext}(\mathbf{r})$ can give rise to the same ground state density $n_0(\mathbf{r})$, i.e., the ground state density determines the external potential $V_{ext}(\mathbf{r})$, except for a constant. That is to say, there is a one-to-one mapping between the ground state density $n_0(\mathbf{r})$ and the external potential $V_{ext}(\mathbf{r})$, although the exact formula is unknown.

The HK theorem II:

There exists a universal functional $F[n(\mathbf{r})]$ of the density, independent of the external potential $V_{ext}(\mathbf{r})$, such that the global minimum value of the energy functional $E[n(\mathbf{r})] \equiv \int n(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r} + F[n(\mathbf{r})]$ is the exact ground state energy of the system and the exact ground state density $n_0(\mathbf{r})$ minimizes this functional. Thus the exact ground state energy and density are fully determined by the functional $E[n(\mathbf{r})]$.

Proof of the HK theorem II:

The universal functional $F[n(r)]$ can be written as

$$F[n(\mathbf{r})] \equiv T[n(\mathbf{r})] + E_{int}[n(\mathbf{r})] \qquad (2.1.15)$$

where $T[n(\mathbf{r})]$ is the kinetic energy and $E_{int}[n(\mathbf{r})]$ is the interaction energy of the particles. According to variational principle, for any wavefunction Ψ', the energy functional $E[\Psi']$:

$$E[\Psi'] \equiv \langle \Psi' \mid \hat{T} + \hat{V}_{int} + \hat{V}_{ext} \mid \Psi' \rangle \qquad (2.1.16)$$

has its global minimum value only when Ψ' is the ground state wavefunction Ψ_0, with the constraint that the total number of the particles is conserved. According to HK theorem I, Ψ' must correspond to a ground state with particle density $n'(\mathbf{r})$ and external potential $V'_{ext}(\mathbf{r})$, then $E[\Psi']$ is a functional of $n'(\mathbf{r})$. According to variational principle:

$$
\begin{aligned}
E[\Psi'] &\equiv \langle \Psi' \mid \hat{T} + \hat{V}_{int} + \hat{V}_{ext} \mid \Psi' \rangle \\
&= E[n'(\mathbf{r})] \\
&= \int n'(\mathbf{r}) V'_{ext}(\mathbf{r}) d\mathbf{r} + F[n'(\mathbf{r})] \\
&> E[\Psi_0] \\
&= \int n_0(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + F[n_0(\mathbf{r})] \\
&= E[n_0(\mathbf{r})]
\end{aligned}
\qquad (2.1.17)
$$

Thus the energy functional $E[n(\mathbf{r})] \equiv \int n(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + F[n(\mathbf{r})]$ evaluated for the correct ground state density $n_0(\mathbf{r})$ is indeed lower than the value of this functional for any other density $n(\mathbf{r})$. Therefore by minimizing the total energy functional of the system with respect to variations in the density $n(\mathbf{r})$, one would find the exact ground state density and energy.

The HK theorems can be generalized to spin density functional theory with spin degrees of freedom.[12] In this theory, there are two types of densities, namely, the particle density $n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r})$ and the spin density $s(\mathbf{r}) = n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r})$ where $\uparrow$ and $\downarrow$ denote the two different kinds of spins. The energy functional is generalized to $E[n(\mathbf{r}), s(\mathbf{r})]$. In systems with magnetic order or atoms with net spins, the spin density functional theory should be used instead of the original one-spin density functional theory. DFT can also be generalized to include temperature dependence[13] and time dependence known as time-dependent density functional theory (TD-DFT).[14]

Although HK theorems put particle density $n(\mathbf{r})$ as the basic variable, it is still impossible to calculate any property of a system because the universal functional $F[n(\mathbf{r})]$ is unknown. This difficulty was overcome by Kohn and Sham[3] in 1965, who proposed the well known Kohn-Sham ansatz.

2.1.4   The Kohn-Sham (KS) Ansatz

It is the Kohn-Sham (KS) ansatz[3] that puts Hohenberg-Kohn theorems into practical use and makes DFT calculations possible with even a single personal computer. This is part of the reason that DFT became the most popular tool for electronic structure calculations. The KS ansatz was so successful that Kohn was honored the Nobel prize in chemistry in 1998.

The KS ansatz is to replace the original many-body system by an auxiliary independent-particle system and assume that the two systems have exactly the same ground state density. It maps the original interacting system with real potential onto a fictitious non-interacting system whereby the electrons move within an effective Kohn-Sham single-particle potential $V_{KS}(\mathbf{r})$. For the auxiliary independent-particle

system, the auxiliary Hamiltonian is

$$\hat{H}_{KS} = -\frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r}) \tag{2.1.18}$$

in atomic units $\hbar = m_e = e = 4\pi/\epsilon_0 = 1$. For a system with N independent electrons, the ground state is obtained by solving the N one-electron Schrödinger equations,

$$(\frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r}))\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{2.1.19}$$

where there is one electron in each of the N orbitals $\psi_i(\mathbf{r})$ with the lowest eigenvalues ε_i. The density of the auxiliary system is constructed from:

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2 \tag{2.1.20}$$

which is subject to the conservation condition:

$$\int n(\mathbf{r})d\mathbf{r} = N \tag{2.1.21}$$

The non-interacting independent-particle kinetic energy $T_S[n(\mathbf{r})]$ is given by,

$$T_S[n(\mathbf{r})] = -\frac{1}{2}\sum_{i=1}^{N}\int \psi_i^*(\mathbf{r})\nabla^2\psi_i(\mathbf{r})d\mathbf{r} \tag{2.1.22}$$

Then the universal functional $F[n(\mathbf{r})]$ was rewritten as

$$F[n(\mathbf{r})] = T_S[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{XC}[n(\mathbf{r})] \tag{2.1.23}$$

where $E_H[n(\mathbf{r})]$ is the classic electrostatic (Hartree) energy of the electrons,

$$E_H[n(\mathbf{r})] = \frac{1}{2} \int \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \qquad (2.1.24)$$

and $E_{XC}[n(\mathbf{r})]$ is the XC energy, which contains the difference between the exact and non-interacting kinetic energies and also the non-classical contribution to the electron-electron interactions, of which the exchange energy is a part. Since the ground state energy of a many-electron system can be obtained by minimizing the energy functional $E[n(\mathbf{r})] = F[n(\mathbf{r})] + \int n(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r}$, subject to the constraint that the number of electrons N is conserved,

$$\delta\{F[n(\mathbf{r})] + \int n(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r} - \mu(\int n(\mathbf{r})d\mathbf{r} - N)\} = 0 \qquad (2.1.25)$$

and the resulting equation is

$$\begin{aligned}
\mu &= \frac{\delta F[n(\mathbf{r})]}{\delta n(\mathbf{r})} + V_{ext}(\mathbf{r}) \\
&= \frac{\delta T_S[n(\mathbf{r})]}{\delta n(\mathbf{r})} + V_{KS}(\mathbf{r})
\end{aligned} \qquad (2.1.26)$$

where μ is the chemical potential,

$$\begin{aligned}
V_{KS}(\mathbf{r}) &= V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r}) \\
&= V_{ext}(\mathbf{r}) + \frac{\delta E_H[n(\mathbf{r})]}{\delta n(\mathbf{r})} + \frac{\delta E_{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})}
\end{aligned} \qquad (2.1.27)$$

is the KS one-particle potential with the Hartree potential $V_H(\mathbf{r})$

$$
\begin{aligned}
V_H(\mathbf{r}) &= \frac{\delta E_H[n(\mathbf{r})]}{\delta n(\mathbf{r})} \\
&= \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'
\end{aligned}
\tag{2.1.28}
$$

and the XC potential $V_{XC}(\mathbf{r})$

$$
V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})}
\tag{2.1.29}
$$

Equations (2.1.19), (2.1.20), (2.1.27) together are the well-known KS equations, which must be solved self-consistently because $V_{KS}(\mathbf{r})$ depends on the density through the XC potential. In order to calculate the density, the N equations in Eq. (2.1.19) have to be solved in KS theory as opposed to one equation in the TF approach. However an advantage of the KS method is that as the complexity of a system increases, due to N increasing, the problem becomes no more difficult, only the number of single-particle equations to be solved increases.

