UNIVERSITY OF THE PHILIPPINES

FIRST PRINCIPLE CALCULATIONS OF DEFECT STRUCTURES IN ZINC OXIDE

by

CHRISTIAN LOER T. LLEMIT

An undergraduate thesis submitted in partial fulfillment of
the requirements for the degree of

BACHELOR OF SCIENCE IN APPLIED PHYSICS

NATIONAL INSTITUTE OF PHYSICS
University of the Philippines - Diliman

JUNE 2020

This thesis entitled, **FIRST PRINCIPLE CALCULATIONS OF DEFECT STRUCTURES IN ZINC OXIDE**, prepared and submitted by **CHRISTIAN LOER T. LLEMIT**, in partial fulfillment of the requirements for the degree of **Bachelor of Science**, major in **Applied Physics** is hereby accepted.

Roland V. Sarmago, Ph.D., Adviser

Richard Feynman, Ph.D.

Albert Einstein, Ph.D.

ACKNOWLEDGMENT

This thesis would not have been possible without the support of Computing and Archiving Research Environment (COARE) facility under DOST Advanced Science and Technology Institue (ASTI), for the high performance computing needed to run cpu-intensive simulations.

FIRST PRINCIPLE CALCULATIONS OF DEFECT STRUCTURES

IN ZINC OXIDE

Abstract

by Christian Loer T. Llemit, BS
University of the Philippines - Diliman
June 2020

Nullam mollis et leo at pharetra. Nulla efficitur molestie euismod. Sed dapibus metus sed tempus varius. Aenean finibus eros ut urna luctus feugiat. Duis turpis risus, viverra vitae porta et, ullamcorper ac est. Proin in eros nec ipsum interdum tempus. Nam fringilla lectus velit, non posuere ex vehicula ut. Mauris tincidunt, dolor sit amet commodo tempor, erat mi egestas dui, at elementum tellus est rhoncus libero. Ut et rutrum lectus, id viverra tortor. Vivamus nec lacus eros. Donec dictum porta nisi et vestibulum. Mauris luctus ligula ut libero aliquet luctus. Quisque malesuada egestas finibus.

Mauris dictum pharetra fermentum. Maecenas ut felis varius, dapibus sapien imperdiet, dictum dui. Proin feugiat viverra metus non laoreet. Integer pulvinar mi id lacus semper commodo. Praesent vel erat interdum purus scelerisque maximus. Sed enim risus, mollis blandit ligula ac, sagittis venenatis augue. Mauris nisi purus, gravida ac aliquam eu, ullamcorper eget nulla. Proin id finibus purus. Vestibulum leo ante, porta in quam sed, eleifend feugiat arcu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Dedication

This dissertation/thesis is dedicated to my mother and father who provided both emotional and financial support

Chapter One

Theoretical Framework

1.1 Electronic Structure

The problem of electronic structure methods begins with the attempt to solve the general non-relativistic time-independent Schrödinger equation given as [1]

$$\hat{\mathcal{H}}\Psi = E\Psi \tag{1.1}$$

where $\hat{\mathcal{H}}$ is the Hamiltonian operator for a system of electrons, Ψ is the electronic wave-function and E is the energy of the system. Consider a single electron in three dimensional system, the Schrödinger equation can be expressed as

$$\hat{\mathcal{H}}\Psi_n = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\Psi_n + V\Psi_n = \epsilon_n\Psi_n \tag{1.2}$$

where m is the mass of electron, V is the effective potential energy and ϵ_n is the energy of electron in the orbital. The term orbital denotes the solution of the Schrödinger equation for a system of only one electron. This will be useful in later sections because this will allow to distinguish between the exact quantum state of a system of N interacting electrons from the approximate quantum state of N electrons in N orbitals, where each orbital is a solution to one-electron wavefunction in (1.2). If V is zero for the case of free electrons (i.e. non-interacting), then the orbital model is exact.

Since electrons are restricted by the potential inside the atom, the simplest way of solving (1.2) is by considering an infinite potential well. The electrons are confined inside a cube of length L where the potential V inside is zero and infinite at outside must satisfy the boundary condition

$$\Psi_n(L_x, L_y, L_z) = 0 \tag{1.3}$$

where L_x, L_y, L_z can be either 0 or L. The solution will have a sine dependence

$$\Psi_n(x, y, z) = \sqrt{\left(\frac{2}{L}\right)^3} \sin\left(\frac{n_x\pi}{L}x\right) \sin\left(\frac{n_y\pi}{L}y\right) \sin\left(\frac{n_z\pi}{L}z\right) \tag{1.4}$$

where n_x, n_y, n_z are integer quantum states. Provided that $k_i = n_i\pi/L$ where $i = x, y, \text{or } z$; then the energy dispersion relation can be expressed as

$$\epsilon_k = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2) = \frac{\hbar^2}{2m}k^2 \propto k^2 \tag{1.5}$$

Note that energy levels are discretized by the quantum states which arises from imposing the boundary conditions.

1.1.1 Electronic Band structure

Inside the crystal lattice, the periodic arrangement of atoms or ions causes the potential to be periodic which eventually gives rise to the formation of energy bands. The wavefunction Ψ will become periodic in space with a period L and must obey the Born-von Karman boundary condition [2]

$$\Psi_k(x, y, z) = \Psi_k(x + L, y, z) \tag{1.6}$$

and similarly for the y and z coordinates. It can be shown that wavefunctions satisfying (1.2) and (1.6) are the Bloch form of a travelling plane wave

$$\Psi_k(\vec{\mathbf{r}}) = u_k(\vec{\mathbf{r}}) \exp\left(i\vec{\mathbf{k}} \cdot \vec{\mathbf{r}}\right) \tag{1.7}$$

where $u_k(\vec{\mathbf{r}})$ has the period of the crystal lattice with $u_k(\vec{\mathbf{r}}) = u_k(\vec{\mathbf{r}} + \vec{\mathbf{R}})$. Here $\vec{\mathbf{R}}$ is the translation vector which can be simply thought as the periodicity expressed as vector. The Bloch expression can be written as

$$\begin{aligned}
\Psi_k(\vec{\mathbf{r}} + \vec{\mathbf{R}}) &= u_k(\vec{\mathbf{r}} + \vec{\mathbf{R}}) \exp\Big(i\vec{\mathbf{k}} \cdot (\vec{\mathbf{r}} + \vec{\mathbf{R}})\Big) \\
\Psi_k(\vec{\mathbf{r}} + \vec{\mathbf{R}}) &= u_k(\vec{\mathbf{r}}) \exp\Big(i\vec{\mathbf{k}} \cdot \vec{\mathbf{r}}\Big) \exp\Big(i\vec{\mathbf{k}} \cdot \vec{\mathbf{R}}\Big) \\
\Psi_k(\vec{\mathbf{r}} + \vec{\mathbf{R}}) &= \Psi_k(\vec{\mathbf{r}}) \exp\Big(i\vec{\mathbf{k}} \cdot \vec{\mathbf{R}}\Big)
\end{aligned} \tag{1.8}$$

Notice that the wavefunction differs from the plane wave of free electrons only by a periodic modulation given by the new phase factor. This means that the electrons in the crystal lattice are treated as perturbed weakly by the periodic potential of the ion cores.

1.1.1.1 Band structure of free electron

A special case of periodicity is where the potential is set to zero, which is applicable for the free electrons. The wavefunction will be a plane wave

$$\Psi_k(\vec{\mathbf{r}}) = \exp\Big(i\vec{\mathbf{k}} \cdot \vec{\mathbf{r}}\Big) \tag{1.9}$$

that represents travelling wave with a momentum $\vec{\mathbf{p}} = \hbar\vec{\mathbf{k}}$. The energy dispersion relation is still given by (1.5) but this time the allowed energy values are distributed essentially from zero to infinity. Figure 1.1 shows the parabolic dependence of energy with the wavevector k. Since the system is periodic in real space, it must be true for the reciprocal space, in this case by $2\pi/a$ where a is some lattice constant. Figure 1.1a shows the extended zone scheme where there are no restrictions on the values of wavevector $\vec{\mathbf{k}}$. When wavevectors are outside the first Brillouin zone (BZ), they can be translated back to the first zone by subtracting a suitable reciprocal lattice vector. In mathematical sense [3]

$$\vec{\mathbf{k}} + \vec{\mathbf{G}} = \vec{\mathbf{k}'} \tag{1.10}$$

where $\vec{\mathbf{k}}'$ is the unrestricted wavevector, $\vec{\mathbf{k}}$ is in the first Brillouin zone, and $\vec{\mathbf{G}}$ is the translational reciprocal lattice vector. The energy dispersion relation can always be written as

$$\begin{aligned}\epsilon(k_x, k_y, k_z) &= \frac{\hbar^2}{2m}(\vec{\mathbf{k}} + \vec{\mathbf{G}})^2 \\ &= \frac{\hbar^2}{2m}[(k_x + G_x)^2 + (k_y + G_y)^2 + (k_z + G_z)^2] \end{aligned} \tag{1.11}$$

Figure 1.1b shows the reduced zone scheme where the bands are folded into the first BZ by applying (1.10). Any energy state beyond the first BZ is the same to a state inside the first BZ with a different band index n.

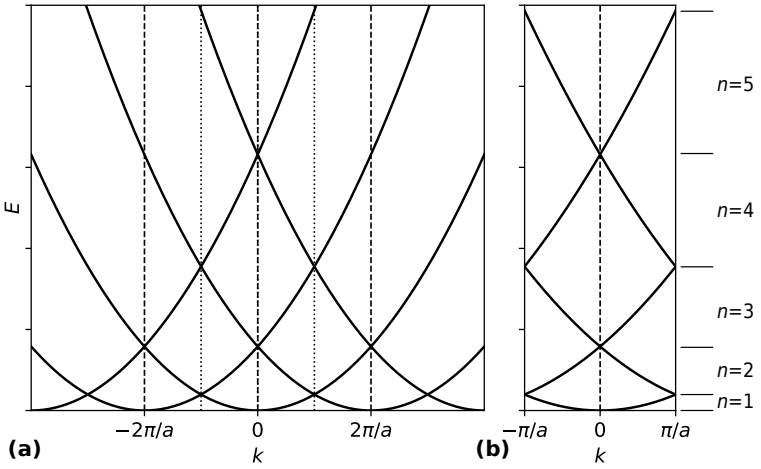

Figure 1.1 Free electron band structure where (a) is in the extended zone scheme and (b) in the reduced zone scheme. The dotted lines in (a) lies the first BZ.

1.1.1.2 Band structure of electrons in solids

When atoms are very far from each other with no interaction, each electron occupies specific discrete orbitals such as 1s, 2p, 3d, etc. When they are bring closer enough, the outermost (valence) electrons interact with each other and will result in the energy level splitting. The innermost (core) electrons remain as they are, since they are closer to the nuclei and bounded by a deep potential well. For a solid containing a large N atoms, there will be N orbitals (i.e. N 3d-orbitals) trying to occupy the same energy level. Pauli's exclusion principle will prevent this from happening, hence what happens is there will be splitting of the energy

level that are closely spaced and this will eventually form a continuous band of energy levels. Figure 1.2 summarizes the evolution of energy levels as the atoms are brought together.

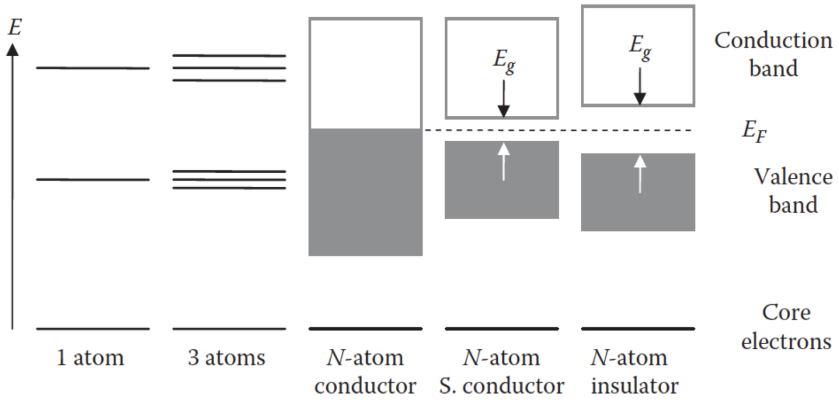

Figure 1.2 Formation of bands and band gaps when isolated atoms are bring closer together. Figure taken from [4]

Another interesting property of band structure is the formation of energy band gaps. This happens when the valence electrons interact with the periodic potential of the nuclei. Assuming a weak periodic potential, most of the band structure will not changed very much, except possibly at the Brillouin zone boundaries with a wavevector of $\vec{\mathbf{k}} = n\pi/a$. The orbitals with the wavevector at zone boundaries, chosen to be at high symmetry points, follows the Bragg diffraction condition and thus are diffracted. The valence electrons are scattered (or reflected) at the zone boundary in which the wavefunction are made up of equal plane waves travelling from the left and from the right. The wavefunction becomes a standing wave that resembles more of those bound states. Hence, there will be a forbidden region where travelling waves are not allowed. If sufficient energy is provided to the electron, they can overcome the binding potential.

The band gap is generally referred to the energy difference between the top of valence band, Valence band maximum (VBM), and the bottom of the conduction band, Conduction band minimum (CBM). If VBM and CBM coincides with each other, the material is said to be a conductor. Electrons can easily occupy the conduction band without any excitation, hence electrons are highly mobile that will lead to high current. For band gaps with a

value comparable to the quantity k_BT, where k_B is the Boltzmann constant and T is the absolute temperature near room temperature, then the material is semiconductor. If band gap is much larger than k_BT, then the material is insulator. However, this criterion is very loose because there are materials with large band gaps such as ZnO, $SrIn_2O_4$, and diamond that are categorized as semiconductors. These materials are generally called wide-band gap semiconductors. If the VBM and CBM are located in the same wavevector k, then the gap is direct. Otherwise, it is indirect.

1.1.2 Density of States

Another useful quantity in describing the electronic structure is the density of states (DOS). In general, the density of states can be defined as [5]

$$D(\epsilon) = 2\sum_n \sum_k \delta(\epsilon - \epsilon_n(k)) \tag{1.12}$$

where for each band index n, the sum is over all allowed values of k lying inside the first Brillouin zone. The factor 2 comes from the allowed values of the spin quantum number for each allowed value of k. In the limit of large crystal, the k points are very close together, and the sum can be replaced by an integral. Since each allowed states will take up a volume of $(\Delta k)^3 = \pi^3/V$ where V is the volume of the solid in real space, it is convenient to write (1.12) as

$$D(\epsilon) = 2\,\frac{V}{\pi^3}\sum_n \sum_k \delta(\epsilon - \epsilon_n(k))(\Delta k)^3 \tag{1.13}$$

for in the limit of $V \to \infty$, $\Delta k \to 0$, so that

$$g(\epsilon) = \lim_{V\to\infty} \frac{1}{V} D(\epsilon) = \frac{2}{\pi^3}\sum_n \int \delta(\epsilon - \epsilon_n(k))\,\mathrm{d}^3k \tag{1.14}$$

Usually, the total DOS is set to be the number of states per unit energy per unit volume.

The DOS can be projected in terms of the orbital contribution of each atoms. This can be expanded in a complete orthonormal basis as [6]

$$D(\epsilon) = \sum_i D_i(\epsilon) \tag{1.15}$$

$$= \sum_i \sum_n \int \langle \psi_n | \, \alpha \, | \psi_n \rangle \, \delta(\epsilon - \epsilon_n(k)) \, \mathrm{d}^3 k \tag{1.16}$$

where $D_i(\epsilon)$ is the projected density of states (PDOS) of orbital i with state α.

1.2 Many-body Physics

Despite the simplicity of Schrödinger equation in (1.1), solving it is a formidable task when dealing with many-electron systems. Analytical solutions to this equation only exist for the very simplest systems (i.e. hydrogenic atoms). Solving beyond '2 particle' system (electron and nucleus) is already intractable. In addition, solid state systems typically contains more than hundreds of particles, resulting in hundreds of simultaneous equations. Even the use of computational methods relies on a number of approximations just to make computations feasible enough. Hence, this section will discuss various levels of approximations without neglecting the parameter-free of first-principles calculations.

1.2.1 Many-particle Hamiltonian Operator

The exact many-particle Hamiltonian is consist of five operators which can be expressed as

$$\hat{\mathcal{H}} = \hat{\mathcal{T}}_n + \hat{\mathcal{T}}_e + \hat{\mathcal{V}}_{en} + \hat{\mathcal{V}}_{ee} + \hat{\mathcal{V}}_{nn} \tag{1.17}$$

where the $\hat{\mathcal{T}}$ and $\hat{\mathcal{V}}$ refer to kinetic energy and potential energy, respectively, and the labels e and n denotes the electronic and nuclear coordinates and their derivatives, respectively. This equation can be expanded as

$$\begin{aligned} \hat{\mathcal{H}} = & -\frac{\hbar^2}{2} \sum_I \frac{\nabla^2_{\vec{\boldsymbol{R}}_I}}{M_I} - \frac{\hbar^2}{2} \sum_i \frac{\nabla^2_{\vec{\boldsymbol{r}}_i}}{m_e} \\ & - \frac{1}{4\pi\epsilon_0} \sum_{I,i} \frac{e^2 Z_I}{\left| \vec{\boldsymbol{R}}_I - \vec{\boldsymbol{r}}_i \right|} + \frac{1}{8\pi\epsilon_0} \sum_{i \neq j} \frac{e^2}{|\vec{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_j|} + \frac{1}{8\pi\epsilon_0} \sum_{I \neq J} \frac{e^2 Z_I Z_J}{\left| \vec{\boldsymbol{R}}_I - \vec{\boldsymbol{R}}_J \right|} \end{aligned} \tag{1.18}$$

where M_I is the mass of the Ith nuclei (or usually ions) with charge Z_I located at site $\vec{\boldsymbol{R}_I}$, and electrons have mass m_e located at site $\vec{\boldsymbol{r}_i}$. The first and second terms are the kinetic energy of the atomic nuclei and electrons, respectively. The last three terms describe the Coulomb interaction between electrons and nuclei, between electrons and other electrons, and between nuclei and other nuclei.