Although exact in principle, the KS theory is approximate in practice because of the unknown XC energy functional $E_{XC}[n(\mathbf{r})]$. An implicit definition of $E_{XC}[n(\mathbf{r})]$ can be given as

$$
E_{XC}[n(\mathbf{r})] = T[n(\mathbf{r})] - T_S[n(\mathbf{r})] + E_{int}[n(\mathbf{r})] - E_H[n(\mathbf{r})]
\tag{2.1.30}
$$

where $T[n(\mathbf{r})]$ and $E_{int}[n(\mathbf{r})]$ are the exact kinetic and electron-electron interaction energies of the interacting system respectively. It is crucial to have an accurate XC energy functional $E_{XC}[n(\mathbf{r})]$ or potential $V_{XC}(\mathbf{r})$ in order to give a satisfactory description of a realistic condensed-matter system. The most widely used approx-

imations for the XC potential are the local density approximation (LDA) and the generalized-gradient approximation (GGA).

The KS energy eigenvalues of Eq. (2.1.19) are not for the original interacting many-body system and have no physical meaning. They cannot be interpreted as one-electron excitation energies of the interacting many-body system, i.e., they are not the energies to add or subtract from the interacting many-body system, because the total energy of the interacting system is not a sum of all the eigenvalues of occupied states in equation (2.1.19), i.e., $E_{tot} \neq \sum_i^{occ.} \varepsilon_i$. The only exception is the highest eigenvalue in a finite system which is the negative of the ionization energy, -I, because it determines the asymptotic long-range density of the bound system which is assumed to be exact. No other eigenvalue is guaranteed to be correct by the KS theory. Nevertheless, within the KS theory itself, the eigenvalues have a well-defined meaning and they are used to construct physically meaningful quantities. They have a definite mathematical meaning, often known as the Slater-Janak theorem. The eigenvalue is the derivative of the total energy with respect to occupation of a state, i. e.

$$\begin{aligned} \varepsilon_i &= \frac{dE_{total}}{dn_i} \\ &= \int \frac{dE_{total}}{dn(\mathbf{r})} \frac{dn(\mathbf{r})}{dn_i} d\mathbf{r} \end{aligned}$$

(2.1.31)

2.1.5 Local (Spin) Density Approximation (L(S)DA)

The KS ansatz successfully maps the original interacting many-body system onto a set of independent single-particle equations and makes the problem much easier. In the meantime, without knowing the exact form of the XC energy functional $E_{XC}[n(\mathbf{r})]$, the KS equations are unsolvable. Although the exact XC energy functional $E_{XC}[n(\mathbf{r})]$ should be very complicated, simple but successful approximations to it have been

made, which not only predict various properties of many systems reasonably well but also greatly reduce computational costs, leading to the wide use of DFT for electronic structure calculations. Of these approximations, the local density approximation (LDA) is the most widely used one. In LDA, the XC energy per electron at a point $\mathbf{r}$ is considered the same as that for a homogeneous electron gas (HEG) that has the same electron density at the point $\mathbf{r}$. The total exchange-correlation functional $E_{XC}[n(\mathbf{r})]$ can be written as,

$$
\begin{aligned}
E_{XC}^{LDA}[n(\mathbf{r})] &= \int n(\mathbf{r})\epsilon_{XC}^{hom}(n(\mathbf{r}))d\mathbf{r} \\
&= \int n(\mathbf{r})[\epsilon_{X}^{hom}(n(\mathbf{r})) + \epsilon_{C}^{hom}(n(\mathbf{r}))]d\mathbf{r} \\
&= E_{X}^{LDA}[n(\mathbf{r})] + E_{C}^{LDA}[n(\mathbf{r})]
\end{aligned}
\tag{2.1.32}
$$

for spin unpolarized systems and

$$
E_{XC}^{LSDA}[n_{\uparrow}(\mathbf{r}), n_{\downarrow}(\mathbf{r})] = \int n(\mathbf{r})\epsilon_{XC}^{hom}(n_{\uparrow}(\mathbf{r}), n_{\downarrow}(\mathbf{r}))d\mathbf{r}
\tag{2.1.33}
$$

for spin polarized systems[15], where the XC energy density $\epsilon_{XC}^{hom}(n(\mathbf{r}))$ is a function of the density alone, and is decomposed into exchange energy density $\epsilon_{X}^{hom}(n(\mathbf{r}))$ and correlation energy density $\epsilon_{C}^{hom}(n(\mathbf{r}))$ so that the XC energy functional is decomposed into exchange energy functional $E_{X}^{LDA}[n(\mathbf{r})]$ and correlation energy functional $E_{C}^{LDA}[n(\mathbf{r})]$ linearly. Note that $E_{XC}^{LSDA}[n_{\uparrow}(\mathbf{r}), n_{\downarrow}(\mathbf{r})]$ is not written in the way

$$
E_{XC}^{LSDA}[n_{\uparrow}(\mathbf{r}), n_{\downarrow}(\mathbf{r})] = \int [n_{\uparrow}(\mathbf{r})\epsilon_{XC,\uparrow}^{hom}(n_{\uparrow}(\mathbf{r})) + n_{\downarrow}(\mathbf{r})\epsilon_{XC,\downarrow}^{hom}(n_{\downarrow}(\mathbf{r}))]d\mathbf{r}
\tag{2.1.34}
$$

as one may think. The exchange energy functional $E_X^{LDA}[n(\mathbf{r})]$ employs the expression for a HEG by using it pointwise, which is known analytically as[9]

$$
\begin{aligned}
E_X^{LDA}[n(\mathbf{r})] &= \int n(\mathbf{r})\epsilon_X^{hom}(n(\mathbf{r}))d\mathbf{r} \\
&= -\frac{3}{4}(\frac{3}{\pi})^{1/3}\int n(\mathbf{r})^{4/3}d\mathbf{r}
\end{aligned}
\tag{2.1.35}
$$

where

$$
\epsilon_X^{hom}(n(\mathbf{r})) = -\frac{3}{4}(\frac{3}{\pi})^{1/3}n(\mathbf{r})^{1/3}
\tag{2.1.36}
$$

is the exchange energy density of the unpolarized HEG introduced first by Dirac.[9] Analytic expressions for the correlation energy of the HEG are unknown except in the high and low density limits corresponding to infinitely weak and infinitely strong correlations. The expression of the correlation energy density of the HEG at high density limit has the form

$$
\epsilon_C = A\ln(r_s) + B + r_s(C\ln(r_s) + D)
\tag{2.1.37}
$$

and the low density limit takes the form

$$
\epsilon_C = \frac{1}{2}(\frac{g_0}{r_s} + \frac{g_1}{r_s^{3/2}} + \cdots)
\tag{2.1.38}
$$

where the Wigner-Seitz radius r_s is related to the density as

$$
\frac{4}{3}\pi r_s^3 = \frac{1}{n}.
\tag{2.1.39}
$$

In order to obtain accurate values of the correlation energy density at intermediate density, accurate quantum Monte Carlo (QMC) simulations for the energy of the

HEG are needed and have been performed at several intermediate density values.[16] Most local density approximations to the correlation energy density interpolate these accurate values from QMC simulations while reproducing the exactly known limiting behavior. Depending on the analytic forms used for ϵ_C, different local density approximations were proposed including Vosko-Wilk-Nusair[17] (VWM), Perdew-Zunger[18] (PZ81), Cole-Perdew[19] (CP) and Perdew-Wang[20] (PW92).