1.2.2 Simplifying Assumptions

Solving (1.18) exactly is very impractical and not worth the effort. Hence, we resort to approximations in order to find acceptable eigenstates.

The first level of approximation is the Born-Oppenheimer approximation or the Adiabatic approximation [7]. It begins with the observation that the mass of nuclei is much larger compared to the electron, as such one can assume that electrons moving in a potential much faster than the nuclei and that the nuclei can be treated as fixed or 'frozen' with respect to motion. As a consequence, the nuclear kinetic energy will be zero and the nuclear interaction with the electron cloud can be treated as an external parameter. Hence, the first term in (1.18) will vanish and the last term reduces to a constant which can be neglected. The third term will become the external potential. The Hamiltonian reduces to

$$\hat{\mathcal{H}} = \hat{\mathcal{T}} + \hat{\mathcal{V}} + \hat{\mathcal{V}}_{ext} \tag{1.19}$$

and using Hartree atomic units $\hbar = m_e = e = 4\pi/\epsilon_0 = 1$ for simplicity

$$\hat{\mathcal{H}} = -\frac{1}{2}\sum_i \nabla^2_{\vec{\boldsymbol{r}_i}} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\vec{\boldsymbol{r}_i} - \vec{\boldsymbol{r}_j}|} + \sum_i V_I(\left|\vec{\boldsymbol{R}_I} - \vec{\boldsymbol{r}_j}\right|) \tag{1.20}$$

1.2.3 Hartree Method

Since the second term in (1.20) includes electron-electron interaction which is difficult to evaluate, Hartree (1928) had proposed a simplified model where he treated each electrons to be independent and interacts with others in an averaged way [8].This implies that each

electron does not recognize others as single entities but rather as a mean Coulomb field. The second term will be replaced by Hartree energy given as

$$\hat{\mathcal{V}}_H = \frac{1}{2} \iint \frac{\rho(\vec{\boldsymbol{r}})\rho(\vec{\boldsymbol{r}}')}{|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'|} \, \mathrm{d}^3 r \, \mathrm{d}^3 r' \tag{1.21}$$

where $\rho(\vec{\boldsymbol{r}})$ is the electron density. The total energy will be sum of N numbers of one-electron energies

$$E = E_1 + E_2 + \cdots + E_N \tag{1.22}$$

then, the N-electron wavefunction can be approximated as a product of one-electron wavefunctions

$$\Psi = \Psi_1 \times \Psi_2 \times \cdots \times \Psi_N \tag{1.23}$$

Hartree model successfully predicts the ground-state energy of Hydrogen atom to be around -13.6 eV. However, for other systems, Hartree model produced crude estimations because it does not take into account the quantum mechanical effects such as antisymmetry principle and the Pauli's exclusion principle. Moreover, the model does not include the exchange and correlation energies of every interacting electrons in the actual systems.

1.2.4 Hartree-Fock Method

Due to the limitations of Hartree Model, Fock (1930) has taken into account the antisymmetric property of electron wavefunctions [9]. Pauli's exclusion principle posits that no two fermions can occupy the same quantum state because the wavefunction is antisymmetric upon particle exchange [10]. The many-electron wavefunction will be expressed in terms of Slater determinant [11]

$$\boldsymbol{\Psi} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Psi_1(\vec{\boldsymbol{r_1}}) & \Psi_2(\vec{\boldsymbol{r_1}}) & \cdots & \Psi_N(\vec{\boldsymbol{r_1}}) \\ \Psi_1(\vec{\boldsymbol{r_2}}) & \Psi_2(\vec{\boldsymbol{r_2}}) & \cdots & \Psi_N(\vec{\boldsymbol{r_2}}) \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_1(\vec{\boldsymbol{r_N}}) & \Psi_2(\vec{\boldsymbol{r_N}}) & \cdots & \Psi_N(\vec{\boldsymbol{r_N}}) \end{vmatrix} \tag{1.24}$$

Using the Slater determinant form of the wavefunction, the Hamiltonian can be written as before with the addition of exchange term

$$\hat{\mathcal{H}}_{HF} = \hat{\mathcal{T}} + \hat{\mathcal{V}}_{ext} + \hat{\mathcal{V}}_H + \hat{\mathcal{V}}_x \tag{1.25}$$

where

$$\hat{\mathcal{V}}_x = -\sum_j \int \frac{\psi_j^*(\vec{\boldsymbol{r}}')\psi(\vec{\boldsymbol{r}}')}{|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'|} \frac{\psi_j(\vec{\boldsymbol{r}})}{\psi(\vec{\boldsymbol{r}})} \, \mathrm{d}\vec{\boldsymbol{r}} \tag{1.26}$$

$\hat{\mathcal{V}}_H$ comes from the Hartree approximation of electron-elecron interaction and $\hat{\mathcal{V}}_x$ comes from the antisymmetric nature of wave function.

1.3 Density Functional Theory (DFT)

Density Functional Theory reframes the problem of calculating electronic properties in terms of the ground state electron density instead of the traditional electronic wavefunctions [12]. The incredible success of DFT in predicting ground state properties have led to widespread applications in materials modelling research.

1.3.1 Hohenberg-Kohn (HK) Formalism

The modern formulations of DFT started in the seminal work of Hohenberg and Kohn in 1964 [13]. Hohenberg and Kohn have shown that the ground state properties can be written as unique functional of the ground state electron density. This statement has large implication because the problem of solving $3n$-dimensional equation simultaneously can be replaced by n separate three-dimensional equations with the use of electron density, $\rho(x, y, z)$.

1.3.1.1 First HK Theorem

The first theorem shows that electron density is a unique functional of the external potential. It states that there is a one-to-one correspondence between the ground state density $\rho_0(r)$

of a many-electron system and the external potential V_{ext}, to within an additive constant. Alternatively, it is impossible to have two external potentials, $V_{ext}(r)$ and $V'_{ext}(r)$, acting on an electron whose difference is not a constant, that give rise to the same ground state electron density, $\rho_0(r)$. That is,

$$\rho(r) = \rho'(r) \qquad \Longleftrightarrow \qquad V'_{ext}(r) - V_{ext}(r) = \text{constant} \tag{1.27}$$

If the external potential is known beforehand, then the ground state electron density can be obtained and vice versa. As the ground state electron density uniquely determines the Hamiltonian of the system, it follows that all measurable properties of the system can be expressed as a functional of the electron density.

1.3.1.2 Second HK Theorem

The second theorem proves the existence of the energy as a functional of the electron density. It states that there exists a universal functional for the energy $E[\rho]$ such that for any given $V_{ext}(r)$, the exact ground-state energy is the global minimum of this functional, and the ground-state density $\rho_0(r)$ is the density $\rho(r)$ that minimizes the functional. Note that the total energy in HK formulation gives an exact form and not approximate ones. The form of the energy functional can be expressed as

$$E_{HK}[\rho(r)] = \langle\psi|\,\hat{\mathcal{T}} + \hat{\mathcal{V}} + \hat{\mathcal{V}}_{ext}\,|\psi\rangle \tag{1.28}$$

$$= \langle\psi|\,\hat{\mathcal{T}} + \hat{\mathcal{V}}\,|\psi\rangle + \langle\psi|\,\hat{\mathcal{V}}_{ext}\,|\psi\rangle \tag{1.29}$$

$$= F[\rho(r)] + \int V_{ext}(r)\rho(r)\,\mathrm{d}^3r \tag{1.30}$$

where $F[\rho(r)]$ is the unknown functional that includes all internal energies, kinetic, and potential, that are independent of the external potential. The HK theorems only asserts the existence of energy functional but it does not provide a practical solution on solving the energy functional.

1.3.2 Kohn Sham (KS) Formulation

Kohn and Sham (1965) introduced an artificial system of non-interacting electrons with the same ground state electron density as the many-body Schrödinger equation [14]. Instead of using the fully interacting multi-electron wavefunctions, the KS formulation resorts to single-particle wavefunctions for solving the many-body problem. The Kohn-Sham Hamiltonian is just an extension of Hartree-Fock Hamiltonian described in (1.25). However, it was implicitly assumed that $\hat{\mathcal{T}}$ is the kinetic energy operator of non-interacting electrons. This assumption neglects the correlation of the interacting system, hence a correction factor must be added. The kinetic energy of the real interacting system can be rewritten as

$$\hat{\mathcal{T}} = \hat{\mathcal{T}}_{KS} + \hat{\mathcal{V}}_c \tag{1.31}$$

where $\hat{\mathcal{T}}_{KS}$ is kinetic energy of the non-interacting electron, and $\hat{\mathcal{V}}_c$ is the correlation energy that measures how much movement of one electron is influenced by the presence of other electrons. The total KS Hamiltonian has the form

$$\begin{aligned} \hat{\mathcal{H}}_{KS} &= (\hat{\mathcal{T}}_{KS} + \hat{\mathcal{V}}_c) + \hat{\mathcal{V}}_{ext} + \hat{\mathcal{V}}_H + \hat{\mathcal{V}}_x \\ &= \hat{\mathcal{T}}_{KS} + \hat{\mathcal{V}}_{ext} + \hat{\mathcal{V}}_H + \hat{\mathcal{V}}_{xc} \end{aligned} \tag{1.32}$$

where $\hat{\mathcal{V}}_{xc} = \hat{\mathcal{V}}_x + \hat{\mathcal{V}}_c$ is the combined exchange-correlation energy. It is instructive to see that the difference between Hartree Hamiltonian from Hartree-Fock Hamiltonian gives the exchange term while the difference between Hartree-Fock Hamiltonian and Kohn-Sham Hamiltonian gives the correlation term [15]. The theorem of Kohn and Sham can be formally formulated as follows:

The exact ground state density $\rho(\vec{\boldsymbol{r}})$ of an N-electron system is

$$\rho(\vec{\boldsymbol{r}}) = \sum_{i=1}^{N} \phi_i(\vec{\boldsymbol{r}})^* \phi_i(\vec{\boldsymbol{r}}) \tag{1.33}$$

where the single-particle KS orbitals $\phi_i(\vec{\boldsymbol{r}})$ are the N lowest energy solutions of the Kohn-Sham equation

$$\hat{\mathcal{H}}_{KS}\, \phi_i(\vec{\boldsymbol{r}}) = \epsilon_i\, \phi_i(\vec{\boldsymbol{r}}) \tag{1.34}$$

1.3.3 Self Consistent Field Calculation

In order to solve the KS equation (1.34), the Hamiltonian $\hat{\mathcal{H}}_{KS}$ must be known beforehand. However, the Hamiltonian depends entirely on the electron density $\rho(\vec{\boldsymbol{r}})$ that can only be solved from single-particle KS orbital $\phi_i(\vec{\boldsymbol{r}})$ given in (1.33). The orbital $\phi_i(\vec{\boldsymbol{r}})$ are in turn calculated from the KS equation and the cycle continues on. This infinite loop is visualized in Figure 1.3

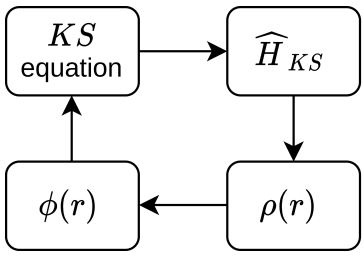

Figure 1.3 Solving Kohn Sham equation leads to a circular argument

To circumvent this, an iterative scheme was developed in which a trial electron density is introduced and the KS equation is iteratively solved to achieve convergence. This iterative process is often referred as Self Consistent Field (SCF) calculation [16]. Specific steps are illustrated in Figure 1.4. First, an initial trial electron density is provided. The trial electron density is usually derived from the superposition of known atomic potentials. Second, the KS equation is solved using the trial electron density. The resulting eigenfunction, in this case the orbital $\phi_i(\vec{\boldsymbol{r}})$, will then be used to calculate the new electron density. The new electron density is compared to the previous electron density and if the error is less than some acceptable deviation, then this will be the ground state density. Otherwise, the electron density is updated and the iteration is repeated kth times until convergence is achieved. Factors that affect the rate of convergence will be discussed on the next chapter.

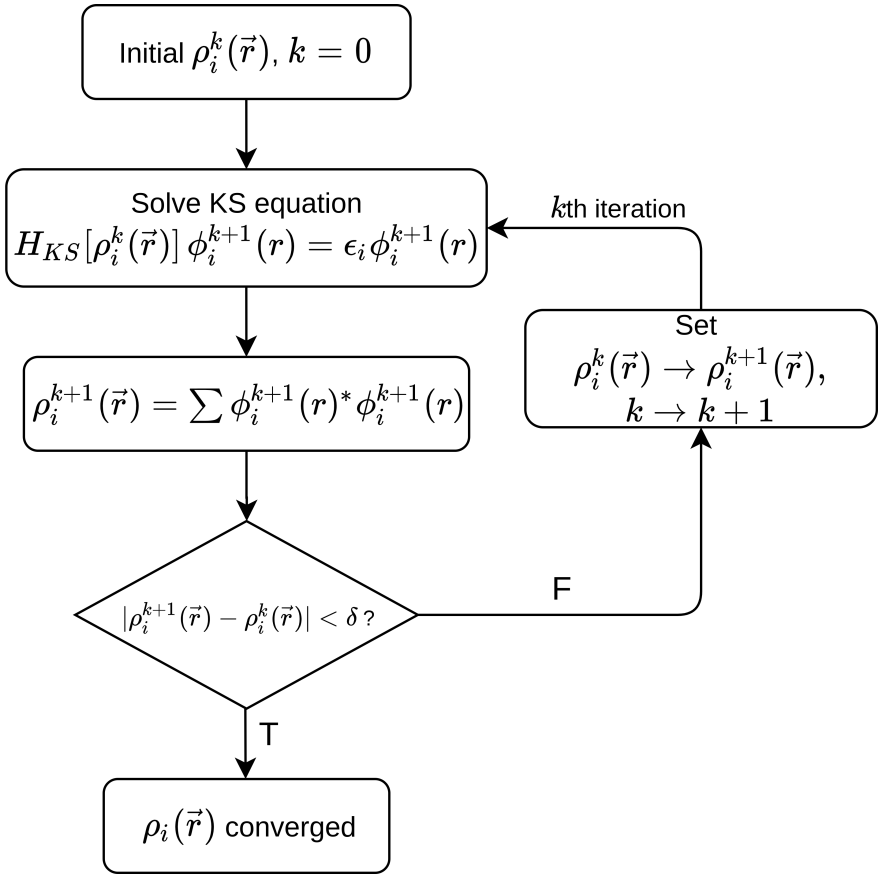

Figure 1.4 Convergence of electron density and other observable quantities using Self Consistent Field calculation

1.4 Exchange-Correlation Functional

So far, no analytical form for the exchange-correlation functional has been found yet that perfectly describes any interacting system [17–19]. The success of DFT depends on the improvement and refinement of the exchange-correlation functional and how it enables to predict many observable properties. Hence, the search for the universal functional is a hot topic of ongoing research. The choice of XC functional varies from different applications of DFT. Thus, there is no one particular functional in the literature which universally performs better than others across all applications.

1.4.1 Local Density Approximation (LDA)

The simplest commonly used exchange-correlation functional is the so called Local Density Approximation (LDA). LDA assumes that the electronic contribution to the exchange-correlation energy from each point in space is the same as to what it would be for a homogeneous electron gas with the uniform density throughout the whole system. This approximation was originally introduced by Kohn and Sham, and holds for a slowly varying density [14]. Using the approximation, the XC energy functional is given by

$$E_{xc}^{\mathrm{LDA}}[\rho] = \int \rho(\vec{\boldsymbol{r}})\, \epsilon_{xc}[\rho(\vec{\boldsymbol{r}})]\, \mathrm{d}^3 r \tag{1.35}$$

where $\epsilon_{xc}[\rho(\vec{\boldsymbol{r}})]$ is the exchange-correlation energy per particle of a uniform electron gas of density $\rho(\vec{\boldsymbol{r}})$. The quantity $\epsilon_{xc}[\rho(\vec{\boldsymbol{r}})]$ can be further split into exchange and corelation contributions

$$\epsilon_{xc}[\rho(\vec{\boldsymbol{r}})] = \epsilon_X[\rho(\vec{\boldsymbol{r}})] + \epsilon_C[\rho(\vec{\boldsymbol{r}})] \tag{1.36}$$

The exchange part was expressed analytically by Dirac [20]

$$\epsilon_x[\rho(\vec{\boldsymbol{r}})] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \rho(\vec{\boldsymbol{r}}) \tag{1.37}$$

while the correlation part has been found numerically by Ceperley and Alder [21] using a stochastic quantum Monte Carlo method [22]. Later, an accurate parametrization of this data was published Perdew and Zunger (LDA-PZ) which is still used in DFT calculations [23]. LDA was expected to be best for solids with slowly varying densities like a nearly-free-electron metals and worst for inhomogeneous systems such as atoms where the density must go continuously to zero just outside the atom. The partial success of LDA in inhomogeneous systems is due to systematic error cancellation in which the correlation is underestimated but the exchange is overestimated resulting to a good value of E_{xc}^{LDA} [24, 25]. However, LDA tends to overestimate cohesive energies and binding energies for metals and insulators

[26–28]. Errors in LDA are severely exaggerated for weakly bonded systems such as van der Waals and H-bond systems [29–31]. Nevertheless, LDA is fairly accurate in predicting elastic properties, such as bulk modulus [32, 33].