For spin polarized systems, the exchange energy functional is known exactly from the result of spin-unpolarized functional:

$$E_X[n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})] = \frac{1}{2}(E_X[2n_\uparrow(\mathbf{r})] + E_X[2n_\downarrow(\mathbf{r})]) \qquad (2.1.40)$$

The spin-dependence of the correlation energy density is approached by the relative spin-polarization:

$$\zeta(\mathbf{r}) = \frac{n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r})}{n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r})} \qquad (2.1.41)$$

The spin correlation energy density $\epsilon_C(n(\mathbf{r}), \zeta(\mathbf{r}))$ is so constructed to interpolate extreme values $\zeta = 0, \pm 1$, corresponding to spin-unpolarized and ferromagnetic situations.

The XC potential $V_{XC}(\mathbf{r})$ in LDA is

$$\begin{aligned} V_{XC}^{LDA} &= \frac{\delta E_{XC}^{LDA}}{\delta n(\mathbf{r})} \\ &= \epsilon_{XC}(n(\mathbf{r})) + n(\mathbf{r})\frac{\partial \epsilon_{XC}(n(\mathbf{r}))}{\partial n(\mathbf{r})} \end{aligned} \qquad (2.1.42)$$

Within LDA, the total energy of a system is:

$$E_{tot}[n(\mathbf{r})] = T_S[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{XC}[n(\mathbf{r})] + \int n(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r}$$

$$= \sum_i^{occ.} \langle \psi_i(\mathbf{r}) \mid -\frac{1}{2}\nabla^2 \mid \psi_i(\mathbf{r}) \rangle + E_H[n(\mathbf{r})] + E_{XC}[n(\mathbf{r})] + \int n(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r}$$

$$= \sum_i^{occ.} \langle \psi_i(\mathbf{r}) \mid -\frac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{XC}(\mathbf{r}) + V_{ext}(\mathbf{r}) \mid \psi_i(\mathbf{r}) \rangle$$

$$- \sum_i^{occ.} \langle \psi_i(\mathbf{r}) \mid V_H(\mathbf{r}) \mid \psi_i(\mathbf{r}) \rangle - \sum_i^{occ.} \langle \psi_i(\mathbf{r}) \mid V_{XC}(\mathbf{r}) \mid \psi_i(\mathbf{r}) \rangle$$

$$- \sum_i^{occ.} \langle \psi_i(\mathbf{r}) \mid V_{ext}(\mathbf{r}) \mid \psi_i(\mathbf{r}) \rangle + E_H[n(\mathbf{r})] + E_{XC}[n(\mathbf{r})] + \int n(\mathbf{r})V_{ext}(\mathbf{r})d\mathbf{r}$$

$$= \sum_i^{occ.} \varepsilon_i - \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + \int n(\mathbf{r})(\epsilon_{XC}(\mathbf{r}) - V_{XC}(\mathbf{r}))d\mathbf{r}$$

$$= \sum_i^{occ.} \varepsilon_i - \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' - \int n(\mathbf{r})^2\frac{\partial \epsilon_{XC}(n(\mathbf{r}))}{\partial n(\mathbf{r})}d\mathbf{r}.$$

$$(2.1.43)$$

As mentioned before, $E_{tot} \neq \sum_i^{occ.} \varepsilon_i$.

The LDA is very simple, corrections to the exchange-correlation energy due to the inhomogeneities in the electronic density are ignored. However it is surprisingly successful and even works reasonably well in systems where the electron density is rapidly varying. One reason is that LDA gives the correct sum rule to the exchange-correlation hole. That is, there is a total electronic charge of one electron excluded from the neighborhood of the electron at $\mathbf{r}$. In the meantime, it tends to underestimate atomic ground state energies and ionization energies, while overestimating binding energies. It makes large errors in predicting the energy gaps of some semiconductors. Its success and limitations lead to approximations of the XC energy functional beyond the LDA, through the addition of gradient corrections to incorporate longer range gradient effects (GGA), as well as LDA+U method to account for

the strong correlations of the d electrons in transition elements and f electrons in lanthanides and actinides.

2.1.6 Generalized-Gradient Approximation (GGA)

As mentioned above, the LDA neglects the inhomogeneities of the real charge density which could be very different from the HEG. The XC energy of inhomogeneous charge density can be significantly different from the HEG result. This leads to the development of various generalized-gradient approximations (GGAs) which include density gradient corrections and higher spatial derivatives of the electron density and give better results than LDA in many cases. Three most widely used GGAs are the forms proposed by Becke[21] (B88), Perdew *et al.*[22], and Perdew, Burke and Enzerhof[23] (PBE).

The definition of the XC energy functional of GGA is the generalized form in Eq. (2.1.33) of LSDA to include corrections from density gradient $\nabla n(\mathbf{r})$ as

$$E_{XC}^{GGA}[n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})] = \int n(\mathbf{r})\epsilon_{XC}^{hom}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}), |\nabla n_\uparrow(\mathbf{r})|, |\nabla n_\downarrow(\mathbf{r})|, \cdots)d\mathbf{r}$$

$$= \int n(\mathbf{r})\epsilon_X^{hom}(n(\mathbf{r}))F_{XC}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}), |\nabla n_\uparrow(\mathbf{r})|, |\nabla n_\downarrow(\mathbf{r})|, \cdots)d\mathbf{r}$$

$$(2.1.44)$$

where F_{XC} is dimensionless and $\epsilon_X^{hom}(n(\mathbf{r}))$ is the exchange energy density of the unpolarized HEG as given in Eq. (2.1.36). F_{XC} can be decomposed linearly into exchange contribution F_X and correlation contribution F_C as $F_{XC} = F_X + F_C$. For a detailed treatment of F_X and F_C in different GGAs, please refer to Chapter 8 of the book by Martin.[24]

GGA generally works better than LDA, in predicting bond length and binding energy of molecules, crystal lattice constants, and so on, especially in systems where

the charge density is rapidly varying. However GGA sometimes overcorrects LDA results in ionic crystals where the lattice constants from LDA calculations fit well with experimental data but GGA will overestimate it. Nevertheless, both LDA and GGA perform badly in materials where the electrons tend to be localized and strongly correlated such as transition metal oxides and rare-earth elements and compounds. This drawback leads to approximations beyond LDA and GGA.

2.1.7   LDA+U Method

Strongly correlated systems usually contain transition metal or rare-earth metal ions with partially filled d or f shells. Because of the orbital-independent potentials in L(S)DA and GGA, they cannot properly describe such systems. For example, L(S)DA predicts transition metal oxides to be metallic with itinerant d electrons because of the partially filled d shells. Instead, these transition metal oxides are Mott insulators and the d electrons are well localized. In order to properly describe these strongly correlated systems, orbital-dependent potentials should be used for d and f electrons. There are several approaches available nowadays to incorporate the strong electron-electron correlations between d electrons and f electrons. Of these methods including the self-interaction correction (SIC) method [25], Hartree-Fock (HF) method [26], and GW approximation [27], LDA+U method [28] is the most widely used one.

In the LDA+U method, the electrons are divided into two classes: delocalized s, p electrons which are well described by LDA (GGA) and localized d or f electrons for which an orbital-dependent term $\frac{1}{2}U\sum_{i\neq j}n_i n_j$ should be used to describe Coulomb $d-d$ or $f-f$ interaction, where n_i are $d-$ or $f-$orbital occupancies. The total

energy in L(S)DA+U method is given as[28]:

$$E_{tot}^{LDA+U}[\rho_\sigma(\mathbf{r}), \{n_\sigma\}] = E^{LSDA}[\rho_\sigma(\mathbf{r})] + E^U[\{n_\sigma\}] - E_{dc}[\{n_\sigma\}] \qquad (2.1.45)$$

where σ denotes the spin index, $\rho_\sigma(\mathbf{r})$ is the electron density for spin-σ electrons and $\{n_\sigma\}$ is the density matrix of d or f electrons for spin-σ, the first term is the standard LSDA energy functional, the second term is the electron-electron Coulomb interaction energy given by[28]

$$
\begin{aligned}
E^U[\{n\}] = \frac{1}{2} \sum_{\{m\},\sigma} & \{\langle m, m'' \mid V_{ee} \mid m', m''' \rangle n_{mm',\sigma} n_{m''m''',-\sigma} \\
& - (\langle m, m'' \mid V_{ee} \mid m', m''' \rangle - \langle m, m'' \mid V_{ee} \mid m''', m' \rangle) n_{mm',\sigma} n_{m''m''',\sigma}\}
\end{aligned}
$$
$$(2.1.46)$$

where m denotes the magnetic quantum number, and V_{ee} are the screened Coulomb interactions among the d or f electrons. The last term in Eq. (2.1.45) is the double-counting term which removes an averaged LDA energy contribution of these d or f electrons from the LDA energy. It is given by[28]

$$E_{dc}[\{n_\sigma\}] = \frac{1}{2}UN(N-1) - \frac{1}{2}J[N_\uparrow(N_\uparrow - 1) + N_\downarrow(N_\downarrow - 1)] \qquad (2.1.47)$$

where $N_\sigma = Tr(n_{mm',\sigma})$ and $N = N_\uparrow + N_\downarrow$. U and J are screened Coulomb and exchange parameters.

As a simple approximation, if the exchange and non-sphericity is neglected, Eq. (2.1.45) is simplified to[28]

$$E_{tot}^{LDA+U} = E_{LDA} + \frac{1}{2}U \sum_{i \neq j} n_i n_j - \frac{1}{2}UN(N-1) \qquad (2.1.48)$$

The orbital energies ε_i are derivatives of Eq. (2.1.48) with respect to orbital occupations n_i:

$$\varepsilon_i = \frac{\partial E}{\partial n_i} = \varepsilon_{LDA} + U(\frac{1}{2} - n_i) \qquad (2.1.49)$$

In this simple consideration, the LDA orbital energies are shifted by $-U/2$ for occupied orbitals ($n_i = 1$) and by $+U/2$ for unoccupied orbitals ($n_i = 0$), resulting in lower and upper Hubbard bands separated by U, which opens a gap at the Fermi energy in transition metal oxides.

In the general case, the effective single-particle potential is

$$V_{mm',\sigma} = \frac{\partial(E^U[\{n_\sigma\}] - E_{dc}[\{n_\sigma\}])}{\partial n_{mm',\sigma}}$$

$$= \sum_{\{m\}}\{\langle m, m'' \mid V_{ee} \mid m', m'''\rangle n_{m''m''',-\sigma} - (\langle m, m'' \mid V_{ee} \mid m'm'''\rangle \qquad (2.1.50)$$

$$- \langle m, m'' \mid V_{ee} \mid m''', m'\rangle) n_{m''m''',\sigma}\} - U(N - \frac{1}{2}) + J(N_\sigma - \frac{1}{2})$$

which is used in the effective single-particle Hamiltonian

$$\hat{H} = \hat{H}_{LSDA} + \sum_{mm'} \mid inlm\sigma\rangle V_{mm',\sigma}\langle inlm'\sigma \mid \qquad (2.1.51)$$

where i denotes the site, n the main quantum number, and l the orbital quantum number.

The matrix elements of V_{ee} can be expressed in terms of complex spherical harmonics and effective Slater integrals F_k as[29]

$$\langle m, m'' \mid V_{ee} \mid m', m'''\rangle = \sum_k a_k(m, m', m'', m''')F_k \qquad (2.1.52)$$

where $0 \leq k \leq 2l$ and

$$a_k(m, m', m'', m''') = \frac{4\pi}{2k+1} \sum_{q=-k}^{k} \langle lm \mid Y_{kq} \mid lm' \rangle \langle lm'' \mid Y_{kq}^* \mid lm''' \rangle \qquad (2.1.53)$$

$$F_k \approx \int \int_0^{\infty} dr_1 dr_2 (r_1 R_i(r_1))^2 (r_2 R_i(r_2))^2 \frac{r_<^k}{r_>^{k+1}}, \quad \text{for } k > 0 \qquad (2.1.54)$$

Here, $r_<$ is the smaller of r_1 and r_2 and $r_>$ the larger. The relations between the Slater integrals and the screened Coulomb and exchange parameters U and J are:

$$U = F_0; \; J = (F_2 + F_4)/14, \quad \text{for 3d or 4d systems,}$$

$$U = F_0; \; J = (286F_2 + 195F_4 + 250F_6)/6435, \quad \text{for 4f or 5f systems,} \qquad (2.1.55)$$

The ratio F_4/F_2 and F_6/F_2 are taken from atomic situations. $F_4/F_2 \sim 0.625$ for $3d$ transition elements[30] and $F_4/F_2 \sim 2/3, F_6/F_2 \sim 1/2$ for $4f$ lanthanides. The screened Coulomb parameter U can be calculated from the constraint LDA method[31], so that the LDA+U method remains a first principle method (no adjustable parameters).

For the double-counting term, there are two different treatments: the so-called around mean field (AMF) and fully localized limit (FLL) (also called atomic limit) approaches. The former is more suitable for small U systems[32] and the latter is more suitable for large U systems.[33] The energies for the double counting are given by[34]

$$\begin{aligned} E_{AMF}^{dc} &= \frac{1}{2} \sum_{m \neq m', \sigma\sigma'} [U_{mm'} - \delta_{\sigma,\sigma'} J_{mm'}] \bar{n}\bar{n} \\ &= \frac{1}{2} U N^2 - \frac{U + 2lJ}{2l+1} \frac{1}{2} \sum_{\sigma} N_{\sigma}^2 \end{aligned} \qquad (2.1.56)$$

$$E_{FLL}^{dc} = \frac{1}{2} \sum_{m \neq m', \sigma\sigma'} [U_{mm'} - \delta_{\sigma,\sigma'} J_{mm'}] \bar{n}_\sigma \bar{n}_{\sigma'}$$

$$= \frac{1}{2} U N (N-1) - \frac{1}{2} J \sum_\sigma N_\sigma (N_\sigma - 1) \tag{2.1.57}$$

where $\bar{n} = N/2(2l+1)$ is the average occupation of the correlated orbitals and $\bar{n}_\sigma = N_\sigma/(2l+1)$ is the average occupation of a single spin of the correlated orbitals. Note that, Eq. (2.1.57) is the same as Eq. (2.1.47). For a detailed comparison of the different double counting terms, please refer to [34].

2.1.8 Solving Kohn-Sham Equations

By using independent-particle methods, the KS equations provide a way to obtain the exact density and energy of the ground state of a condensed matter system. The KS equations must be solved consistently because the effective KS potential V_{KS} and the electron density $n(\mathbf{r})$ are closely related. This is usually done numerically through some self-consistent iterations as shown in Fig. 2.1.1. The process starts with an initial electron density, usually a superposition of atomic electron density, then the effective KS potential V_{KS} is calculated and the KS equation is solved with single-particle eigenvalues and wavefunctions, a new electron density is then calculated from the wavefunctions. After this, self-consistent condition(s) is checked. Self-consistent condition(s) can be the change of total energy or electron density from the previous iteration or total force acting on atoms is less than some chosen small quantity, or a combination of these individual conditions. If the self-consistency is not achieved, the calculated electron density will be mixed with electron density from previous iterations to get a new electron density. A new iteration will start with the new electron density. This process continues until self-consistency is reached. After the

self-consistency is reached, various quantities can be calculated including total energy, forces, stress, eigenvalues, electron density of states, band structure, etc..

The most timing consuming step in the whole process is to solve KS equation with a given KS potential V_{KS}. There are several different schemes to the calculation of the independent-particle electronic states in solids where boundary conditions are applied. They are basically classified into three types[24]:

1.Plane waves.

In this method, the wavefunctions (eigenfunctions of the KS equations) are expanded in a complete set of plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$ and the external potential of nuclei are replaced by pseudopotentials which include effects from core electrons. Such pseudopotentials have to satisfy certain conditions. Most widely used pseudopotentials nowadays include norm-conserving pseudopotentials[35] (NCPPs) and ultrasoft pseudopotentials[36] (USPPs). In norm-conserving pseudopotentials, five requirements should be satisfied:

a. the pseudo valence eigenvalues should agree with all-electron valence eigenvalues for the chosen atomic reference configuration;

b. the pseudo valence wavefunctions should match all-electron valence wavefunctions beyond a chosen core radius R_c;

c. the logarithmic derivatives of the pseudo and the all-electron wavefunctions should agree at R_c,

d. the integrated charge inside R_c for each wavefunction agrees (norm-conservation); and

e. the first energy derivative of the logarithmic derivatives of the all-electron and pseudo wavefunctions agree at R_c, and therefore for all $r \leq R_c$.