1.4.2 Generalized Gradient Approximation (GGA)

Attempts to improve the shortcomings of LDA has led to the use of gradient corrections. These so called Generalized Gradient Approximations (GGA) systematically calculate gradient corrections of the form $|\boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})|$, $|\boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})|^2$, $|\nabla^2\rho(\vec{\boldsymbol{r}})|$, etc. to the LDA. Such functionals can be generalized as

$$E_{xc}^{\mathrm{GGA}}[\rho] = \int f^{GGA}[\rho(\vec{\boldsymbol{r}}), \boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})]\, \mathrm{d}^3r \tag{1.38}$$

where f^{GGA} is some arbitrary function of electron density and its gradient. GGA functionals are often term as semi-local because of their $\boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})$ dependence. Because of the flexibility in choosing f^{GGA}, a plethora of functionals have been developed and depending on the system under study, various results can be obtained. A more specific form of the GGA functional can be written as [27]

$$E_{xc}^{\mathrm{GGA}}[\rho] = \int \rho(\vec{\boldsymbol{r}})\, \epsilon_{xc}[\rho(\vec{\boldsymbol{r}})]\, F_{xc}[s]\, \mathrm{d}^3r \tag{1.39}$$

where $\epsilon_{xc}[\rho(\vec{\boldsymbol{r}})]$ is the exchange-correlation energy per particle of an electron gas in a uniform electron density $\rho(\vec{\boldsymbol{r}})$ (i.e. similar to LDA). F_{xc} is the enhancement factor that tells how much XC energy is enhanced over its LDA value for a given $\rho(\vec{\boldsymbol{r}})$. Note the resemblance of GGA functional in (1.39) to the LDA functional in (1.35) which differ only by an enhancement factor. Here s is a dimensionless reduced gradient

$$s = \frac{|\boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})|}{2(3\pi^2)^{1/3}\rho(\vec{\boldsymbol{r}})^{4/3}} \tag{1.40}$$

The most popular GGA functionals used in the literature are Perdew-Burke-Ernzerhof (PBE) [34], PBEsol [35],Becke88 (B88) [36], Perdew-Wang (PW91) [37], Lee-Yang-Parr (LYP) [38], OptX (O) [39] and Xu (X) [40]. Among the functionals, PBE is the simplest and has exchange

enhancement factor of the form

$$F_x^{\mathrm{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \tag{1.41}$$

where κ and μ are parameters obtained from physical constraints. When the density gradient approaches to zero ($|\boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})| \to 0, s \to 0$), $F_{xc}^{PBE}(s)$ will become unity and (1.39) reduces to LDA formulation. The form of the correlation functional is a complicated function of s and its discussion is beyond the scope of this thesis.

GGA functionals retained most of the correct features of LDA with much greater accuracy [41]. In addition, GGA tends to give better total energies, atomization energies, and energy barriers [36, 42–44]. However, GGA-based schemes typically fail on the region of weak interatomic interactions such as weak Hydrogen bonds, van der Waals interaction, and charge-transfer complexes [45–47].

1.5 Corrections to DFT

One important limitation of DFT that matters most in solid-state physics is the underestimation of band gap in semiconductors. A good theory must successfully predict the properties of wide range of materials including those novel materials as it is critical for applications in optoelectronics and nanotechnology. Hence, this section will discuss the inherent band gap problem and existing methods on improving the band gap.

1.5.1 Band Gap Problem

Given a set of eigenvalues, the band gap E_g is the difference in energy between the lowest unoccupied and highest occupied states

$$E_g^{KS} = \epsilon_{\mathrm{CBM}} - \epsilon_{\mathrm{VBM}} \tag{1.42}$$

where the CBM and VBM refer to conduction band minimum and valance band maximum, respectively. Note that E_g^{KS} is obtained from the calculation of Kohn-Sham band structure.

The CBM and VBM can be approximated as

$$\epsilon_{\text{CBM}} \approx E_{N+1} - E_N \tag{1.43}$$

$$\epsilon_{\text{VBM}} \approx E_N - E_{N-1} \tag{1.44}$$

where E_N and $E_{N\pm1}$ are the ground-state total energies of the neutral system and with one electron added or removed, respectively. By combining equations (1.42)-(1.44), the band gap can be calculated as [48]

$$\begin{aligned} E_g^{KS} &\approx (E_{N-1} - E_N) - (E_N - E_{N+1}) \\ &\approx I - A \end{aligned} \tag{1.45}$$

The first term is precisely the ionization energy. In the case of solids, the same quantity is referred as work function which can be measured directly from photoelectron spectroscopy (PES) experiments. The second term is the electron affinity and can similarly be measured.

The origin of the band gap problem stems from the fact that (1.45) is an approximation to the 'quasiparticle gap' or 'electrical gap' [49]

$$E_g^{qp} = (E_{N-1} - E_N) - (E_N - E_{N+1}) \tag{1.46}$$

For the case of molecules and atoms, the calculation of total energies under DFT in the neutral state (E_N), cationic state (E_{N-1}), and anionic state (E_{N+1}) is possible. Therefore, E_g^{qp} can be calculated directly from differences in total energies without invoking to Kohn-Sham eigenvalues. However, in the case of extended systems such as solids, the change in electron density upon the addition or removable of one electron is extremely small ($\Delta\rho \sim 10^{-20}\rho$). By taking the limit $\Delta\rho \to 0$, it can be shown that [50, 51]

$$\lim_{\Delta\rho\to 0} E_g^{qp} = E_g^{KS} + \Delta_{xc} \tag{1.47}$$

where the correction factor Δ_{xc} is given by

$$\Delta_{xc} = \lim_{\Delta\rho \to 0} V_{xc}[\rho + \Delta\rho] - V_{xc}[\rho - \Delta\rho] \tag{1.48}$$

This implies that quasiparticle gap and the Kohn–Sham band gap differs by a constant, Δ_{xc}. This also means that Δ_{xc} must not be zero, suggesting Δ_{xc} has a discontinuity at the specified limit [52, 53]. The problem with this formulation is that the exact exchange-correlation functional is not yet known. If LDA or GGA functional is used instead, V_{xc} will be a continuous function by construction and therefore there is no discontinuity (i.e. $\Delta_{xc} = 0$). The band gap problem of DFT is a result of the Kohn-Sham formulation of DFT, and in particular to the approximations made in exchange-correlation functional.

1.5.2 GW Approximation

The most suitable method for studying single particle excitation spectra such as ionization energies and electron affinities of extended systems is the Green's function. The Green's function relies on the calculation of the self-energy operator which is non-local, energy dependent, and non-Hermitian [54]. The self-energy is best approximated by the so called quasiparticle GW approximation, after the pioneering works of Hedin and Lundqvist [55, 56]. GW stands for the single-particle Green's function (G) and the dynamically screened Coulomb interaction (W). In practice, the exchange-correlation functional is replaced by the self-energy Σ [57]

$$\hat{\mathcal{V}}_{xc}\phi_i(\vec{\boldsymbol{r}}) \to \int \Sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}', \epsilon_i)\phi_i(\vec{\boldsymbol{r}}')\ \mathrm{d}r' \tag{1.49}$$

so that the Kohn-Sham equation in (1.34) is modified as

$$(\hat{\mathcal{T}}_{KS} + \hat{\mathcal{V}}_{ext} + \hat{\mathcal{V}}_H)\phi_i(\vec{\boldsymbol{r}}) + \int \Sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}', \epsilon_i)\phi_i(\vec{\boldsymbol{r}}')\ \mathrm{d}r' = \epsilon_i\phi_i(\vec{\boldsymbol{r}}) \tag{1.50}$$

The GW approximation for Σ is [58]

$$\Sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}', \omega) = \frac{i}{4\pi} \int G(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}', \omega + \omega')W(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}', \omega')\ \mathrm{d}\omega' \tag{1.51}$$

where ω is the angular frequency related to energy as $\epsilon = \hbar\omega$. The precise meaning of G and W can be found in the seminal work of Hedin and Lundqvist [56] which involves the use of six coupled equations that are solved self-consistently. The self-energy Σ takes into account the finite discontinuity of Δ_{xc} in (1.48), thus yielding the correct quasiparticle band gap. Figure 1.5 illustrates the effectiveness of GW Approximation in improving the band gaps of semiconductors. Clearly, the band gaps calculated using LDA are greatly underestimated. The price to pay in using GW Approximation is that such calculations are considerably more computationally expensive, due partly to complications in convergence of total energies and unfavorable scaling with respect to the system size [59–61].

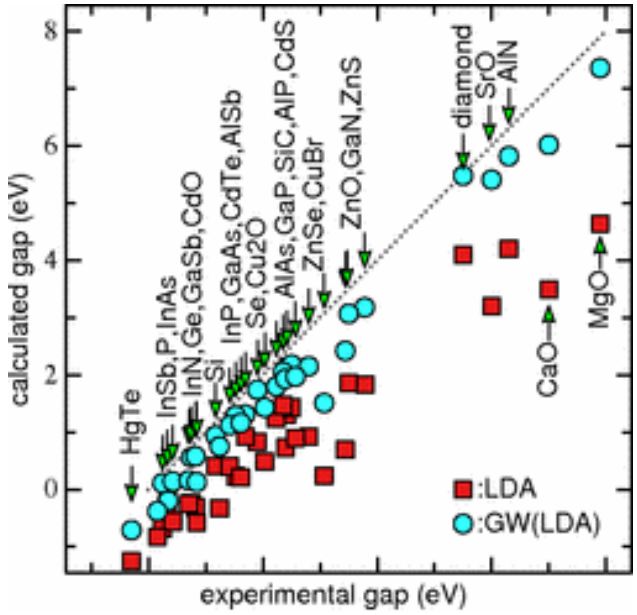

Figure 1.5 Improvement of band gap of semiconductors under GW Approximation. Squares correspond to band gaps calculated using LDA while circles correspond to GW Approximation. If the data point is below the dotted line, the calculated band gap is underestimated. Otherwise, it is overestimated. Illustration taken from [62].

1.5.3 Hybrid Functionals

Hybrid functional admixes a fixed amount of non-local Hartree-Fock exchange with the local or semi-local DFT exchange. The exchange functional in standard DFT was only

approximated under LDA and GGA functional (i.e. eqtn (1.37)). However, the HF exchange is given in exact form in eqtn (1.26). The simplest hybrid functional is the linear combination of the two exchange

$$E_{xc}^{\text{hybrid}} = a_0 E_x^{\text{HF}} + (1 - a_0) E_x^{\text{DFT}} + E_c, \quad 0 \leq a_0 \leq 1 \tag{1.52}$$

The hybrid functional used by Becke [63] has the form $E_x^{\text{HF}} = E_x^{\text{DFT}}$ with $a_0 = 0.5$ using LDA formulation. The PBE0 [64, 65] hybrid functional is constructed by a rational mixing of 25% HF exchange and 75% PBE exchange, with 100% PBE correlation having the form

$$E_{xc}^{\text{PBE0}} = \frac{1}{4} E_x^{\text{HF}} + \frac{3}{4} E_x^{\text{PBE}} + E_c^{\text{PBE}} \tag{1.53}$$

By far the most commonly used functional is the B3LYP (Becke,3-parameter,Lee,Yang,Par) [36, 38] hybrid functional which has the form [66]

$$E_{xc}^{\text{B3LYP}} = E_x^{\text{LDA}} + a_0(E_x^{\text{HF}} - E_x^{\text{LDA}}) + a_x(E_x^{\text{B88}} - E_x^{\text{LDA}}) + E_c^{\text{LDA}} + a_c(E_c^{\text{LYP}} - E_c^{\text{LDA}}) \tag{1.54}$$

where $a_0 = 0.20$,$a_x = 0.72$, and $a_c = 0.81$. The parameters were determined by fitting to a data of measured atomization energies [64].

A critical feature of Hartree-Fock exchange is that it is nonlocal, that is, it cannot be evaluated at one particular spatial location, unless the electron density is known for all locations. Introducing nonlocality greatly increases the computational cost in solving Kohn-Sham equation. These type of functional are very difficult to apply in bulk and spatially extended systems. As a result, HF exact exchange find almost its use in quantum chemistry calculations involving molecules. However, progress is being done in developing the screened hybrid functionals in which the exchange interaction is split into two regions, a long-range (i.e. interstitial region) and a short-range (i.e. core region) interaction. The HF exchange is only incorporated to the short-range portion while standard DFT exchange acts on all portion. The Heyd, Scuseria, and Ernzerhof (HSE) functional is based on this approach which is calculated as [67, 68]

$$E_{xc}^{\text{HSE}} = \frac{1}{4} E_x^{\text{HF,SR}}(\omega) + \frac{3}{4} E_x^{\text{PBE,SR}}(\omega) + E_x^{\text{PBE,LR}}(\omega) + E_c^{\text{PBE}} \tag{1.55}$$

where the screening parameter ω defines the separation range, SR and LR refer to short range and long range, respectively.

1.5.4 Meta-GGA

Meta-GGA is an extension of the GGA in which the local kinetic energy density is included in the input to the functional. The GGA exchange-correlation functional in (1.39) is modified to include the non-interacting kinetic energy density τ [26]

$$E_{xc}^{\mathrm{MGGA}}[\rho] = \int \rho(\vec{\boldsymbol{r}})\, \epsilon_{xc}[\rho(\vec{\boldsymbol{r}})]\, F_{xc}[\rho(\vec{\boldsymbol{r}}), \boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}}), \tau(\vec{\boldsymbol{r}})]\ \mathrm{d}^3 r \tag{1.56}$$

where $\tau(\vec{\boldsymbol{r}})$ is defined as

$$\tau(\vec{\boldsymbol{r}}) = \frac{1}{2}\sum_{i=1}^{N} |\boldsymbol{\nabla}\phi_i(\vec{\boldsymbol{r}})|^2 \tag{1.57}$$

The implementation of Tao, Perdew, Staroverov, and Scuseria (TPSS) functional is based on meta-GGA functional [69]. Its exchange enhancement factor has a similar form as the PBE-GGA functional in (1.41) [26]

$$F_x^{\mathrm{TPSS}}(s) = 1 + \kappa - \frac{\kappa}{1 + \chi/\kappa} \tag{1.58}$$

where χ is a complicated function of $\rho(\vec{\boldsymbol{r}})$, $\boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}})$, and $\tau(\vec{\boldsymbol{r}})$. Other meta-GGAs have also been proposed recently such as Tran, Blaha-modified Becke, Johnson (TB-mBJ) functional [70]; Perdew, Kurth, Zupan, and Blaha functional[71]; and Strongly Constrained and Appropriately Normed Density functional (SCAN) [72].

Since there are kinetic energy density corrections incorporated in the functional, the accuracy of meta-GGAs can compete with the computationally expensive hybrid or GW calculations. It is as cheap as the standard DFT such as LDA or GGA and hence can be scaled to large systems efficiently [70]. It also improves the band gaps of various insulators, semiconductors, oxides and halides but it fails sometimes on materials containing d and f orbitals [73, 74].

1.5.5 Hubbard-U Correction

One of the corrective approach used in the DFT electronic band gap problem is the DFT+U method. One pertinent problem in DFT is the description of strongly correlated systems in which the exchange-correlation functional tends to over-delocalize valence electrons. This problem is more pronounced to systems whose ground state energies are characterized by localized valence electrons such as d orbitals, f orbitals, and Mott insulators [75]. The inability of XC functionals to fully cancel the self-interaction in the Hartree term leads to an excessive delocalization, hence, creating a larger dispersion (i.e. larger valence bandwidth) and smaller band gap than what expected. The strong correlation stems from the Coulomb repulsion between electrons that forces them to localize [76]. This Coulomb potential is described by the term "U". Various models have been proposed to treat correlated systems, and one of the simplest is the "Hubbard-U" model which takes into account the "on-site" repulsion of electrons at the same atomic orbitals [77]. The total energy can be written as [78–80]

$$E_{\mathrm{DFT}+U}[\rho(\vec{\boldsymbol{r}})] = E_{\mathrm{DFT}}[\rho(\vec{\boldsymbol{r}})] + E_{\mathrm{HUB}}[n^I_{mm'}] - E_{\mathrm{dc}}[n^I] \tag{1.59}$$

where $E_{\mathrm{DFT}}[\rho(\vec{\boldsymbol{r}})]$ is the eigenvalue of the Kohn-Sham equation in (1.34), E_{HUB} is the energy of Hubbard functional that describes the correlated systems, and E_{dc} is the double-counting correction when treating electronic interactions as a mean field. Based on this formulation, DFT+U energy can be expanded as

$$E_{\mathrm{DFT}+U}[\rho(\vec{\boldsymbol{r}})] = E_{\mathrm{DFT}}[\rho(\vec{\boldsymbol{r}})] + \sum_I \left[\frac{U^I}{2} \sum_{m \neq m'} n^I_m n^I_{m'} - \frac{U^I}{2} n^I (n^I - 1) \right] \tag{1.60}$$

where n^I_m are the occupation numbers of localized orbitals identified by the atomic site index I and state index m (magnetic quantum number for a particular angular quantum number l). The occupations are computed from the projection of Kohn-Sham orbitals onto the localized basis set such as the atomic orbital states [81]

$$n^I_{mm'} = \sum_{k,i} f_{ki} \left\langle \phi_{ki} \middle| \psi^I_m \right\rangle \left\langle \psi^I_{m'} \middle| \phi_{ki} \right\rangle \tag{1.61}$$

where ϕ_{ki} are the Kohn-Sham states (labeled by k-point k and band index i), f_{ki} represents their occupations according to Fermi-Dirac distribution, and ψ_m^I are the atomic orbitals. In some published works, the on-site Coulomb term U is replaced by an effective potential $U_{eff} = U + J$, where J is the site exchange term that accounts for Hund's rule coupling [78, 79]. The effective potential is proved to be crucial in describing strong spin-orbit coupling.

Hubbard-U calculations depend on the values of U, where it can be either formulated from first principles or achieved empirically by tuning it such that it agrees with the experimental results. The former can be achieved through linear response method in which the response of the localized states to a small perturbation is calculated [80, 82]. The latter is usually compared to the experimental band gap. Nevertheless, the empirical tuning is much preferred because of the significant computational cost in doing linear response calculations, and also the calculated U is not necessarily better than the empirical one [83]. As stated earlier, Hubbard-U calculation can be used to correct the band gaps of strongly correlated systems. This is applicable to semiconductors containing of d and f orbitals such as ZnO, CeO_2, TiO_2, etc. Figure 1.6 shows the improvement of band gaps of transition-metal oxides using Hubbard-U correction. Note that the band gap of MnO is underestimated while FeO is incorrectly predicted as metallic when LDA is used.

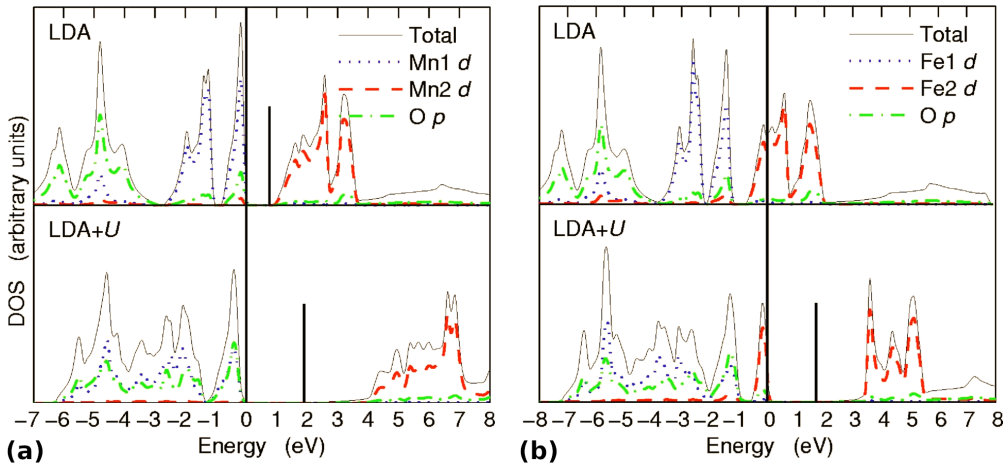

Figure 1.6 Comparison of density of states calculated by LDA and LDA+U for (a) MnO and (b) FeO. The solid vertical bars indicate the end of fundamental band gap. Fermi energy is set at 0 eV. Illustration taken from [84].

Chapter Two

DFT Calculation of Solids

2.1 Basis Sets

Solving the Kohn-Sham equation in (1.34) requires the use of mathematical representations to describe the single-particle Kohn-Sham orbitals $\phi_i(\vec{\boldsymbol{r}})$. One possibility is to express these orbitals as a basis sets that are known and numerically solvable. One starts by expressing orbitals as a linear combination of generic basis set

$$\phi_i(\vec{\boldsymbol{r}}) = \sum_{\alpha}^{M} c_{\alpha}^{i} \, |\chi_\alpha\rangle \tag{2.1}$$

where i is the band index, the sum runs over all the basis functions up to the dimension M, and c_{α}^{i} is the expansion coefficient of a known basis function $|\chi_\alpha\rangle$. Since $\phi_i(\vec{\boldsymbol{r}})$ spans the whole infinite space, M must be in principle infinite. However, in practice the basis set is truncated just enough for an accurate description of the orbital. The choice of the basis set depends on several factors such as (a) efficiency and (b) unbiased [85]. A basis set is efficient if it resembles $\phi_i(\vec{\boldsymbol{r}})$ closely, hence requiring less expansion coefficients and smaller dimension size. However, this assumes that the solution to the problem must be known beforehand. Such basis set can never be general because it will quickly yield a solution for a specific problem but will poorly perform for other cases. The problem is that optimizing a basis set for a specific system can cause bias. This means that if a property of a system is

calculated, but the basis set is optimized for only one particular system, the result will be biased towards that one [86, 87]. It is the goal of theoretical condensed matter physics to find a basis set that is simultaneously efficient and unbiased. There are three types of basis sets that are commonly used for expansions, namely: local, nonlocal, and augmented basis.

2.1.1 Local Basis Set

A local basis has its peak centered on a local point and is well applicable to orbitals around individual atoms in real space. Gaussian basis sets or any atom-centered basis orbitals are examples of this type. It is a popular choice for atoms and molecules whose orbitals are highly localized around each atom. Hence, less than 20 basis functions per atom are sufficient enough to achieve acceptable accuracy. As an example, the Slater type basis orbitals (STO) are written as [88]

$$\left|\chi_\alpha^{\mathrm{STO}}\right\rangle = Ae^{-\alpha r} \tag{2.2}$$

where A is some normalization constant. Note that STO exponentially decays away from an atom centered at $\vec{\boldsymbol{r}}$. On the other hand, Gaussian type basis orbitals (GTO) are written as [89]

$$\left|\chi_\alpha^{\mathrm{GTO}}\right\rangle = Ae^{-\alpha r^2} \tag{2.3}$$

GTO has the advantage that all integrals associated with it can be performed analytically. Since these basis sets are localized, they cannot properly described the long-range interaction of metals and the periodicity of crystalline solids.

2.1.2 Nonlocal Basis Set

Nonlocal basis set span the whole space. An important class of basis orbital under this category is the plane wave basis (PW) described as

$$\left|\chi_\alpha^{\mathrm{PW}}\right\rangle = Ae^{i\alpha r} \tag{2.4}$$

which can be generalized as a dot product of wavevector $\vec{\mathbf{k}}$ and position vector $\vec{\mathbf{r}}$

$$\left|\chi_{\vec{\mathbf{k}}}^{\mathrm{PW}}\right\rangle = Ae^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}} \tag{2.5}$$

PWs are the most commonly used in DFT because of the following reasons: PWs are already solutions to periodic systems satisfying the Bloch condition; PWs are convenient in taking gradients and integrals because of their exponential form; changing the domain of PWs from real space to reciprocal space are easily executed using Fourier transformation; PWs are orthogonal which simplifies calculation; and lastly, PWs are independent of the atomic positions because of its nonlocal nature. However, there are also disadvantages of using it. It requires an enormous amount of PWs to properly describe the rapid fluctuation of orbital wavefunctions near the core region of an atom or ion. A direct fix to this problem is the application of pseudopotentials to smoothen the strong Coulomb potential of the nucleus, which will be the topic of later section [90].

2.1.3 Augmented Basis Set

Augmented basis sets are combinations of local and nonlocal basis sets. Under this category is the Augmented plane waves (APW) basis. The APW divides the space into two regions: the core region,where the orbitals are atomic-like; and the interstital region, where the orbitals resemble plane waves [91]. The basis orbitals are taken to be [92]

$$\left|\chi_{\vec{\mathbf{k}}}^{\mathrm{APW}}\right\rangle = \begin{cases} \text{atomic basis} & , \left|\vec{\mathbf{r}} - \vec{\mathbf{R}}\right| \leq r_c \\ Ae^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}} & , \left|\vec{\mathbf{r}} - \vec{\mathbf{R}}\right| > r_c \end{cases} \tag{2.6}$$

where $\vec{\mathbf{R}}$ is the center of the atom and r_c is the core radius. Outside the core, the orbital wavefunction is a plane wave because the potential is constant there. Inside the core, the orbital wavefunction is atomic-like and can be solved by the appropriate Schrödinger equation. The potential involved in this type of basis is usually called muffin-tin potential due its resemblance to muffin tins. APWs must satisfy the boundary conditions at $\left|\vec{\mathbf{r}} - \vec{\mathbf{R}}\right| = r_c$.

That is, the basis orbital must be continuous at the boundary value and its slope exists [93]. Augmented plane waves are very accurate because it describes both core electrons and valence electrons well. However, accuracy is always associated with computational costs.

2.2 Matrix Formulation for KS equation

The use of basis sets transforms the Kohn-Sham equation into an ordinary matrix algebra that can be solved numerically. The Kohn-Sham equation in (1.34) is expanded in terms of basis sets using (2.1)

$$\hat{\mathcal{H}}_{KS} \sum_{\alpha}^{M} c_{\alpha}^{i} \left|\chi_{\alpha}\right\rangle = \epsilon_i \sum_{\alpha}^{M} c_{\alpha}^{i} \left|\chi_{\alpha}\right\rangle \tag{2.7}$$

Left "multiply" with $\langle\chi_\beta|$:

$$\sum_{\alpha}^{M} \left\langle\chi_{\beta}\right| \hat{\mathcal{H}}_{KS} \left|\chi_{\alpha}\right\rangle c_{\alpha}^{i} = \sum_{\alpha}^{M} \left\langle\chi_{\beta}|\chi_{\alpha}\right\rangle c_{\alpha}^{i} \epsilon_i \tag{2.8}$$

which can be simplified as

$$\sum_{\alpha}^{M} H_{\beta\alpha} c_{\alpha}^{i} = \sum_{\alpha}^{M} S_{\beta\alpha} c_{\alpha}^{i} \epsilon_i \tag{2.9}$$

where $H_{\beta\alpha}$ and $S_{\beta\alpha}$ are the energy-independent Hamiltionian and the overlap matrix, respectively [94]. The elements of these matrices are defined as

$$H_{\beta\alpha} = \left\langle\chi_{\beta}\right| \hat{\mathcal{H}}_{KS} \left|\chi_{\alpha}\right\rangle = \int \chi_{\beta}^{*}(\vec{\boldsymbol{r}}) \hat{\mathcal{H}}_{KS} \chi_{\alpha}(\vec{\boldsymbol{r}}) \, \mathrm{d}^3 r \tag{2.10}$$

$$S_{\beta\alpha} = \left\langle\chi_{\beta}|\chi_{\alpha}\right\rangle = \int \chi_{\beta}^{*}(\vec{\boldsymbol{r}}) \chi_{\alpha}(\vec{\boldsymbol{r}}) \, \mathrm{d}^3 r \tag{2.11}$$

The overlap matrix $S_{\beta\alpha}$ takes into account the possible non-orthogonality of the basis functions [16]. Note that for plane wave (PW) basis sets, which are orthonormal, the overlap matrix $S_{\beta\alpha}$ becomes an unit matrix. The general matrix eigenvalue problem can be recast into a compact form [95]

$$\boldsymbol{Hc = ScA} \tag{2.12}$$

where $\mathbf{\Lambda}$ is the diagonal matrix containing energy eigenvalues and $\boldsymbol{c}$ has the eigenfunction (expansion coefficients of the KS orbital) as columns. In solving (2.12), the normalization condition must be taken into account

$$\int \phi_i(\vec{\boldsymbol{r}})\phi_i^*(\vec{\boldsymbol{r}}) = \int \sum_{\alpha}^{M}\sum_{\beta}^{M} c_\alpha^{i*}\chi_\alpha^* c_\beta^i \chi_\beta \, \mathrm{d}^3 r = 1 \tag{2.13}$$

$$= \sum_{\alpha}^{M}\sum_{\beta}^{M} c_\alpha^{i*} c_\beta^i \int \chi_\alpha^* \chi_\beta \, \mathrm{d}^3 r = 1 \tag{2.14}$$

$$= \sum_{\alpha}^{M}\sum_{\beta}^{M} c_\alpha^{i*} c_\beta^i S_{\alpha\beta} = 1 \tag{2.15}$$

There are $M \times N$ elements of $\boldsymbol{c}$ needed to be solved, where M is the total number of basis functions used and N is the total number of lowest-energy orbitals. In addition, there are N unknown energy eigenvalues to be solved. Fortunately, there are $M \times N$ independent equations in (2.12) and N equations coming from the normalization condition (2.15) so that $N(M+1)$ equations are simultaneously solved [85]. It is obvious that increasing either M or N will increase the computational power needed. This does not inlude yet the iterative self consistent field calculation, as shown in Figure 1.3, needed to have converged electron density. Common numerical algorithms used in matrix diagonalization are Davidson iterative diagonalization [96, 97], Residual minimization/direct inversion in the iterative subspace (RMM–DIIS) [98, 99], and Conjugate-gradient-like band-by-band diagonalization [100, 101]. Numerical algorithms must be efficient and optimized since most DFT codes spend substantial amount of time in matrix diagonalization.

2.3 Pseudopotential (PP) Approach

The idea behind the use of pseudopotentials is to replace the strong Coulomb potential of the nucleus by an effective potential acting on the valence electrons [102–104]. When atoms bond together to form a solid, the core electrons are so localized in a deep potential well that they remain invariant. Thus, their contribution to bonding is negligible and the

replacement of its potential by a simple fictitious potential is justified. Figure 2.1 illustrates the action of pseudopotential on the wavefunction and potential of an atomic orbital. The all-electron wavefunction contains nodes which are computationally difficult to solve. On the other hand, pseudo wavefunction is nodeless everywhere, and therefore it greatly reduces the number of plane waves required for the calculation by a significant amount. Note that at large distances away from the nucleus, both potential becomes constant and the wavefunction is expected to be a plane wave. By effectively neglecting the core electrons from the calculation, the Kohn-Sham orbitals needed is dramatically reduced. This will substantially reduce the computational time required to calculate orbital-dependent quantities.

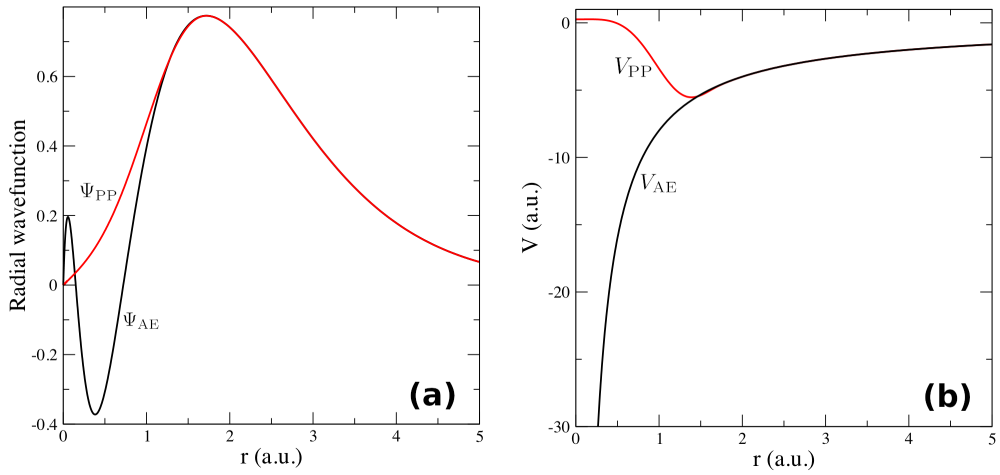

Figure 2.1 Schematic illustration of a (a) pseudo wavefunction pseudized from a 3s wavefunction of Si orbital and the (b) corresponding pseudo- (PP) and all-electron (AE) potentials. The all-electron approach takes into account all electrons including core and valence electrons. Illustration taken from [105].

There are two criteria for choosing a good pseudopotential, namely: softness and transferability [106, 107]. A pseudopotential is soft if it requires few plane waves to model the system. This is similar to efficiency of basis sets. A pseudopotential is transferable if it can be used in whatever environment (e.g. molecule, solid, cluster, surface, metal, insulator, etc). The choice depends on which pseudopotential is advantageous to use and the type of calculation being done. The common pseudopotentials used in DFT codes are Norm-Conserving pseudopotential, Ultrasoft pseudopotential, and Projector Augmented Wave.

2.3.1 Norm-Conserving Pseudopotential (NCPP)

In norm-conserving pseudopotentials, the pseudopotential and all-electron charge densities are set equal so that the norm is conserved in both potentials [108, 109]. Pseudopotentials are generated to meet this criterion

$$\int_0^{r_c} |\phi_{\mathrm{PP}}(\vec{\boldsymbol{r}})|^2 \, \mathrm{d}\vec{\boldsymbol{r}} = \int_0^{r_c} |\phi_{\mathrm{AE}}(\vec{\boldsymbol{r}})|^2 \, \mathrm{d}\vec{\boldsymbol{r}} \tag{2.16}$$

where r_c is the chosen cutoff radius that separates the core region from the valence region. The constraint imposed on this pseudopotential leads to an improvement in transferability of potentials to different chemical environments. In addition, reducing r_c improves the transferability because in this way the pseudo wavefunction becomes closer to the all-electron result. However, the cutoff radius should be chosen outside the location of the maximum node of the all-electron wavefunction. Note that this pseudopotential gives only the valence charge density and not the total charge density. Other norm-conserving schemes were proposed by Troullier and Martin (TM) [109], and by Rappe, Rabe, Kaxiras, and Joannopoulos (RRKJ) [110].

2.3.2 Ultrasoft Pseudopotential (USPP)

Ultrasoft pseudopotentials were introduced in order for the calculations to have lowest numbers of plane waves basis set used since it was shown that the norm of the all-electron and pseudo wavefunction was not necessary requirement for transferability. Hence, this was done by Vanderbilt [111] who showed that smoother but highly transferable pseudopotentials are possible. The cutoff radius r_c is situated farther than the equivalent norm-conserving pseudopotential and the pseudo wavefunction is flatter. This leads to fewer plane waves that gives significant reduction in computational time. Similar to norm-conserving pseudopotentials, the ultrasoft pseudopotential only gives valence charge densities, not total charge densities.

2.3.3 Projector Augmented Wave (PAW)

The Projector Augmented Wave takes into account both all-electron and pseudo wavefunction into calculations. It aims for both efficiency of using pseudopotential and the accuracy of using all-electron potential [112, 113]. However, the all-electron wavefunction is limited only on the core region and will be truncated beyond the cut-off radius r_c. A correction factor is added to subtract the overlapping part of the pseudo wavefunction in the core region [114]. Hence, the PAW wavefunction involves three terms

$$\psi_{\mathrm{PAW}} = \psi_{\mathrm{AE}} + \psi_{\mathrm{PP}} - \psi_{\mathrm{net}} \tag{2.17}$$

The actions of the terms in the equation above are visualized in Figure 2.2. The ψ_{PP} is expanded in plane wave basis sets while ψ_{AE} is only defined within the cutoff radius r_c. The ψ_{net} subtracts the overlapping part of ψ_{PP} in the core region.

Note that in Kohn-Sham formulation, these wavefunctions ψ become the independent Kohn-Sham orbitals ϕ. PAW calculations are accurate as all-electron calculations with much less computational effort. Unlike the two pseudopotentials mentioned before, PAW pseudopotential returns both the core and valence charge densities.

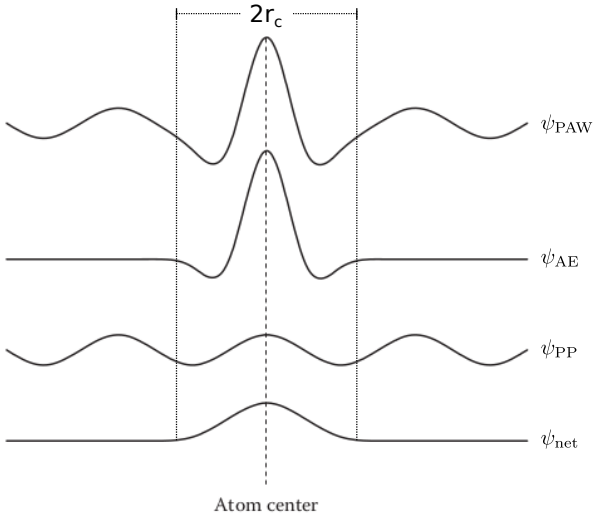

Figure 2.2 Schematic illustration of the wavefunctions used in PAW pseudopotential. Illustration taken from [4].