In ultrasoft pseudopotentials, the norm-conservation condition is not required so

that the pseudo wavefunctions are much softer than pseudo wavefunctions in norm-conserving pseudopotentials. As a result, it significantly reduces the number of plane waves needed to expand the wavefunctions (smaller energy cutoff for wavefunctions).

Plane waves have played an important role in the early orthogonalized plane wave[37, 38, 39] (OPW) calculations and are generalized to modern projector augmented wave[40, 41, 42] (PAW) method. Because of the simplicity of plane waves and pseudopotentials, computational load is significantly reduced in these methods and therefore it is most suitable for calculations of large systems. In this method, forces can be easily calculated and it can be easily developed to quantum molecular dynamics simulations as well as response to (small) external perturbations. However, results from plane wave methods using pseudopotentials are usually less accurate than results from all-electron full potential methods. And great care should be taken when one generates a pseudopotential and it should be tested to match results from all-electron calculations. The most widely used codes using plane waves and pseudopotentials are plane wave self-consistent field (now known as Quantum ESPRESSO)[43] (PWscf), ABINIT[44], VASP[45] (which uses PAW method too).

2. Localized atomic(-like) orbitals.

The most well-known methods in this category are linear combination of atomic orbitals[46] (LCAO), also called tight-binding[46] (TB) and full potential non-orthogonal local orbital[47] (FPLO). The basic idea of these methods is to use atomic orbitals as the basis set to expand the one-electron wavefunction in KS equations.

In FPLO, in addition to the spherical average of the crystal potential, a so-called confining potential $V_{con} = (r/r_0)^m$ is used to compress the long range tail of the local orbitals (wave functions), where m is the confining potential exponent with a typical value of 4, $r_0 = (x_0 r_{NN}/2)^{3/2}$ is a compression parameter with x_0 being

a dimensionless parameter and r_{NN} the nearest neighbor distance. Therefore, the atomic-like potential is written as

$$V_{at}(r) = -(1/4\pi) \int V(\mathbf{r} - \mathbf{R} - \tau)d^3\mathbf{r} + V_{con}(r), \qquad (2.1.58)$$

where the first term is the spherical average of the crystal potential mentioned above. For systems containing atom(s) with partially filled $4f$ and $5f$ shells, the confining potential exponent m needs to be increased to 5 or 6. In practice, the dimensionless parameter x_0 is taken as a variational parameter in the self-consistent procedure.

3. Atomic sphere methods.

Methods in the class can be considered as a combination of plane wave method and localized atomic orbitals. It uses localized atomic orbital presentation near the nuclei and plane waves in the interstitial region. The most widely used methods are (full potential) linear muffin-tin orbital[48] (LMTO) as implemented in LMTART[49] by Dr. Savrasov and (full potential) linear augment plane wave[48, 50] (LAPW) as implemented in WIEN2K[51].

2.2 Linear Response Calculations and Superconductivity

2.2.1 Lattice Dynamics and Phonons

To calculate the lattice dynamical properties, we have linear response method[52] and density functional perturbation theory (DFPT) [53], which are closely related. In both methods, it is essential to calculate the second-order perturbation of DFT total energy, i.e., $\delta^2 E$, in the framework of density functional theory. The perturbation
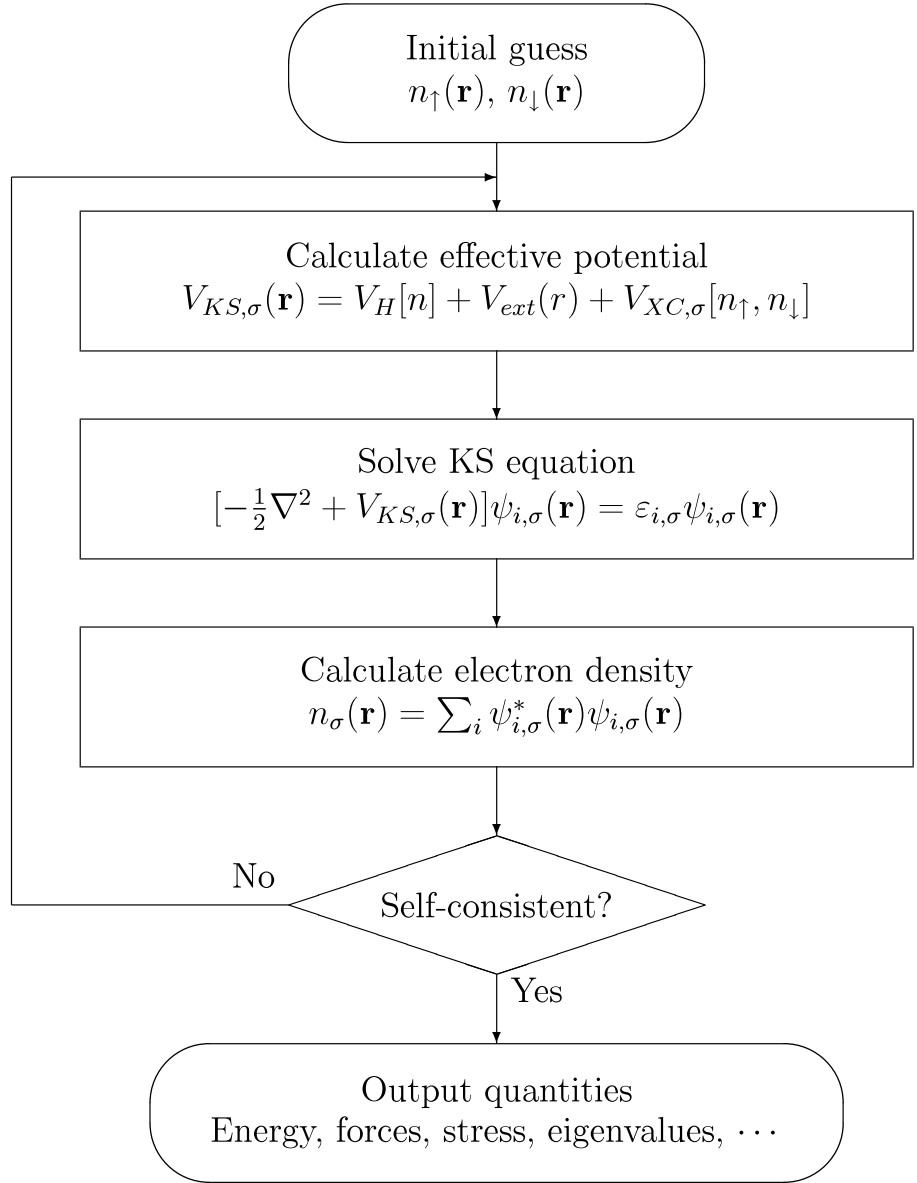
Figure 2.1.1: Flowchart of self-consistency loop for solving KS equations.

is induced by small displacements $\delta\mathbf{R}$ of the nuclei from their equilibrium positions, which result in changes in the external potential V_{ext}, the wave functions Ψ of the KS equations and hence the electron charge density. $\delta^2 E$ is obtained by expanding the DFT total energy with respect to the changes in the wave functions to first order and external potentials up to second order. Detailed expressions can be found in Ref[52, 53].

Phonon spectra can be obtained by first calculating the dynamical matrix

$$D_{ij}(\mathbf{q}) = \sum_{\mathbf{R}'} e^{-i\mathbf{q}\cdot\mathbf{R}'} \frac{\partial^2 E}{\partial u_i(\mathbf{R}+\mathbf{R}')\partial u_j(\mathbf{R})} \tag{2.2.1}$$

with respect to the atomic displacements $\mathbf{u}(\mathbf{R})$ for each atom in each direction, (i, j=1, 2, 3, corresponding to x, y and z directions), and then by solving the equation

$$D(\mathbf{q})\vec{\varepsilon} = M\omega_{\mathbf{q}}^2\vec{\varepsilon} \tag{2.2.2}$$

which gives the phonon frequencies $\omega_{\mathbf{q}}$ of the phonons with wave vector $\mathbf{q}$, where M is a diagonal matrix with the atomic masses on the diagonal.