2.4 Supercells

Most solids are characterized by its regular repeating three-dimensional structure called a crystal lattice. Hence, it is possible to study solids by just looking at the building block, which is ordinarily called the unit cell. In order to model solids that are feasible for computational simulation, repeating unit cells that are stack together must be needed. These stacked unit cells are collectively called supercell. When implementing DFT, the periodic boundary conditions must be taken into account. In this case, the supercell is duplicated periodically throughout the whole space. However, the actual calculation is applied only on a single supercell while the rest (called images) simply copies it with no significant computational cost.

When defects are introduced into the supercell, it forms a periodic array of defects across all the images of the supercell. A supercell must be large enough so that the calculation is independent of the location of the defect inside the supercell and also to reduce the interaction with its images. Thus, the Kohn-Sham equation and other pertinent calculations are solved only within a single supercell.

2.5 DFT Calculation in Reciprocal Space

Working in reciprocal space is greatly convenient if functions are expressed in terms of plane waves. Plane waves propagate in real space but they become point in the reciprocal space, wherein each point corresponds to a particular wavevector $\vec{\mathbf{k}}$. The lattice points in real space will define the allowed wavevectors $\vec{\mathbf{G}}$, the reciprocal lattice vector which is subset of the reciprocal space. See Appendix A for discussions about reciprocal lattice and Brillouin zones. Since real space and reciprocal space have inverse relationship, increasing the supercell size by a certain factor will cause the supercell in reciprocal space to shrink by a same factor, and vice versa. Figure 2.3 illustrates this relationship. Note that no information is lost when

transforming between the two spaces. In addition, bigger supercells require fewer k-points but in the expense of many atoms included in calculations.

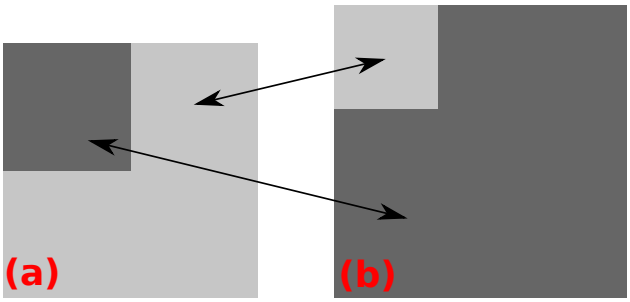

Figure 2.3 Relationship between a supercell in (a) real space and the corresponding (b) reciprocal space.

In previous discussions, $\vec{\mathbf{k}}$ was defined as any wavevector in the reciprocal space. However, it can always be transform to $\vec{\mathbf{k}} \rightarrow \vec{\mathbf{k}} + \vec{\mathbf{G}}$ so that the new $\vec{\mathbf{k}}$ is in the first Brillouin zone and that any wavevectors are equivalent to the new one by a reciprocal lattice vector $\vec{\mathbf{G}}$. This transformation will limit DFT calculations inside the first Brillouin zone instead of the whole reciprocal space.

Furthermore, one can take advantage of the symmetry of the solid to reduce the first Brillouin zone into what is called irreducible Brillouin zone (IBZ). Hence, DFT calculations will be narrowed down further into this Brillouin zone. Note that each k-points will have a weight factor that depends on how many times it was folded during symmetry operations (e.g. rotation and inversion). Figure 2.4 summarizes the various techniques employed in simplifying DFT calculations in solids starting from a bulk solid, then a supercell simulation, transformed to reciprocal space confined in first Brillouin zone, and further reduced to a irreducible zone by symmetry operations. All these techniques and the pseudopotential approximations made DFT calculations computationally feasible.

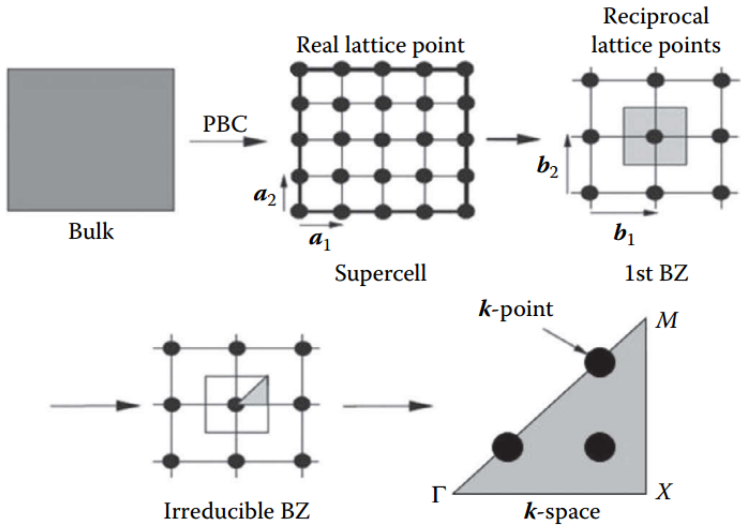

Figure 2.4 Various techniques used in treating solids in DFT calculations. Figure taken from [4].

2.6 k-point Sampling

It was shown that DFT calculations can be solved within the irreducible Brillouin zone. However, there are infinite numbers of k-points inside IBZ that are well qualified for a plane wave. One way to deal with this problem is to sample finite number of k-points that represent each region well. This sampling technique is justified by the fact that orbitals and other quantities vary smoothly in the IBZ. Sampling in IBZ must satisfy two goals: select few k-points as possible to reduce computational time, and select enough so that they represent the actual quantities well. Convergence tests in which the number of k-points are varied until quantities such as total energy does not change anymore must be conducted. The quantity is said to be converged with respect to k-points. Convergence tests are very helpful in finding the optimum number of k-points with minimum error. Any integrated function $f(\vec{\mathbf{r}})$ can be written over the Brillouin zone

$$f(\vec{\mathbf{r}}) = \int_{BZ} F(\vec{\mathbf{k}})\, \mathrm{d}\vec{\mathbf{k}} \tag{2.18}$$

where $F(\vec{\mathbf{k}})$ is the Fourier transform of $f(\vec{\mathbf{r}})$. To evaluate computationally, the integral can be approximated by weighted sum over special k-points

$$\int_{BZ} F(\vec{\mathbf{k}})\,\mathrm{d}\vec{\mathbf{k}} \approx \sum_j w_j F(\vec{\mathbf{k}}_j) \tag{2.19}$$

where w_j are the weighted factors. There are standard schemes for generating k-points grid mesh and probably the most popular one is the Monkhorst-Pack method.

2.6.1 Monkhorst-Pack method

The Monkhorst-Pack method generates a grid of uniform special k-points along the three lattice vectors in the reciprocal space [115]. The construction of the special points is based on the formula

$$u_r = \frac{2r - q_r - 1}{2q_r} \qquad , r = 1, \ldots, q_r \tag{2.20}$$

where q_r determines the number of k-points used along one of the axis in $r = x, y,$ or z, and r varies from 1 to q_r. The k-point is given by

$$\vec{\mathbf{k}} = u_x\vec{\mathbf{b}}_1 + u_y\vec{\mathbf{b}}_2 + u_z\vec{\mathbf{b}}_3 \tag{2.21}$$

wherein $\vec{\mathbf{b}}_1, \vec{\mathbf{b}}_2, \vec{\mathbf{b}}_3$ are the primitive lattice vectors of the reciprocal lattice. u_r gives the fractional part of the corresponding component of the reciprocal lattice vector. For instance, in a $4 \times 4 \times 4$ grid this will correspond to $q_x = q_y = q_z = 4$ with a total number of $q_x \times q_y \times q_z = 64$ k-points. The number of k-points will be further reduced by symmetry operations inside the irreducible Brillouin zone.

The center of the mesh can be centered on the origin ($\vec{\mathbf{k}} = 0$ or Γ point) or shifted by a fixed amount away from the origin. The former is important if one needs to know the electronic states at the Γ point. The latter generally breaks the symmetry, hence, it must be use with care.

2.6.2 Gamma Point Sampling

For very large supercell, the associated first Brillouin zone are very small that it approaches the zone center or the Γ point. Thus, it would be practical to use only one k-point with high weight factor. Calculations based on the sampling at the Γ point reduce significant computational cost because the real and reciprocal space coincide with each other at the origin and the KS orbital will real quantity so that any consideration for complex numbers is not necessary. Γ point calculation are routinely used in massive calculations.

2.7 Bloch Representations in DFT

The Bloch expression for the Kohn-Sham orbital is similar to the many-particle wavefunction derived from Schrödinger equation. See section 1.1.1 for the discussion about periodicity. Here, the Bloch form of KS orbital is

$$\phi_{n,k}(\vec{\mathbf{r}}) = u_{n,k}(\vec{\mathbf{r}})e^{i\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}} \tag{2.22}$$

where $\phi_{n,k}(\vec{\mathbf{r}})$ now depends both on the band index n and the wavevector $\vec{\mathbf{k}}$ confined in the first Brillouin zone. Expanding the periodic function $u_{n,k}$ in terms of plane waves whose wavevectors are reciprocal lattice vector $\vec{\mathbf{G}}$

$$u_{n,k}(\vec{\mathbf{r}}) = \sum_{\vec{G}} C_{n,k}(\vec{\mathbf{G}})\, e^{i\vec{\mathbf{G}}\cdot\vec{\mathbf{r}}} \tag{2.23}$$

where $C_{n,k}(\vec{\mathbf{G}})$ is the expansion coefficient of plane wave basis sets. The phase factor $\exp\left(i\vec{\mathbf{G}}\cdot\vec{\mathbf{r}}\right)$ represents a plane wave travelling in space, perpendicular to $\vec{\mathbf{G}}$. Thus, the KS orbital can be rewritten as

$$\phi_{n,k}(\vec{\mathbf{r}}) = \sum_{\vec{G}} C_{n,k}(\vec{\mathbf{G}})\, e^{i(\vec{\mathbf{k}}+\vec{\mathbf{G}})\cdot\vec{\mathbf{r}}} \tag{2.24}$$

The coefficients $C_{n,k}(\vec{\mathbf{G}})$ can be solved by taking the inverse Fourier transform of $\phi_{n,k}(\vec{\mathbf{r}})$

$$C_{n,k}(\vec{\mathbf{G}}) = \mathcal{F}^{-1}[\phi_{n,k}(\vec{\mathbf{r}})] \tag{2.25}$$

$$= \int \phi_{n,k}(\vec{\mathbf{r}}) e^{-i(\vec{\mathbf{k}}+\vec{\mathbf{G}})\cdot\vec{\mathbf{r}}}\, \mathrm{d}\vec{\boldsymbol{r}} \tag{2.26}$$

$$= \phi_{n,k}(\vec{\mathbf{G}}) \tag{2.27}$$

Similarly, the electron density in real space and reciprocal space are Fourier transform of each other

$$\rho(\vec{\mathbf{r}}) = \sum_{\vec{\boldsymbol{G}}} \rho(\vec{\mathbf{G}}) e^{i\vec{\mathbf{G}}\cdot\vec{\mathbf{r}}} \tag{2.28}$$

$$\rho(\vec{\mathbf{G}}) = \int \rho(\vec{\mathbf{r}}) e^{-i\vec{\mathbf{G}}\cdot\vec{\mathbf{r}}}\, \mathrm{d}\vec{\boldsymbol{r}} \tag{2.29}$$

2.8 Energy Operators in Reciprocal Space

Since DFT calculations take place in the reciprocal space, the Hamiltonian of the KS equation in (1.34) must be transform from the real space to reciprocal space. In the Kohn-Sham formulation, both non-interacting kinetic energy and Hartree potential are easily evaluated because they are local in reciprocal space. The action of the kinetic energy operator on the KS orbital is

$$\hat{\mathcal{T}}_{KS} \left|\phi_{n,k}(\vec{\mathbf{r}})\right\rangle = -\frac{1}{2}\nabla^2 \left(\sum_{\vec{\boldsymbol{G}}} C_{n,k}(\vec{\mathbf{G}})\, e^{i(\vec{\mathbf{k}}+\vec{\mathbf{G}})\cdot\vec{\mathbf{r}}} \right) \tag{2.30}$$

$$= \frac{1}{2} \sum_{\vec{\boldsymbol{G}}} (\vec{\mathbf{k}}+\vec{\mathbf{G}})^2 C_{n,k}(\vec{\mathbf{G}})\, e^{i(\vec{\mathbf{k}}+\vec{\mathbf{G}})\cdot\vec{\mathbf{r}}} \tag{2.31}$$

Hence, the effect of the kinetic energy operator in the reciprocal space is to multiply each coefficient by one-half times the square of its wavevector. Its matrix representation is

$$\hat{\mathcal{T}}_{KS}(\vec{\mathbf{G}}, \vec{\mathbf{G}}') = \left\langle \phi_{n,k}(\vec{\mathbf{r}}) \right| \hat{\mathcal{T}}_{KS} \left| \phi_{n,k}(\vec{\mathbf{r}}) \right\rangle \tag{2.32}$$

$$= \frac{1}{2} \left| \vec{\mathbf{k}} + \vec{\mathbf{G}} \right|^2 \delta_{\vec{\mathbf{G}},\vec{\mathbf{G}}'} \tag{2.33}$$

In the above equation, the bra-term is expanded in terms of $\vec{\mathbf{G}}$ while the ket-term is expanded in $\vec{\mathbf{G}}'$. The Hartree potential in reciprocal space is given by

$$\hat{\mathcal{V}}_H = \frac{1}{2} \sum_{\vec{\boldsymbol{G}}} \left| \rho(\vec{\mathbf{G}}) \right|^2 \tag{2.34}$$

which is more simple compared to its real space counterpart given in (1.21). Its matrix representation is given by

$$\hat{\mathcal{V}}_H(\vec{\mathbf{G}}, \vec{\mathbf{G}}') = \langle \phi_{n,k}(\vec{\mathbf{r}}) | \, \hat{\mathcal{V}}_H \, | \phi_{n,k}(\vec{\mathbf{r}}) \rangle \tag{2.35}$$

$$= \hat{\mathcal{V}}_H(\vec{\mathbf{G}} - \vec{\mathbf{G}}') \tag{2.36}$$

The remaining external potential and exchange-correlation energy can be obtained from their Fourier transform, respectively

$$\hat{\mathcal{V}}_{ext}(\vec{\mathbf{G}}) = \int \hat{\mathcal{V}}_{ext}(\vec{\mathbf{r}}) e^{-i\vec{\mathbf{G}}\cdot\vec{\mathbf{r}}} \, \mathrm{d}\vec{\boldsymbol{r}} \tag{2.37}$$

$$\hat{\mathcal{V}}_{xc}(\vec{\mathbf{G}}) = \int \hat{\mathcal{V}}_{xc}(\vec{\mathbf{r}}) e^{-i\vec{\mathbf{G}}\cdot\vec{\mathbf{r}}} \, \mathrm{d}\vec{\boldsymbol{r}} \tag{2.38}$$

Their matrix representations are similar to (2.36). Taking all the derivations above, the complete Kohn-Sham equation in reciprocal space is

$$\sum_{\vec{\boldsymbol{G}}'} \left[\frac{1}{2} \left| \vec{\mathbf{k}} + \vec{\mathbf{G}} \right|^2 \delta_{\vec{\mathbf{G}},\vec{\mathbf{G}}'} + \hat{\mathcal{V}}_{ext}(\vec{\mathbf{G}} - \vec{\mathbf{G}}') + \hat{\mathcal{V}}_H(\vec{\mathbf{G}} - \vec{\mathbf{G}}') + \hat{\mathcal{V}}_{xc}(\vec{\mathbf{G}} - \vec{\mathbf{G}}') \right] C_{n,k}(\vec{\mathbf{G}}')$$

$$= \epsilon_{n,k} \, C_{n,k}(\vec{\mathbf{G}}) \tag{2.39}$$

which can be rewritten as a general matrix equation

$$\hat{\mathcal{H}}_{\vec{\mathbf{G}},\vec{\mathbf{G}}'} C_{n,k}(\vec{\mathbf{G}}') = \epsilon_{n,k} \, C_{n,k}(\vec{\mathbf{G}}) \tag{2.40}$$

This is similar to (2.12) but the overlapping matrix is set to identity matrix.

2.9 Cutoff Energy

The plane wave basis expansion of both KS orbital and the electron density in (2.24) and (2.28) is evaluated in the complete set of reciprocal lattice vector $\vec{\mathbf{G}}$, which is infinite. This

means that it will take infinitely long to compute desired properties. Nevertheless, orbitals and electron densities tend to become smoothly varying at large $\vec{\mathbf{G}}$-vectors. Thus, their plane wave components become negligible for large $\vec{\mathbf{G}}$. The expansion can be truncated by introducing kinetic energy cut-off E_{cut} defined as

$$E_{cut} = \left|\vec{\mathbf{k}} + \vec{\mathbf{G}}_{cut}\right|^2 \tag{2.41}$$

the $\vec{\mathbf{G}}_{cut}$ serves as the upper bound for the expansion series of KS orbital in (2.24). This means that plane waves whose kinetic energy is less than this cut-off energy are the only ones included in DFT calculations. The cut-off radius of electron densities is usually a multiple of E_{cut} quantified by

$$mE_{cut} = \left|\vec{\mathbf{k}} + \vec{\mathbf{G}}_{rho}\right|^2 \tag{2.42}$$

where m is called a dual, a multiplier of E_{cut}, and $\vec{\mathbf{G}}_{rho}$ serves as the upper bound for the expansion series of electron density in (2.28). The number of plane waves can be estimated as

$$N_{\mathrm{PW}} \approx \frac{1}{2\pi^2} V E_{cut}^{3/2} \tag{2.43}$$

where V is the volume of the supercell. The cut-off energy that is appropriate for a given calculation is not usually known in advance, as it varies on the configuration of the system. However, convergence tests can be conducted where the cut-off energies is increased until the desired properties stop changing. Note that wavevectors $\vec{\mathbf{k}}$ were discretized by using k-point sampling in the IBZ, $\vec{\mathbf{G}}$-vectors become finite by energy cutoff, and the band index n depends on the number of orbital states, which is also finite. Hence, solving Kohn-Sham in (2.39) becomes computationally tractable.

2.10 Ionic Relaxation

For every DFT calculations, the system must be fully relaxed both electronically and structurally. The electronic relaxation is given by the self-consistent field calculation in section

1.3.3. The structural relaxation or ionic relaxation computes the forces of each atom and are moved to directions of minimum forces for the next electronic relaxation. The force on the Ith atom positioned at $\vec{\mathbf{R}}_I$ can be calculated from Hellman-Feynman theorem as [116, 117]

$$\vec{\mathbf{F}}_I = -\frac{\partial E}{\partial \vec{\mathbf{R}}_I} = -\langle\phi| \frac{\partial \hat{\mathcal{H}}}{\partial \vec{\mathbf{R}}_I} |\phi\rangle - \frac{\partial E_{II}}{\partial \vec{\mathbf{R}}_I} \tag{2.44}$$

$$= -\int \rho(\vec{\mathbf{r}}) \frac{\partial \hat{\mathcal{V}}_{ext}(\vec{\mathbf{r}})}{\partial \vec{\mathbf{R}}_I} \, \mathrm{d}\vec{\boldsymbol{r}} - \frac{\partial E_{II}}{\partial \vec{\mathbf{R}}_I} \tag{2.45}$$

Note that the calculation of forces is given strictly in terms of the electron density and the external potential, independent of electron kinetic energy, Hartree potential, and exchange-correlation terms. Thus, forces can be calculated by taking simple derivative operations on two potential terms. This is why force calculations are very fast that they are almost unnoticed in a DFT calculation. The atoms will move in the direction of least force and the process is repeated again until the total force of the system is negligible. The schematic diagram of the complete relaxation in DFT calculations is shown in Figure 2.5. The inner loop bounded by dashed lines is the self-consistent field calculation that was shown previously in Figure 1.4. The outer loop is the ionic relaxation. If the total force of the system is almost zero, then desired material properties such total energy, pressure, stress, etc. can be calculated.

Common numerical algorithms used to implement ionic relaxation are the quasi-Newton method [118], the conjugate gradient (CG) method [119], and the damped molecular dynamics method [120].

2.11 Density Mixing Schemes

In each iteration of the self consistent field calculation, it starts with a given electron density $\phi_i(\vec{\boldsymbol{r}})$ and obtain the corresponding Kohn-Sham Hamiltonian and its eigenstates (see section 1.3.3). A new electron density can be computed from the occupied eigenstates. Afterwards, the new input density is updated from the old ones and will be used for the next iteration.

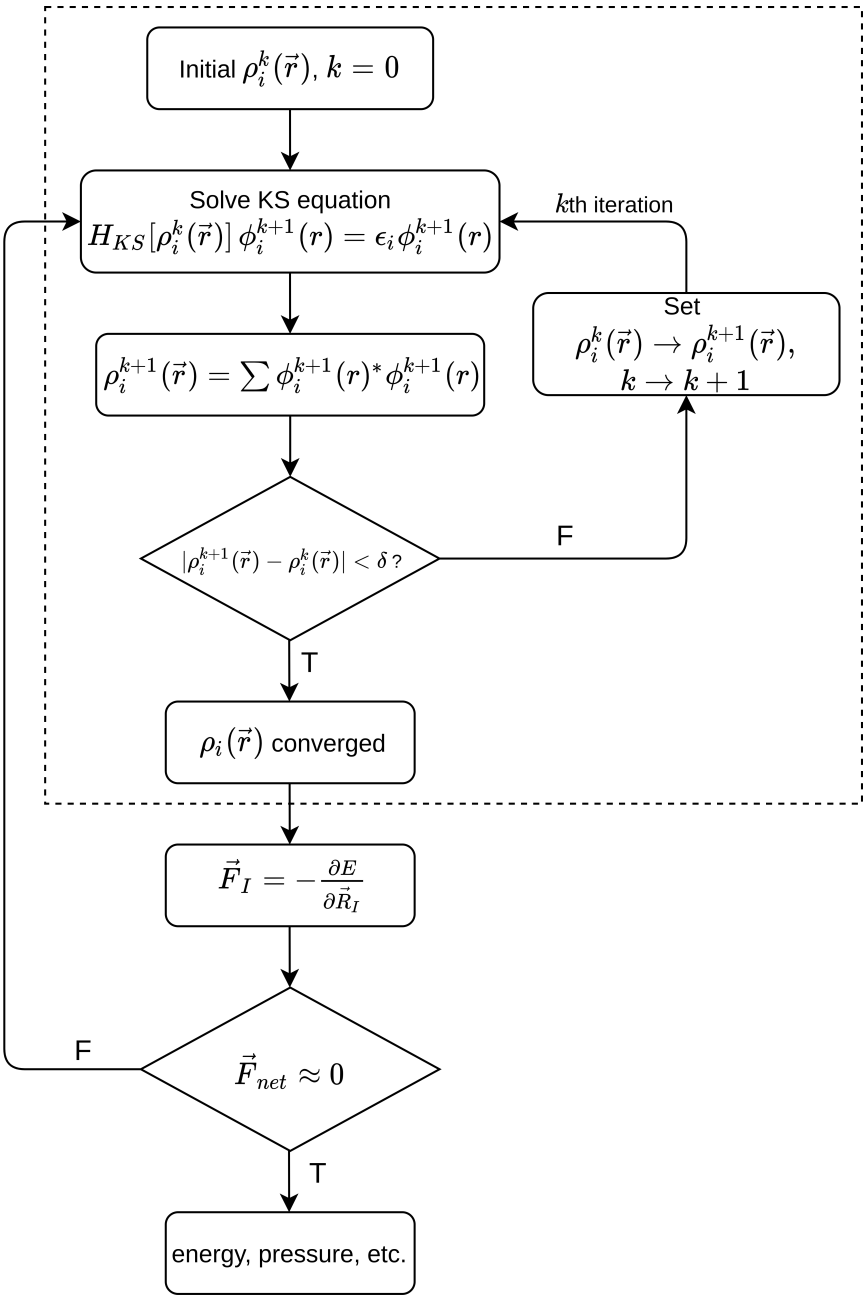

Figure 2.5 Schematic diagram of the complete relaxation in DFT simulations.

The end goal is to reach the self consistent solution, i.e. $\rho_{out}(\vec{\mathbf{r}}) = \rho_{in}(\vec{\mathbf{r}})$, which can be thought as fixed point problem of the form $x_{n+1} = f(x_n)$. If the procedure converges, then the final value is the fixed point $\bar{x} = f(\bar{x})$. The simplest strategy is the linear mixing scheme [121]

$$\rho_{in}^{n+1}(\vec{\mathbf{r}}) = \alpha \rho_{out}^{n}(\vec{\mathbf{r}}) + (1-\alpha)\rho_{in}^{n}(\vec{\mathbf{r}}) \tag{2.46}$$

where α is the empirical mixing parameter adjusted to minimize the number of iterations needed for self consistency. The larger the α, the more is contributed from the output density. The purpose of density mixing is to prevent the charge sloshing. Charge sloshing is the consistent charge overshooting, or the large charge redistribution, that occurs from one iteration to the next. Density mixing will damp out these charge displacements leading to a better convergence [122]. Other advanced mixing schemes commonly used are Broyden mixing [123], Thomas-Fermi charge mixing [124], and Pulay mixing [101, 125]. The detailed explanations of these mixing schemes will not be discussed and interested readers are referred to the corresponding references.

2.12 Smearing

For materials that have band gap such as insulators and semiconductors, the electron densities decay smoothly near the band gap. However, metals have abrupt change of occupations from 0 to 1 at the Fermi level. This implies that any integration of functions that are discontinuous at the Fermi level will require a dense grid of k-points to achieve an acceptable accuracy. This will slow down the speed of convergence for a given set of k-points. The best way to deal with this problem is to use smearing. For instance, consider the total energy

$$E = \sum_i \int_{BZ} \epsilon_{ik}\, \Theta(\epsilon_{ik} - \mu)\, \mathrm{d}\vec{\mathbf{k}} \tag{2.47}$$

where $\Theta(\epsilon_{ik} - \mu)$ is the Dirac step function defined as

$$\Theta(x) = \begin{cases} 1 & , x \leq 0 \\ 0 & , x > 1 \end{cases} \tag{2.48}$$

Here, ϵ_{ik} is the energy of the ith band state located at wavevector $\vec{\mathbf{k}}$. Due to finite computer resources, the integral can be approximated by weighted sum over special k-points similar to (2.19)

$$E = \sum_i \sum_{k \in \mathrm{IBZ}} w_k \epsilon_{ik}\, \Theta(\epsilon_{ik} - \mu) \tag{2.49}$$

where w_k are the weighted factors of each sampled k-points. The next step is to replace the Dirac step function by a smearing function. This will result to a much faster convergence speed without destroying the accuracy. The final approximate form will be

$$E = \sum_{k \in \text{IBZ}} w_k \sum_i f_{ik} \epsilon_{ik} \tag{2.50}$$

where f_{ik} is the smearing function which is also known as the partial occupancy. In similar vein, the electron density in (1.33) is reformulated to include the partial occupancy of the KS orbitals

$$\rho(\vec{\boldsymbol{r}}) = \sum_{k \in \text{IBZ}} w_k \sum_{i=1}^{N} f_{ik}\, \phi_{ik}(\vec{\boldsymbol{r}})^* \phi_{ik}(\vec{\boldsymbol{r}}) \tag{2.51}$$

The equation above implies that for a general system, in principle, there should be a set orbitals for every possible value of k-points. There are different smearing methods used in the literature. For example, the Fermi smearing method replaces the smearing function using Fermi-Dirac distribution [126, 127]

$$f_{ik} = \frac{1}{\exp[(\epsilon_{ik} - E_f)/k_B T] + 1} \tag{2.52}$$

where E_f is the Fermi energy, k_B is the Boltzmann's constant, and T is the absolute temperature. The $k_B T$ is also called the broadening parameter that quantifies the degree of broadening of the Fermi-Dirac distribution. Figure 2.6 illustrates the partial occupancies as a function of energy. Observe that there are eigenstates above the Fermi energy when smearing is applied. This means that eigenstates are partially filled inside the Fermi surface but the discontinuity at E_f is removed by a smooth function. Note that T has no physical meaning in DFT, unless the system under study is really at finite electronic temperature. Gaussian smearing is also possible which is given by [128]

$$f_{ik} = \frac{1}{2}\left[1 - erf\left(\frac{\epsilon_{ik} - E_f}{k_B T}\right)\right] \tag{2.53}$$

where $erf(x)$ is the Gauss error function. The Methfessel–Paxton smearing method [129] approximates the smearing function f_{ik} by a hierarchy of increasingly accurate smooth functions based on Hermite polynomials. Smaller k-points are adequate to accurately describe DFT quantities.

Lastly, the linear tetrahedron method divides the irreducible Brillouin zone (IBZ) into many small tetrahedrons [130]. The energy eigenvalues ϵ_{ik} inside each tetrahedron are linearly interpolated and integrated within these tetrahedrons. This method is especially suited for transition metals and rare earths whose delicate details of the Fermi surface requires a finer resolution. It is also preferred method for band structure and DOS calculations of semiconductors. The smearing method can also be used to broaden the Dirac delta function. For instance, the δ function is replaced by a Gaussian distribution

$$\delta(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-x^2/2\sigma^2\right) \tag{2.54}$$

where σ is the broadening parameter. This broadening is very useful for calculations requiring delta functions such as density of states (DOS) in (1.14).

As a final note, the amount of broadening of smearing function must be optimized. Too large smearing might result in large error in calculation, whereas too small smearing requires a much finer k-point mesh, a computationally demanding task.

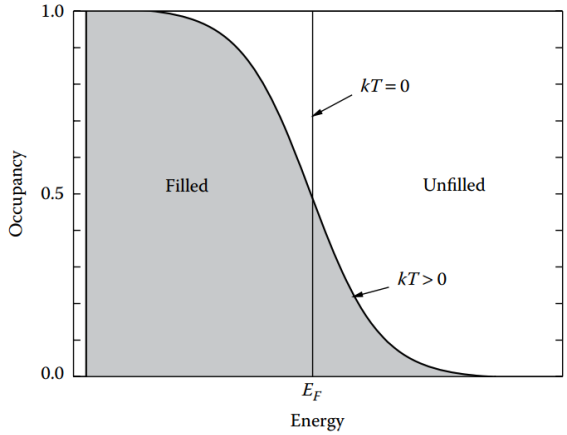

Figure 2.6 Partial occupancies near the Fermi energy using Fermi smearing. Illustration taken from [4].

Chapter Three

Software Implementation

3.1 QUANTUM ESPRESSO

The Quantum opEn Source Package for Research in Electronic Structure, Simulation, and Optimization (QUANTUM ESPRESSO) is an integrated suite of open source codes for electronic structure calculations and materials modeling based on Density functional theory, plane waves, and pseudopotential approach [131, 132]. cbands, cegterg, cdiaghg

3.2 Executables

3.3 Computational Details

Zn has 30 electrons and O has 8 electrons but the valence electrons are the only one taken into account

3.3.1 Supercell creation

3.3.2 Convergence Testing

The optimum values of the k-points, and kinetic energy-cutoffs to be used in subsequent calculations were determine from convergence tests. Initially, k-points were varied for a fixed

plane wave cutoff energy and electron density cutoff energy. These fixed values were set at $E_{\text{wfc}} = 40$ Ry and $E_{\text{rho}} = 320$ Ry, which are the suggested minimum values set by the pseudopotential used [133]. Since the c-axis of ZnO is longer, its corresponding reciprocal lattice is smaller. Hence, the grid points in the k_z direction were scaled down properly. The Monkhorst-Pack grid was set at $(k + 5) \times (k + 5) \times k$.

3.3.3 Hubbard correction parameters

3.3.4 Slab Model

3.3.5 Structural relaxation

3.3.6 scf calculation

3.3.7 bandstructure calculation

3.3.8 dos calculation

DOST COARE

Chapter Four

Results and Discussion

4.1 Convergence Tests

Convergence tests on the ZnO unit cell were done by varying the k-points, and the kinetic energy cut-offs of the Kohn-Sham orbital and the electron density, respectively. When the k-points are varied, the cut-off energies were fixed by the minimum energy set by the pseudopotential used. In this case As shown in Figure in 4.1, the pressure and total energy converges at $k \geq 2$ which corresponds to the Monkhorst-Pack grid of $7 \times 7 \times 2$. One can choose the minimum k-point required for convergence to be used for subsequent calculations. However, higher values are preferred as to make sure the convergence is guaranteed for all configuration of the system. In this thesis, the k-point used was $k = 6$ which corresponds to the Monkhorst-Pack grid $11 \times 11 \times 6$. Next,

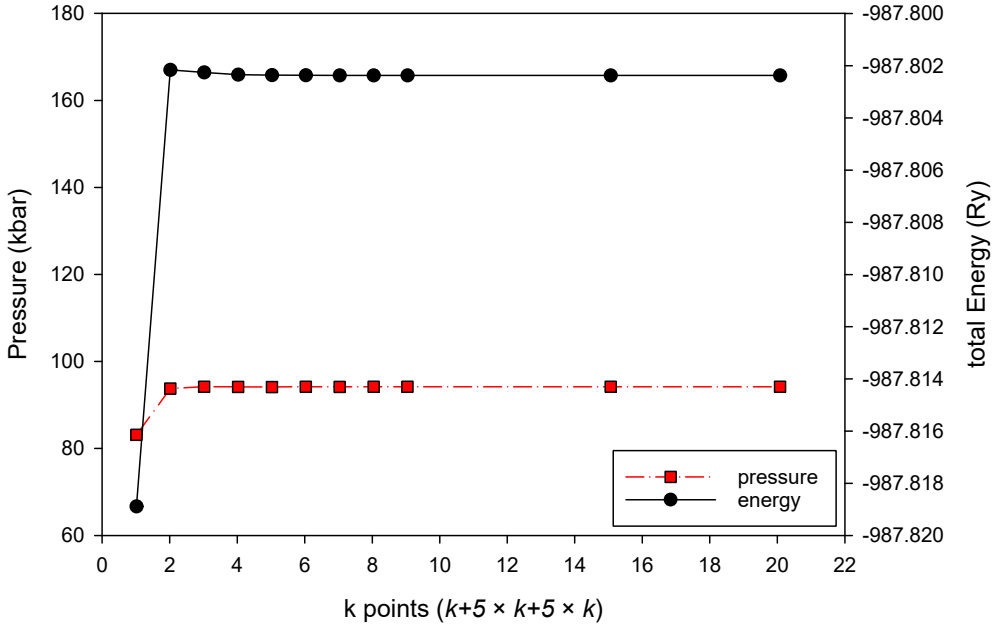

Figure 4.1 Convergence of pressure and total energies with respect to k-points. The x-axis denotes the k-point k in the Monkhorst-Pack grid of $(k+5) \times (k+5) \times k$

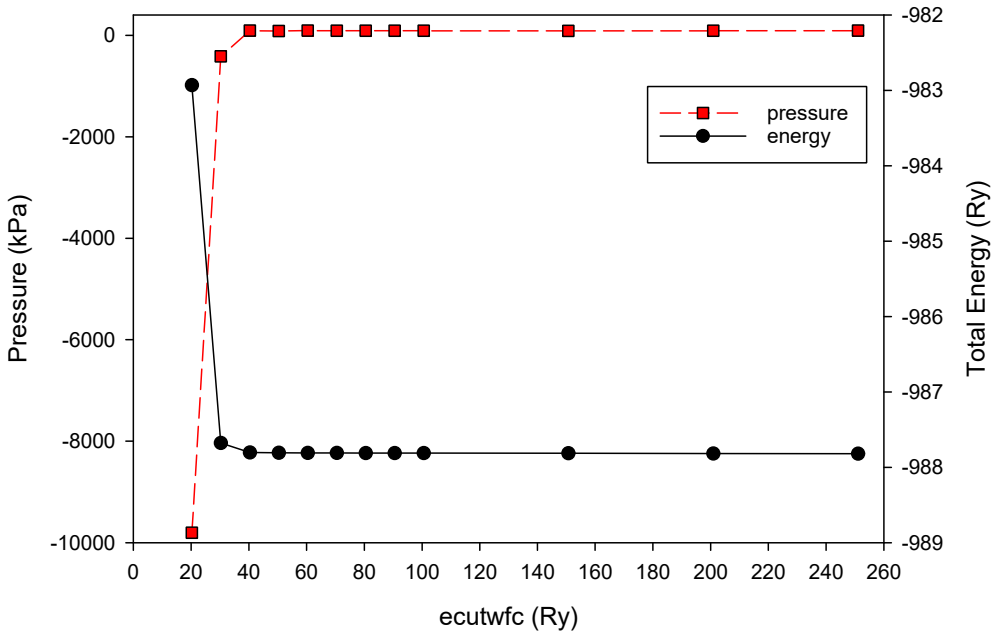

Figure 4.2 Convergence of pressure and total energies with respect to cutoff energy of the Kohn-Sham orbital