2.2.2 Electron-Phonon Interaction and T_c

Electron-phonon interaction plays a crucial role in conventional superconductivity. It provides the attractive interaction between electrons needed in BCS theory in order to form Cooper pairs. The physical picture is that the first electron attracts its nearby positive ions (to form phonons) which polarizes its nearby environment leading to a domain with positive charges, which in turn attracts the second electron, resulting in an effective attractive interaction between electrons. A net attractive interaction between electrons can be obtained if the above attraction is strong enough to override

the repulsive screened Coulomb interaction.

In the framework of density functional linear-response method, the electron-phonon (EP) interaction can be calculated by evaluating the EP matrix element[54]

$$g^{\mathbf{q},\nu}_{\mathbf{k}+\mathbf{q}j',\mathbf{k}j} = \langle \mathbf{k} + \mathbf{q}j' | \delta^{\mathbf{q}\nu} V_{eff} | \mathbf{k}j \rangle \qquad (2.2.3)$$

which is the interaction between electronic potential and phonon mode $\omega_{\mathbf{q}\nu}$, or in other words, it is the probability of scattering from the one-electron state $|\mathbf{k}j\rangle$ to the state $|\mathbf{k} + \mathbf{q}j'\rangle$ via the phonon $\omega_{\mathbf{q}\nu}$, where $\delta^{\mathbf{q}\nu} V_{eff}$ is the change in the effective potential induced by the presence of a phonon mode $\omega_{\mathbf{q}\nu}$, and $\psi_{\mathbf{k}j}$ and $\psi_{\mathbf{k}+\mathbf{q}j'}$ have the Fermi energy ε_F.

The phonon linewidth $\gamma_{\mathbf{q}\nu}$ can be evaluated as

$$\gamma_{\mathbf{q}\nu} = 2\pi \omega_{\mathbf{q}\nu} \sum_{\mathbf{k}jj'} |g^{\mathbf{q}\nu}_{\mathbf{k}+\mathbf{q}j',\mathbf{k}j}|^2 \delta(\varepsilon_{\mathbf{k}j} - \varepsilon_F)\delta(\varepsilon_{\mathbf{k}+\mathbf{q}j'} - \varepsilon_F). \qquad (2.2.4)$$

The electron-phonon spectral distribution functions $\alpha^2 F(\omega)$ can be written in terms of $\gamma_{\mathbf{q}\nu}$,

$$\alpha^2 F(\omega) = \frac{1}{2\pi N(\varepsilon_F)} \sum_{\mathbf{q}\nu} \frac{\gamma_{\mathbf{q}\nu}}{\omega_{\mathbf{q}\nu}} \delta(\omega - \omega_{\mathbf{q}\nu}), \qquad (2.2.5)$$

where $N(\varepsilon_F) = \sum_{\mathbf{k}} \delta(\varepsilon_{\mathbf{k}} - \varepsilon_F)$ is the electronic density of states (DOS) per atom per spin at the Fermi level, $F(\omega) = \sum_{\mathbf{q}} \delta(\omega - \omega_{\mathbf{q}})$ is the phonon density of states.

The electron-phonon coupling parameter λ is given by

$$\lambda = 2 \int \frac{\alpha^2 F(\omega)}{\omega} d\omega, \qquad (2.2.6)$$

and the average phonon frequency $\langle \omega^n \rangle$ is given by

$$\langle \omega^n \rangle = \frac{2}{\lambda} \int \omega^{n-1} \alpha^2 F(\omega) d\omega. \tag{2.2.7}$$

Alternatively, λ can be obtained from the "mode" λ

$$\begin{aligned} \lambda_{\mathbf{q}\nu} &= \frac{\gamma_{\mathbf{q}\nu}}{\pi \omega_{\mathbf{q}\nu}^2 N(\varepsilon_F)} \\ &= \frac{2}{\omega_{\mathbf{q}\nu} N(\varepsilon_F)} \sum_{\mathbf{k}jj'} |g_{\mathbf{k}+\mathbf{q}j',\mathbf{k}j}^{\mathbf{q},\nu}|^2 \delta(\varepsilon_{\mathbf{k}j} - \varepsilon_F)\delta(\varepsilon_{\mathbf{k}+\mathbf{q}j'} - \varepsilon_F). \end{aligned} \tag{2.2.8}$$

by adding up the mode λ

$$\lambda = \sum_{\mathbf{q}\nu} \lambda_{\mathbf{q}\nu}. \tag{2.2.9}$$

The parameter λ plays the role of the BCS parameter $N(\varepsilon_F)V_{ph}$ and is the most important single number characterizing electron-phonon coupling.[55] In BCS theory, the superconducting critical temperature T_c is given by

$$T_c = 1.13\langle \omega \rangle e^{-1/N(\varepsilon_F)V_{ph}}, \tag{2.2.10}$$

where V_{ph} is an effective interaction parameter to simplify the complicated net attractive interaction of electrons near Fermi energy, while here it is estimated by the McMillan equation[56]

$$T_c = \frac{\langle \omega \rangle}{1.20} e^{-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}} \tag{2.2.11}$$

where μ^* is the screened Coulomb pseudopotential. In the strong coupling limit ($\lambda > 1$), the above McMillan equation is generalized to Allen-Dynes equation[57]

$$T_c = \frac{\omega_{log}}{1.20} f_1 f_2 e^{-\frac{1.04(1+\lambda)}{\lambda - \mu^*(1+0.62\lambda)}} \tag{2.2.12}$$

where ω_{log} is the average phonon frequency $\langle\omega^n\rangle^{1/n}$ with $n \to 0$, and

$$f_1 = (1 + (\lambda/\Lambda_1)^{3/2})^{1/3}, \quad f_2 = 1 + \frac{(\omega_2/\omega_{log} - 1)\lambda^2}{\lambda^2 + \Lambda_2^2} \tag{2.2.13}$$

contain the strong-coupling corrections which are important for $\lambda > 1$, with $\Lambda_1 = 2.46(1 + 3.8\mu^*)$, $\Lambda_2 = 1.82(1 + 6.3\mu^*)(\omega_2/\omega_{log})$, where ω_2 being the average phonon frequency $\langle\omega^n\rangle^{1/n}$ with $n = 2$.

2.2.3   Nesting Function

Note that $\lambda_{\mathbf{q}\nu}$, or $\gamma_{\mathbf{q}\nu}$, incorporates a phase space factor, the "nesting function" [58] $\xi(\mathbf{q})$ describing the phase space that is available for electron-hole scattering across the Fermi surface($\varepsilon_F = 0$),

$$\xi(\mathbf{q}) = \frac{1}{N}\sum_{\mathbf{k}}\delta(\varepsilon_{\mathbf{k}})\delta(\varepsilon_{\mathbf{k}+\mathbf{q}}) \propto \oint_{\mathcal{L}}\frac{d\mathcal{L}_{\mathbf{k}}}{|\vec{v}_{\mathbf{k}} \times \vec{v}_{\mathbf{k}+\mathbf{q}}|}. \tag{2.2.14}$$

Here $\mathcal{L}$ is the line of intersection of an undisplaced Fermi surface and one displaced by $\mathbf{q}$, and $\vec{v}_{\mathbf{k}}$ is the electron velocity at $\mathbf{k}$. These equations presume the adiabatic limit, in which the phonon frequencies are small compared to any electronic energy scale. This limit applies to elemental Y and Ca, which will be discussed in details in chapter 4.