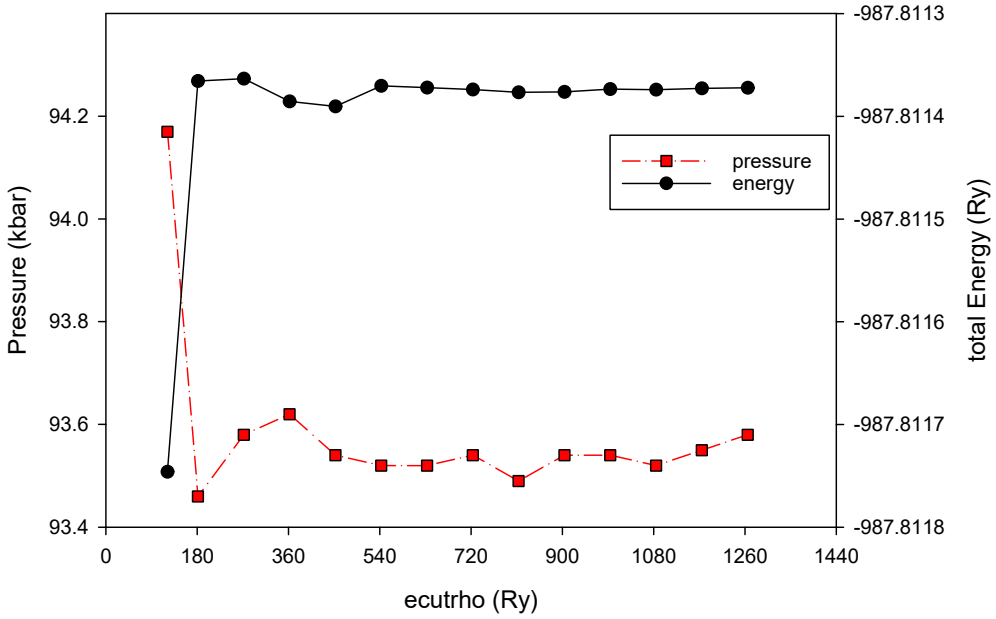

Figure 4.3 Convergence of pressure and total energies with respect to cutoff energy of the electron density

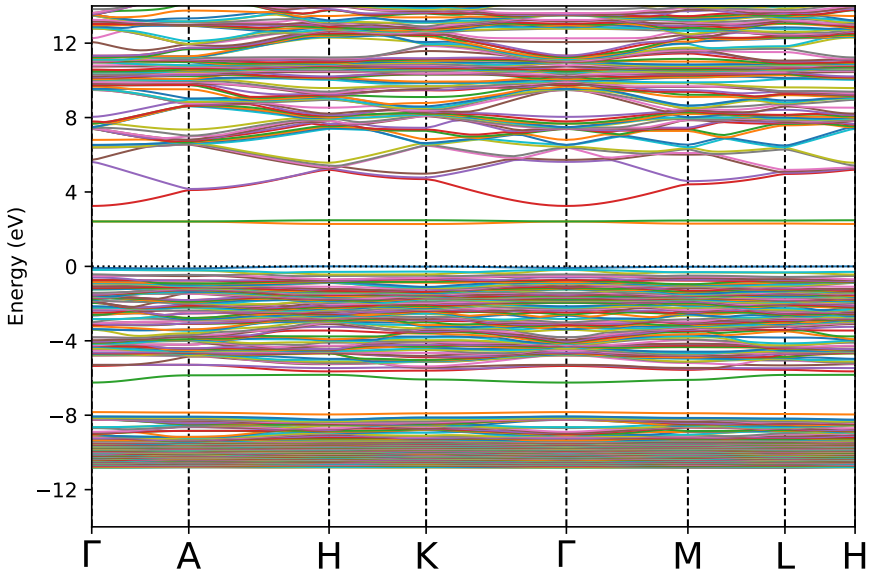

Figure 4.4 Bandstructure of Oxygen antisite

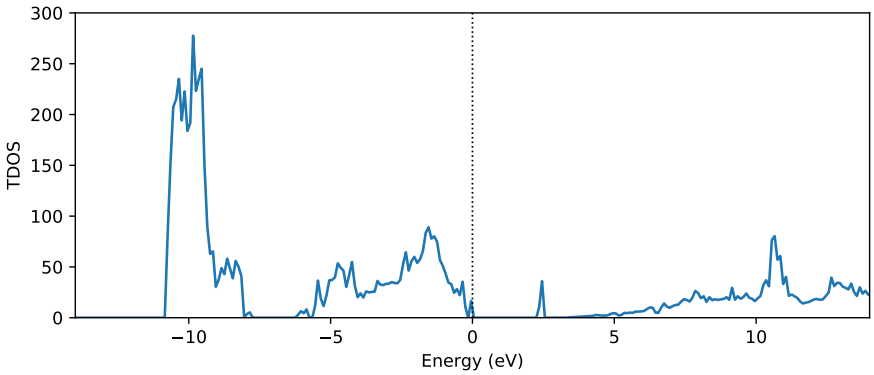

Figure 4.5 Density of States of Oxygen antisite

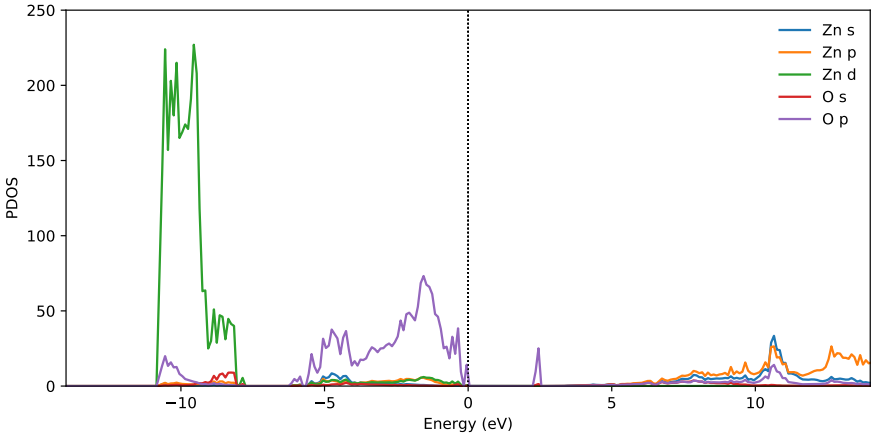

Figure 4.6 Projected Density of States of Oxygen antisite

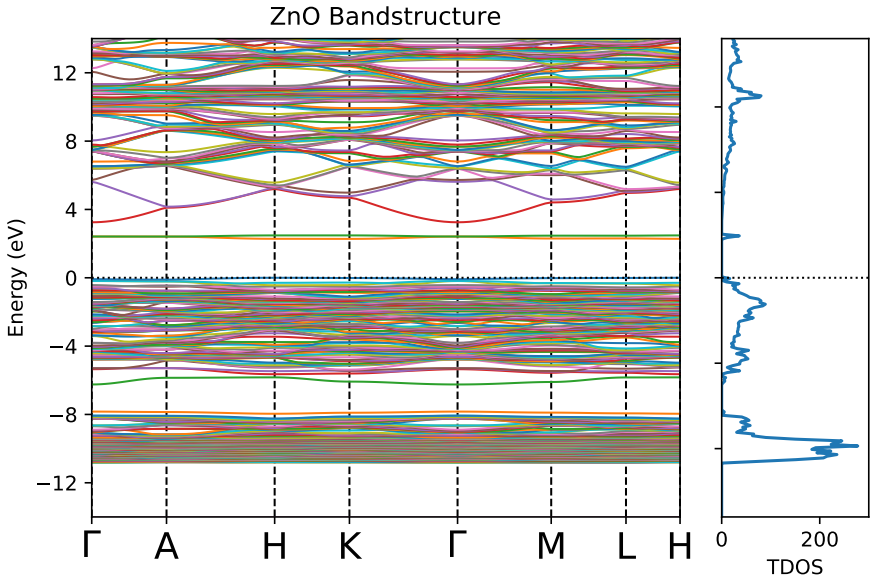

Figure 4.7 Projected Density of States of Oxygen antisite

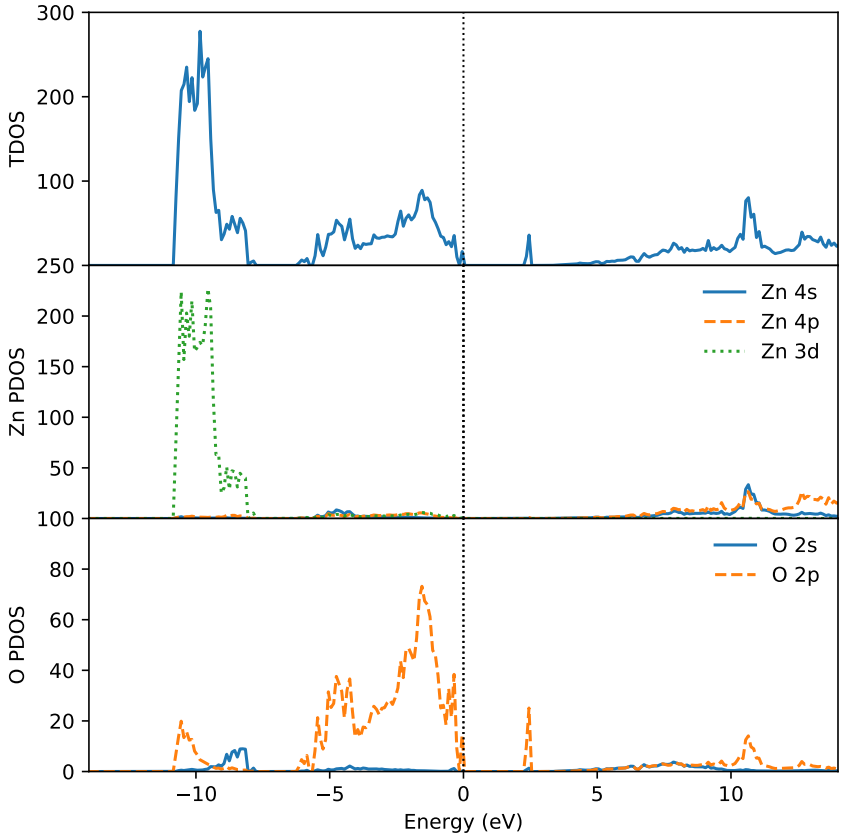

Figure 4.8 Combined Density of states of Oxygen antisite and PDOS

Chapter Five

Conclusion

REFERENCES

1. Schrödinger, E. Quantisierung als eigenwertproblem. *Annalen der physik* **385,** 437–490 (1926).

2. Herman, F. Lattice vibrational spectrum of germanium. *Journal of Physics and Chemistry of Solids* **8,** 405–418 (Jan. 1959).

3. Kittel, C. *Introduction to Solid State Physics* 704 pp. ISBN: 047141526X (John Wiley & Sons Inc, Oct. 28, 2004).

4. Lee, J. G. *Computational Materials Science:An Introduction* 351 pp. ISBN: 1498749739 (Taylor & Francis Inc, Nov. 28, 2016).

5. Neil W. Ashcroft, N. M. *Solid State Physics* 848 pp. ISBN: 0030839939 (Cengage Learning, Inc, Jan. 2, 1976).

6. Enkovaara, J. *et al.* Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *Journal of Physics: Condensed Matter* **22,** 253202 (June 2010).

7. Born, M. & Oppenheimer, R. Zur quantentheorie der molekeln. *Annalen der physik* **389,** 457–484 (1927).

8. Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Mathematical Proceedings of the Cambridge Philosophical Society* **24,** 89–110 (Jan. 1928).

9. Fock, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik* **61,** 126–148 (Jan. 1930).

10. Pauli, W. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik* **31,** 765–783 (1925).

11. Slater, J. C. The theory of complex spectra. *Physical Review* **34,** 1293 (1929).

12. Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Reviews of Modern Physics* **71,** 1253 (1999).

13. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Physical review* **136,** B864 (1964).

14. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical review* **140,** A1133 (1965).

15. Cottenier, S. *et al.* Density Functional Theory and the family of (L) APW-methods: a step-by-step introduction. *Instituut voor Kern-en Stralingsfysica, KU Leuven, Belgium* **4,** 41 (2002).

16. Woods, N. D., Payne, M. C. & Hasnip, P. J. Computing the self-consistent field in Kohn–Sham density functional theory. *Journal of Physics: Condensed Matter* **31,** 453001 (Aug. 2019).

17. Verma, P. & Truhlar, D. G. Status and Challenges of Density Functional Theory. *Trends in Chemistry* **2,** 302–318 (2020).

18. Marques, M. A., Oliveira, M. J. & Burnus, T. Libxc: A library of exchange and correlation functionals for density functional theory. *Computer Physics Communications* **183,** 2272–2281 (Oct. 2012).

19. Segala, M. & Chong, D. P. An evaluation of exchange-correlation functionals for the calculations of the ionization energies for atoms and molecules. *Journal of Electron Spectroscopy and Related Phenomena* **171,** 18–23 (Apr. 2009).

20. Dirac, P. A. M. Note on Exchange Phenomena in the Thomas Atom. *Mathematical Proceedings of the Cambridge Philosophical Society* **26,** 376–385 (July 1930).

21. Ceperley, D. M. & Alder, B. J. Ground State of the Electron Gas by a Stochastic Method. *Physical Review Letters* **45,** 566–569 (Aug. 1980).

22. Foulkes, W. M. C., Mitas, L., Needs, R. J. & Rajagopal, G. Quantum Monte Carlo simulations of solids. *Reviews of Modern Physics* **73,** 33–83 (Jan. 2001).

23. Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B* **23,** 5048–5079 (May 1981).

24. Gunnarsson, O. & Lundqvist, B. I. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Physical Review B* **13,** 4274–4298 (May 1976).

25. Gunnarsson, O., Jonson, M. & Lundqvist, B. Exchange and correlation in inhomogeneous electron systems. *Solid State Communications* **24,** 765–768 (Dec. 1977).

26. Staroverov, V. N., Scuseria, G. E., Tao, J. & Perdew, J. P. Tests of a ladder of density functionals for bulk solids and surfaces. *Physical Review B* **69** (Feb. 2004).

27. Csonka, G. I. *et al.* Assessing the performance of recent density functionals for bulk solids. *Physical Review B* **79** (Apr. 2009).

28. Harl, J., Schimka, L. & Kresse, G. Assessing the quality of the random phase approximation for lattice constants and atomization energies of solids. *Physical Review B* **81** (Mar. 2010).

29. Lee, C., Vanderbilt, D., Laasonen, K., Car, R. & Parrinello, M. Ab initio studies on the structural and dynamical properties of ice. *Physical Review B* **47,** 4863–4872 (Mar. 1993).

30. Hamann, D. R. H2O hydrogen bonding in density-functional theory. *Physical Review B* **55,** R10157–R10160 (Apr. 1997).

31. Feibelman, P. J. Lattice match in density functional calculations: ice Ih vs. β-AgI. *Physical Chemistry Chemical Physics* **10,** 4688 (2008).

32. Froyen, S. & Cohen, M. L. Structural properties of III-V zinc-blende semiconductors under pressure. *Physical Review B* **28,** 3258–3265 (Sept. 1983).

33. Tan, J., Li, Y. & Ji, G. Elastic constants and bulk modulus of semiconductors: Performance of plane-wave pseudopotential and local-density-approximation density functional theory. *Computational Materials Science* **58,** 243–247 (June 2012).

34. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **77,** 3865–3868 (Oct. 1996).

35. Perdew, J. P. *et al.* Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. *Physical Review Letters* **100** (Apr. 2008).

36. Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **38,** 3098–3100 (Sept. 1988).

37. Perdew, J. P. & Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B* **45,** 13244–13249 (June 1992).

38. Lee, C., Yang, W. & Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **37,** 785–789 (Jan. 1988).

39. Handy, N. C. & Cohen, A. J. Left-right correlation energy. *Molecular Physics* **99,** 403–412 (Mar. 2001).

40. Xu, X. & Goddard, W. A. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proceedings of the National Academy of Sciences* **101,** 2673–2677 (Feb. 2004).

41. Burke, K., Perdew, J. P. & Ernzerhof, M. Why the generalized gradient approximation works and how to go beyond it. *International Journal of Quantum Chemistry* **61,** 287–293 (1997).

42. Langreth, D. C. & Mehl, M. J. Beyond the local-density approximation in calculations of ground-state electronic properties. *Physical Review B* **28,** 1809–1834 (Aug. 1983).

43. Perdew, J. P. *et al.* Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B* **46,** 6671–6687 (Sept. 1992).

44. Proynov, E. I., Ruiz, E., Vela, A. & Salahub, D. R. Determining and extending the domain of exchange and correlation functionals. *International Journal of Quantum Chemistry* **56,** 61–78 (Feb. 1995).

45. Kim, K. & Jordan, K. D. Comparison of Density Functional and MP2 Calculations on the Water Monomer and Dimer. *The Journal of Physical Chemistry* **98,** 10089–10094 (Oct. 1994).

46. Pérez-Jordá, J. & Becke, A. A density-functional study of van der Waals forces: rare gas diatomics. *Chemical Physics Letters* **233,** 134–137 (Feb. 1995).

47. Ruiz, E., Salahub, D. R. & Vela, A. Defining the Domain of Density Functionals: Charge-Transfer Complexes. *Journal of the American Chemical Society* **117,** 1141–1142 (Jan. 1995).

48. Mori-Sánchez, P., Cohen, A. J. & Yang, W. Localization and Delocalization Errors in Density Functional Theory and Implications for Band-Gap Prediction. *Physical Review Letters* **100** (Apr. 2008).

49. Perdew, J. P. *et al.* Understanding band gaps of solids in generalized Kohn–Sham theory. *Proceedings of the National Academy of Sciences* **114,** 2801–2806 (Mar. 2017).

50. Perdew, J. P., Parr, R. G., Levy, M. & Balduz, J. L. Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy. *Physical Review Letters* **49,** 1691–1694 (Dec. 1982).

51. Perdew, J. P. & Levy, M. Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities. *Physical Review Letters* **51,** 1884–1887 (Nov. 1983).

52. Sham, L. J. & Schlüter, M. Density-Functional Theory of the Energy Gap. *Physical Review Letters* **51,** 1888–1891 (Nov. 1983).

53. Baerends, E. J. From the Kohn–Sham band gap to the fundamental gap in solids. An integer electron approach. *Physical Chemistry Chemical Physics* **19,** 15639–15656 (2017).

54. Van Setten, M. J., Weigend, F. & Evers, F. The GW-Method for Quantum Chemistry Applications: Theory and Implementation. *Journal of Chemical Theory and Computation* **9,** 232–246 (Dec. 2012).

55. Hedin, L. New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Physical Review* **139,** A796–A823 (Aug. 1965).

56. Hedin, L. & Lundqvist, S. Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids. *Solid State Physics* **23,** 1–181 (1970).

57. Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Physical Review B* **34,** 5390–5413 (Oct. 1986).

58. Godby, R. W., Schlüter, M. & Sham, L. J. Accurate Exchange-Correlation Potential for Silicon and Its Discontinuity on Addition of an Electron. *Physical Review Letters* **56,** 2415–2418 (June 1986).

59. Samsonidze, G., Jain, M., Deslippe, J., Cohen, M. L. & Louie, S. G. Simple Approximate Physical Orbitals for GW Quasiparticle Calculations. *Physical Review Letters* **107** (Oct. 2011).

60. Deslippe, J., Samsonidze, G., Jain, M., Cohen, M. L. & Louie, S. G. Coulomb-hole summations and energies for GW calculations with limited number of empty orbitals: A modified static remainder approach. *Physical Review B* **87** (Apr. 2013).

61. Gao, W., Xia, W., Gao, X. & Zhang, P. Speeding up GW Calculations to Meet the Challenge of Large Scale Quasiparticle Predictions. *Scientific Reports* **6** (Nov. 2016).

62. van Schilfgaarde, M., Kotani, T. & Faleev, S. Quasiparticle Self-Consistent GW Theory. *Physical Review Letters* **96** (June 2006).

63. Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *The Journal of Chemical Physics* **98,** 1372–1377 (Jan. 1993).

64. Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *The Journal of Chemical Physics* **105,** 9982–9985 (Dec. 1996).

65. Adamo, C. & Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics* **110,** 6158–6170 (Apr. 1999).

66. Paier, J., Marsman, M. & Kresse, G. Why does the B3LYP hybrid functional fail for metals? *The Journal of Chemical Physics* **127,** 024103 (July 2007).

67. Heyd, J., Scuseria, G. E. & Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *The Journal of Chemical Physics* **118,** 8207–8215 (May 2003).

68. Krukau, A. V., Vydrov, O. A., Izmaylov, A. F. & Scuseria, G. E. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of Chemical Physics* **125,** 224106 (Dec. 2006).

69. Tao, J., Perdew, J. P., Staroverov, V. N. & Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Physical Review Letters* **91** (Sept. 2003).

70. Tran, F. & Blaha, P. Accurate Band Gaps of Semiconductors and Insulators with a Semilocal Exchange-Correlation Potential. *Physical Review Letters* **102** (June 2009).

71. Perdew, J. P., Kurth, S., Zupan, A. & Blaha, P. Accurate Density Functional with Correct Formal Properties: A Step Beyond the Generalized Gradient Approximation. *Physical Review Letters* **82,** 2544–2547 (Mar. 1999).

72. Sun, J., Ruzsinszky, A. & Perdew, J. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Physical Review Letters* **115** (July 2015).

73. Singh, D. J. Structure and optical properties of high light output halide scintillators. *Physical Review B* **82** (Oct. 2010).

74. Singh, D. J. Electronic structure calculations with the Tran-Blaha modified Becke-Johnson density functional. *Physical Review B* **82** (Nov. 2010).

75. Himmetoglu, B., Floris, A., de Gironcoli, S. & Cococcioni, M. Hubbard-corrected DFT energy functionals: The LDA+U description of correlated systems. *International Journal of Quantum Chemistry* **114,** 14–49 (July 2013).

76. Shen, Z.-X. *et al.* Electronic structure of NiO: Correlation and band effects. *Physical Review B* **44,** 3604–3626 (Aug. 1991).

77. Tolba, S. A., Gameel, K. M., Ali, B. A., Almossalami, H. A. & Allam, N. K. in *Density Functional Calculations - Recent Progresses of Theory and Application* (InTech, May 2018).

78. Liechtenstein, A. I., Anisimov, V. I. & Zaanen, J. Density-functional theory and strong interactions: Orbital ordering in Mott-Hubbard insulators. *Physical Review B* **52,** R5467–R5470 (Aug. 1995).

79. Dudarev, S. L., Botton, G. A., Savrasov, S. Y., Humphreys, C. J. & Sutton, A. P. Electron-energy-loss spectra and the structural stability of nickel oxide:An LSDA+U study. *Physical Review B* **57,** 1505–1509 (Jan. 1998).

80. Cococcioni, M. & de Gironcoli, S. Linear response approach to the calculation of the effective interaction parameters in the LDA+U method. *Physical Review B* **71** (Jan. 2005).

81. Himmetoglu, B., Wentzcovitch, R. M. & Cococcioni, M. First-principles study of electronic and structural properties of CuO. *Physical Review B* **84** (Sept. 2011).

82. Kulik, H. J., Cococcioni, M., Scherlis, D. A. & Marzari, N. Density Functional Theory in Transition-Metal Chemistry: A Self-Consistent Hubbard U Approach. *Physical Review Letters* **97** (Sept. 2006).

83. Dompablo, M. E. A.-d., Morales-Garcia, A. & Taravillo, M. DFT + U calculations of crystal lattice, electronic structure, and phase stability under pressure of TiO_2 polymorphs. *The Journal of Chemical Physics* **135,** 054503 (Aug. 2011).

84. Tran, F., Blaha, P., Schwarz, K. & Novák, P. Hybrid exchange-correlation energy functionals for strongly correlated electrons: Applications to transition-metal monoxides. *Physical Review B* **74** (Oct. 2006).

85. Cottenier, S. *Density Functional Theory and the family of (L) APW-methods: a step-by-step introduction* ISBN: 9789080721517 (Instituut voor Kern-en Stralingsfysica, KU Leuven, Belgium, 2013).

86. Junquera, J., Paz, Ó., Sánchez-Portal, D. & Artacho, E. Numerical atomic orbitals for linear-scaling calculations. *Physical Review B* **64** (Nov. 2001).

87. Louwerse, M. J. & Rothenberg, G. Transferable basis sets of numerical atomic orbitals. *Physical Review B* **85** (Jan. 2012).

88. Slater, J. C. Analytic Atomic Wave Functions. *Physical Review* **42,** 33–43 (Oct. 1932).

89. Boys, S. F. & Egerton, A. C. Electronic wave functions - I. A general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **200,** 542–554 (Feb. 1950).

90. Segall, M. D. *et al.* First-principles simulation: ideas, illustrations and the CASTEP code. *Journal of Physics: Condensed Matter* **14,** 2717–2744 (Mar. 2002).

91. Slater, J. C. An Augmented Plane Wave Method for the Periodic Potential Problem. *Physical Review* **92,** 603–608 (Nov. 1953).

92. Sjöstedt, E., Nordström, L. & Singh, D. An alternative way of linearizing the augmented plane-wave method. *Solid State Communications* **114,** 15–20 (Mar. 2000).

93. Andersen, O. K. Linear methods in band theory. *Physical Review B* **12,** 3060–3083 (Oct. 1975).

94. Kohanoff, J. *Electronic structure calculations for solids and molecules: theory and computational methods* ISBN: 9780511755613 (Cambridge University Press, Cambridge, 2006).

95. Requist, R. & Pankratov, O. Generalized Kohn-Sham system in one-matrix functional theory. *Physical Review B* **77** (June 2008).

96. Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *Journal of Computational Physics* **17,** 87–94 (Jan. 1975).

97. Crouzeix, M., Philippe, B. & Sadkane, M. The Davidson Method. *SIAM Journal on Scientific Computing* **15,** 62–76 (Jan. 1994).

98. Wood, D. M. & Zunger, A. A new method for diagonalising large matrices. *Journal of Physics A: Mathematical and General* **18,** 1343–1359 (June 1985).

99. Rayson, M. & Briddon, P. Rapid iterative method for electronic-structure eigenproblems using localised basis functions. *Computer Physics Communications* **178,** 128–134 (Jan. 2008).

100. Der Vorst, H. A. V. & Dekker, K. Conjugate gradient type methods and preconditioning. *Journal of Computational and Applied Mathematics* **24,** 73–87 (Nov. 1988).

101. Kresse, G. & Furthmüller, J. Efficient iterative schemes forab initiototal-energy calculations using a plane-wave basis set. *Physical Review B* **54,** 11169–11186 (Oct. 1996).

102. Phillips, J. C. Energy-Band Interpolation Scheme Based on a Pseudopotential. *Physical Review* **112,** 685–695 (Nov. 1958).

103. Phillips, J. C. & Kleinman, L. New Method for Calculating Wave Functions in Crystals and Molecules. *Physical Review* **116,** 287–294 (Oct. 1959).

104. Cohen, M. L. & Heine, V. in *Solid State Physics* 37–248 (Elsevier, 1970).

105. Jochym, D. B. *Development of non-local density functional methods* PhD thesis (Durham University, 2008). eprint: http://etheses.dur.ac.uk/2174/.

106. Troullier, N. & Martins, J. A straightforward method for generating soft transferable pseudopotentials. *Solid State Communications* **74,** 613–616 (May 1990).

107. Fuchs, M. & Scheffler, M. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory. *Computer Physics Communications* **119,** 67–98 (June 1999).

108. Hamann, D. R., Schlüter, M. & Chiang, C. Norm-Conserving Pseudopotentials. *Physical Review Letters* **43,** 1494–1497 (Nov. 1979).

109. Troullier, N. & Martins, J. L. Efficient pseudopotentials for plane-wave calculations. *Physical Review B* **43,** 1993–2006 (Jan. 1991).

110. Rappe, A. M., Rabe, K. M., Kaxiras, E. & Joannopoulos, J. D. Optimized pseudopotentials. *Physical Review B* **41,** 1227–1230 (Jan. 1990).

111. Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B* **41,** 7892–7895 (Apr. 1990).

112. Blöchl, P. E. Projector augmented-wave method. *Physical Review B* **50,** 17953–17979 (Dec. 1994).

113. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B* **59,** 1758–1775 (Jan. 1999).

114. Rostgaard, C. The Projector Augmented-wave Method. arXiv: 0910.1921v2 (Oct. 10, 2009).

115. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Physical Review B* **13,** 5188–5192 (June 1976).

116. Hellman, H. Einführung in die Quantenchemie. *Franz Deuticke, Leipzig,* 285 (1937).

117. Feynman, R. P. Forces in Molecules. *Physical Review* **56,** 340–343 (Aug. 1939).

118. Curtis, F. E. & Que, X. A quasi-Newton algorithm for nonconvex, nonsmooth optimization with global convergence guarantees. *Mathematical Programming Computation* **7,** 399–428 (May 2015).

119. Dai, Y. H. & Yuan, Y. A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. *SIAM Journal on Optimization* **10,** 177–182 (Jan. 1999).

120. Probert, M. Improved algorithm for geometry optimisation using damped molecular dynamics. *Journal of Computational Physics* **191,** 130–146 (Oct. 2003).

121. Kerker, G. P. Efficient iteration scheme for self-consistent pseudopotential calculations. *Physical Review B* **23,** 3082–3084 (Mar. 1981).

122. Johnson, D. D. Modified Broyden's method for accelerating convergence in self-consistent calculations. *Physical Review B* **38,** 12807–12813 (Dec. 1988).

123. Broyden, C. G. A class of methods for solving nonlinear simultaneous equations. *Mathematics of Computation* **19,** 577–577 (1965).

124. Raczkowski, D., Canning, A. & Wang, L. W. Thomas-Fermi charge mixing for obtaining self-consistency in density functional calculations. *Physical Review B* **64** (Sept. 2001).

125. Pulay, P. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters* **73,** 393–398 (July 1980).

126. Dirac, P. A. M. On the theory of quantum mechanics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **112,** 661–677 (Oct. 1926).

127. Fermi, E. Sulla quantizzazione del gas perfetto monoatomico. *Rendiconti Lincei* **3,** 145–149 (1926).

128. Fu, C. .-L. & Ho, K. .-M. First-principles calculation of the equilibrium ground-state properties of transition metals: Applications to Nb and Mo. *Physical Review B* **28,** 5480–5486 (Nov. 1983).

129. Methfessel, M. & Paxton, A. T. High-precision sampling for Brillouin-zone integration in metals. *Physical Review B* **40,** 3616–3621 (Aug. 1989).

130. Blöchl, P. E., Jepsen, O. & Andersen, O. K. Improved tetrahedron method for Brillouin-zone integrations. *Physical Review B* **49,** 16223–16233 (June 1994).

131. Giannozzi, P. *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter* **21,** 395502 (Sept. 2009).

132. Giannozzi, P. *et al.* Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter* **29,** 465901 (Oct. 2017).

133. Garrity, K. F., Bennett, J. W., Rabe, K. M. & Vanderbilt, D. Pseudopotentials for high-throughput DFT calculations. *Computational Materials Science* **81,** 446–452 (Jan. 2014).

APPENDIX

Appendix A

Reciprocal Lattice and Brillouin Zone

Reciprocal lattice vectors of a lattice are defined to be the wavevectors $\vec{\mathbf{G}}$ that satisfy

$$\exp\left(i\vec{\mathbf{G}}\cdot\vec{\mathbf{R}}\right) = 1 \tag{A.1}$$

for any lattice translation vector $\vec{\mathbf{R}}$ given by

$$\vec{\mathbf{R}} = n_1\vec{\mathbf{a}}_1 + n_2\vec{\mathbf{a}}_2 + n_3\vec{\mathbf{a}}_3 \tag{A.2}$$

Here n_1, n_2, n_3 are arbitrary integers and $\vec{\mathbf{a}}_1, \vec{\mathbf{a}}_2, \vec{\mathbf{a}}_3$ are the primitive lattice vectors of the direct (real) lattice. Similarly, the reciprocal lattice vector can be resolved into its components

$$\vec{\mathbf{G}} = h\vec{\mathbf{b}}_1 + k\vec{\mathbf{b}}_2 + l\vec{\mathbf{b}}_3 \tag{A.3}$$

where h, k, l are also arbitrary integers and $\vec{\mathbf{b}}_1, \vec{\mathbf{b}}_2, \vec{\mathbf{b}}_3$ are the primitive lattice vectors of the reciprocal lattice. The integers h, k, l constitute the so called Miller indices (hkl). Satisfying (A.1) means that

$$\vec{\mathbf{a}}_i \cdot \vec{\mathbf{b}}_j = 2\pi\delta_{ij}$$

It can be shown that the reciprocal vectors b_j can be constructed entirely in terms of real vectors a_i

$$\vec{\mathbf{b}}_1 = 2\pi\frac{\vec{\mathbf{a}}_2 \times \vec{\mathbf{a}}_3}{V} \tag{A.4}$$

$$\vec{\mathbf{b}}_2 = 2\pi\frac{\vec{\mathbf{a}}_3 \times \vec{\mathbf{a}}_1}{V} \tag{A.5}$$

$$\vec{\mathbf{b}}_3 = 2\pi\frac{\vec{\mathbf{a}}_1 \times \vec{\mathbf{a}}_2}{V} \tag{A.6}$$

where $V = \vec{\mathbf{a}}_1 \cdot (\vec{\mathbf{a}}_2 \times \vec{\mathbf{a}}_3)$ is the volume of the unit cell. Reciprocal lattice and real lattice have important relationship. It can be shown that the reciprocal lattice vector defined by the Miller indices (hkl) are perpendicular to the direct (real) lattice plane whose intercepts are the reciprocals of Miller indices (hkl).

The Brillouin Zone (BZ) is defined to be the primitive unit cell of the reciprocal lattice. The first BZ is the Wigner-Seitz cell around the wavevector $\vec{\mathbf{k}} = 0$, the origin of the reciprocal space, which is defined to be the region of reciprocal space closer to $\vec{\mathbf{k}} = 0$ than to any other reciprocal lattice vector. Similarly, all wavevectors that are second closest to $\vec{\mathbf{k}} = 0$, constitute the second BZ, and so forth. The Wigner-Seitz cell contains only one reciprocal lattice point in it. This means that first BZ is constructed as the smallest volume entirely closed by a set of planes that are the perpendicular bisectors of the origin to each reciprocal lattice vectors. The volume of the first BZ is related to the volume of the primitive unit cell of the real lattice

$$\Omega = \frac{(2\pi)^3}{V} \tag{A.7}$$

Brillouin zone plays a vital role in calculation of periodic systems. Owing to the periodic nature of the reciprocal lattice, any wavevectors $\vec{\mathbf{k}}$ outside the first BZ can be folded back into the first BZ by a reciprocal lattice vector $\vec{\mathbf{G}}$. In addition, any wavevectors that are shifted by a reciprocal lattice vector are all equivalent. Hence, any physical properties that have periodicity of the crystal system can be represented entirely inside the first Brillouin zone.