In the case of the free electron limit, $\varepsilon_{\mathbf{k}} = k^2/2$, in two dimension, the "Fermi surface" is a circle with a radius of k_F (or a cylinder in 3D), the nesting function is

$$\begin{aligned}\xi(\mathbf{q}) &\propto \frac{1}{|\mathbf{q}|\sqrt{4k_F^2 - \mathbf{q}^2}}, \quad if \ |\mathbf{q}| < 2k_F; \\ &= 0, \qquad\qquad\qquad if \ |\mathbf{q}| > 2k_F,\end{aligned} \tag{2.2.15}$$

while in 3D, the Fermi surface is a sphere of radius k_F, the nesting function is then

$$\xi(\mathbf{q}) \propto \frac{\pi}{2|\mathbf{q}|}, \quad if \; |\mathbf{q}| < 2k_F;$$
$$= 0, \quad if \; |\mathbf{q}| > 2k_F. \tag{2.2.16}$$

In real materials, the Fermi surfaces are usually very complicated and the nesting function needs to be calculated numerically.

2.3 Tight Binding Method and Wannier Functions

2.3.1 Local Orbitals and Tight Binding

Tight binding method provides a simple way to calculate electronic band structure and ground state energy, by expanding the wave function $\psi(\mathbf{r})$ in terms of atomic orbitals $\phi_n(\mathbf{r} - \mathbf{R})$ of isolated atoms at each atomic site, or in terms of other local orbitals (eg: Wannier functions). This method is sometimes also regarded as linear combination of atomic orbital (LCAO) approach and applies to non-crystalline materials (eg: molecules) and crystalline materials, although the latter is more common where the atoms are located on a periodic lattice. This approach is valid in systems where the electrons are more localized than itinerant, i.e., the electrons are bound to each atom instead of moving through the crystal. Therefore it applies to insulators and some semiconductors but certainly not simple metals. Recently, the tight binding method became a basic tool in the study of strongly correlated systems where the $3d$ and $4f$ electrons are highly localized.

In tight-binding model, the total Hamiltonian $H(\mathbf{r})$ of the crystal is a sum of the atomic Hamiltonians $H_{at}(\mathbf{r} - \mathbf{R} - \vec{\tau}_i)$ located at each atomic site plus an interaction

term $\delta U(\mathbf{r})$, which is considered as a small perturbation,

$$H(\mathbf{r}) = \sum_{\mathbf{R}} H_{at}(\mathbf{r} - \mathbf{R} - \vec{\tau}_i) + \delta U(\mathbf{r}). \tag{2.3.1}$$

where $\vec{\tau}_i$ is the atomic position to the origin of the cell at $\mathbf{R}$.

A basis state $\varphi_{ni,\mathbf{k}}(\mathbf{r})$ with wave vector $\mathbf{k}$ (restricted to the first Brillouin zone) can be constructed from the atomic orbitals $\phi_n(\mathbf{r} - \mathbf{R} - \vec{\tau}_i)$, (which satisfy $H_{at}(\mathbf{r})\phi_n(\mathbf{r}) = E_n\phi_n(\mathbf{r})$), according to Bloch theorem

$$\varphi_{ni,\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\vec{\tau}_i)} \phi_n(\mathbf{r} - \mathbf{R} - \vec{\tau}_i). \tag{2.3.2}$$

The crystal wave function $\psi_{\mathbf{k}}(\mathbf{r})$ is then constructed from the above basis functions

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{ni} b_{ni}(\mathbf{k})\varphi_{ni,\mathbf{k}}(\mathbf{r}), \tag{2.3.3}$$

where $b_{ni}(\mathbf{k})$ are coefficients depending on $\mathbf{k}$. It is easy to show that $\psi_{\mathbf{k}}(\mathbf{r})$ satisfies the Bloch theorem,

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}\psi_{\mathbf{k}}(\mathbf{r}). \tag{2.3.4}$$

From the Schrödinger equation,

$$H\psi_{\mathbf{k}} = \varepsilon_{\mathbf{k}}\psi_{\mathbf{k}}, \tag{2.3.5}$$

we have

$$\sum_{nj} H_{mi,nj}(\mathbf{k})b_{nj}(\mathbf{k}) = \varepsilon_{\mathbf{k}} \sum_{nj} S_{mi,nj}(\mathbf{k})b_{nj}(\mathbf{k}), \tag{2.3.6}$$

where

$$
\begin{aligned}
H_{mi,nj}(\mathbf{k}) &= \int \varphi_{mi,\mathbf{k}}^*(\mathbf{r}) H \varphi_{nj,\mathbf{k}}(\mathbf{r}) d\mathbf{r} \\
&= \frac{1}{N} e^{i\mathbf{k}\cdot(\vec{\tau}_j - \vec{\tau}_i)} \sum_{\mathbf{R}_1,\mathbf{R}_2} e^{i\mathbf{k}\cdot(\mathbf{R}_2 - \mathbf{R}_1)} \int \phi_m^*(\mathbf{r} - \mathbf{R}_1 - \vec{\tau}_i) H \phi_n(\mathbf{r} - \mathbf{R}_2 - \vec{\tau}_j) d\mathbf{r} \\
&= \frac{1}{N} e^{i\mathbf{k}\cdot(\vec{\tau}_j - \vec{\tau}_i)} \sum_{\mathbf{R}_1,\mathbf{R}_2} e^{i\mathbf{k}\cdot(\mathbf{R}_2 - \mathbf{R}_1)} H_{mi,nj}(\mathbf{R}_2 + \vec{\tau}_j - \mathbf{R}_1 - \vec{\tau}_i) \\
&= e^{i\mathbf{k}\cdot(\vec{\tau}_j - \vec{\tau}_i)} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} H_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i)
\end{aligned}
$$

$$(2.3.7)$$

and

$$
\begin{aligned}
S_{mi,nj}(\mathbf{k}) &= \int \varphi_{mi,\mathbf{k}}^*(\mathbf{r}) \varphi_{nj,\mathbf{k}}(\mathbf{r}) d\mathbf{r} \\
&= \frac{1}{N} e^{i\mathbf{k}\cdot(\vec{\tau}_j - \vec{\tau}_i)} \sum_{\mathbf{R}_1,\mathbf{R}_2} e^{i\mathbf{k}\cdot(\mathbf{R}_2 - \mathbf{R}_1)} \int \phi_m^*(\mathbf{r} - \mathbf{R}_1 - \vec{\tau}_i) \phi_n(\mathbf{r} - \mathbf{R}_2 - \vec{\tau}_j) d\mathbf{r} \\
&= \frac{1}{N} e^{i\mathbf{k}\cdot(\vec{\tau}_j - \vec{\tau}_i)} \sum_{\mathbf{R}_1,\mathbf{R}_2} e^{i\mathbf{k}\cdot(\mathbf{R}_2 - \mathbf{R}_1)} S_{mi,nj}(\mathbf{R}_2 + \vec{\tau}_j - \mathbf{R}_1 - \vec{\tau}_i) \\
&= e^{i\mathbf{k}\cdot(\vec{\tau}_j - \vec{\tau}_i)} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} S_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i)
\end{aligned}
$$

$$(2.3.8)$$

which are usually called the Hamiltonian matrix and the overlap matrix. $H_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i)$ is the amplitude that an electron in the orbital ϕ_n at site $(\mathbf{R} + \vec{\tau}_j)$ will hop to the orbital ϕ_m at position $\vec{\tau}_i$ of origin under the action of the Hamiltonian $\hat{H}$, and is usually denoted as hopping parameter

$$
t_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i) \equiv H_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i), \tag{2.3.9}
$$

where the on-site $(\mathbf{R} = 0, \vec{\tau}_j - \vec{\tau}_i = 0)$ term is

$$
t_{mi,ni}(\mathbf{0}) = \varepsilon_n \delta_{m,n}. \tag{2.3.10}
$$

Similarly $S_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i)$ is called the overlap of $\phi_m(\mathbf{r} - \vec{\tau}_i)$ and $\phi_n(\mathbf{r} - \mathbf{R} - \vec{\tau}_j)$, denoted as overlap matrix

$$s_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i) \equiv S_{mi,nj}(\mathbf{R} + \vec{\tau}_j - \vec{\tau}_i); \; s_{mi,ni}(\mathbf{0}) = \delta_{m,n}. \tag{2.3.11}$$

A general Hamiltonian matrix element involves three center integrals –contributions to the Hamiltonian $\hat{H}$ from a third atom and the two sites upon which the orbitals are centered. But in practice, the three center contributions are neglected as in the two-center approximation introduced by Slater and Koster.[46]

In the simplest case with only a single atomic s-level on a lattice with only one atom per primitive cell, $m = n = s$, $i = j = 1$ and $\vec{\tau}_i = \vec{\tau}_j = 0$ are imposed on the above equations. Eq. (2.3.6) is simplified to

$$H_{s,s}(\mathbf{k})b_s(\mathbf{k}) = \varepsilon_{\mathbf{k}} S_{s,s}(\mathbf{k})b_s(\mathbf{k}). \tag{2.3.12}$$

The solution to the energy dispersion $\varepsilon_{\mathbf{k}}$ is then

$$\begin{aligned}
\varepsilon_{\mathbf{k}} &= \frac{H_{s,s}(\mathbf{k})}{S_{s,s}(\mathbf{k})} \\
&= \frac{\sum_{\mathbf{R}} t(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}}{\sum_{\mathbf{R}} s(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}} \\
&= \frac{\varepsilon_s + \sum_{\mathbf{R}\neq 0} t(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}}{1 + \sum_{\mathbf{R}\neq 0} s(\mathbf{R})e^{i\mathbf{k}\cdot\mathbf{R}}}
\end{aligned} \tag{2.3.13}$$

2.3.2 Wannier Functions

As mentioned above, in tight binding method, the Bloch wave functions can also be expanded using other local orbitals instead of atomic orbitals. Wannier functions, first proposed by G. Wannier[59], is a candidate of such local orbitals, although they

are not localized in some cases when the bandwidths are large, where the Wannier functions are not like the atomic wave functions at all. Technically, Wannier functions are Fourier transformations of Bloch wave functions $\psi_{n\mathbf{k}}(\mathbf{r})$. Since $\psi_{n\mathbf{k}}(\mathbf{r})$ is periodic in the reciprocal lattice, i.e., $\psi_{n\mathbf{k}+\mathbf{G}}(\mathbf{r}) = \psi_{n\mathbf{k}}(\mathbf{r})$, where G is a reciprocal lattice vector, $\psi_{n\mathbf{k}}(\mathbf{r})$ can be expanded in plane waves as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} w_n(\mathbf{r} - \mathbf{R})e^{i\mathbf{R}\cdot\mathbf{k}}, \qquad (2.3.14)$$

where the coefficients $w_n(\mathbf{r} - \mathbf{R})$ are Wannier functions, which depend only on $\mathbf{r} - \mathbf{R}$ instead of $\mathbf{r}$ and $\mathbf{R}$ independently due to the Bloch theorem.

The Wannier functions $w_n(\mathbf{r} - \mathbf{R})$ can be obtained by inverse transformations

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{\Omega_{cell}}{(2\pi)^3} \int_{BZ} e^{-i\mathbf{R}\cdot\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r})d\mathbf{k}, \qquad (2.3.15)$$

where Ω_{cell} is the volume of the real-space primitive cell of the crystal.

The Wannier functions so obtained are not unique because any Bloch function $\psi_{n\mathbf{k}}(\mathbf{r})$ doesn't change any physically meaningful quantity under a "gauge transformation"

$$\psi_{n\mathbf{k}}(\mathbf{r}) \rightarrow \tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) = e^{i\phi_n(\mathbf{k})}\psi_{n\mathbf{k}}(\mathbf{r}). \qquad (2.3.16)$$

A more general construction of the Wannier functions is given by

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{\Omega_{cell}}{(2\pi)^3} \int_{BZ} e^{i\phi_n(\mathbf{k})}e^{-i\mathbf{R}\cdot\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r})d\mathbf{k}, \qquad (2.3.17)$$

The non-uniqueness of the Wannier functions is totally due to presence of the phase factor $\phi_n(\mathbf{k})$.

In addition to the freedom in the choice of phase factor $\phi_n(\mathbf{k})$, there is also a degree

of freedom associated with the choice of a full unitary matrix $U_{nm}^{\mathbf{k}}$, which transforms the N Bloch wave functions $\psi_{n\mathbf{k}}(\mathbf{r})$ between themselves at every wavevector k, but leaves the electronic energy functional (in an insulator) invariant. This leads to the most general construction of Wannier functions from Bloch wave functions $\psi_{n\mathbf{k}}(\mathbf{r})$ in the form

$$w_n(\mathbf{r} - \mathbf{R}) = \frac{\Omega_{cell}}{(2\pi)^3} \int_{BZ} \sum_{m=1}^{N} U_{nm}^{\mathbf{k}} e^{-i\mathbf{R}\cdot\mathbf{k}} \psi_{m\mathbf{k}}(\mathbf{r}) d\mathbf{k}, \qquad (2.3.18)$$

where $U^{\mathbf{k}}$ is a $M \times N$ unitary matrix with $M \leq N$. Note that $U^{\mathbf{k}}$ is not necessarily a square matrix, as one can use this procedure to construct M Wannier functions out of N bands. Again, in the procedure, the choice of $U^{\mathbf{k}}$ is not unique. Actually, one can use this freedom to construct Wannier functions with properties of one's own interest, such as the most symmetric, or maximally projected, or maximally localized. A widely used one is the maximally localized Wannier functions proposed by Vanderbilt and coworkers[60, 61], in which the quantity

$$\Omega = \sum_{n=1}^{N} (\langle \mathbf{r}^2 \rangle_n - \langle \mathbf{r} \rangle_n^2) \qquad (2.3.19)$$

is minimized by choosing appropriate $U^{\mathbf{k}}$, where $\langle \cdots \rangle_n$ is the expectation value over the n-th Wannier function in the unit cell. There are also other Wannier functions in use which are constructed by using projections onto local orbitals to emphasize symmetries.[62, 63]

The Wannier functions $w_n(\mathbf{r} - \mathbf{R})$ for all band n and $\mathbf{R}$ form a complete orthogonal set. That is to say, the Wannier functions are orthogonal at different site and/or different band,

$$\int w_m(\mathbf{r} - \mathbf{R}_i) w_n(\mathbf{r} - \mathbf{R}_j) = \delta_{m,n} \delta_{i,j}. \qquad (2.3.20)$$

In the tight binding approach, if Wannier functions are used as the local orbitals, the overlap matrix $s_{m,n}(\mathbf{R})$ is greatly simplified to

$$s_{m,n}(\mathbf{R}) = \delta_{m,n}\delta_{\mathbf{0},\mathbf{R}}. \tag{2.3.21}$$

As a result,

$$\begin{aligned} S_{m,n}(\mathbf{k}) &= \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} s_{m,n}(\mathbf{R}) \\ &= \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \delta_{m,n}\delta_{\mathbf{0},\mathbf{R}} \\ &= \delta_{m,n} \end{aligned} \tag{2.3.22}$$

and Eq. (2.3.6) is simplified to

$$\sum_{n} H_{m,n}(\mathbf{k}) b_n(\mathbf{k}) = \varepsilon_{\mathbf{k}} b_m(\mathbf{k}). \tag{2.3.23}$$

This is a main advantage to use Wannier functions in tight binding method.

In practice, Wannier functions are constructed from the results of DFT calculations and used as the local orbitals in tight binding method. The hopping parameters ε_n (on-site energy) and $t_{m,n}(\mathbf{R})$ are obtained by fitting the $\varepsilon_n(\mathbf{k})$ (from the tight binding method) to the band structures of DFT calculations. (Note that the overlap matrix is the identity matrix due to the use of Wannier functions as local orbitals.) The hopping parameters are then used to construct model Hamiltonians to study many-body effects. As mentioned above, the Wannier functions are not localized if the bands have large bandwidths. Therefore, the above procedure works better for systems with an isolated set of narrow bands.