
A Method of Real Space Spin Density Functional Theory

Chenyi Zhou
Centre for the Physics of Materials
Department of Physics
McGill University
Montréal, Québec
Canada
2016

A Thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Master of Science

CONTENTS

Abstract	vii
Résumé	viii
Statement of Originality	ix
Acknowledgments	x
1 Introduction	1
2 Theories of Electronic Structure Calculations	4
2.1 First Principles Theories	4
2.1.1 Basics of the Interacting Many-Body Problem	4
2.1.2 Hartree-Fock Approximation	8
2.1.3 Kohn-Sham DFT Formalism	10
2.2 Exchange-Correlation Functionals	15
2.2.1 XC Potentials for Collinear Spin	16
2.2.2 XC Potentials for Non-Collinear Spin	18
2.3 Mixing Techniques - Achieving Self-Consistency	20
2.4 Pseudopotential Method	24
2.4.1 Construction of l -Dependent Pseudopotentials	25
2.4.2 Construction of j -Dependent Relativistic Pseudopotentials . .	31
2.5 Summary	36
3 Numerical Methods	37
3.1 Real-space Representation of Kohn-Sham Equation	37

3.1.1	Finite-Difference Method	38
3.1.2	Ionic Potential	42
3.1.3	Hartree Potential	43
3.2	Eigensolvers	45
3.2.1	Power Method and Inverse Iteration	46
3.2.2	Chebyshev-Filtered Subspace Iteration	49
4	Tests and Applications	52
4.1	Collinear-spin Calculations: Ni and Fe	52
4.2	Noncollinear-spin Calculations: Fe ₅ and Cr ₃ clusters	55
4.3	Spin-orbit Coupling Calculations: SnTe and TaAs	56
5	Conclusions	60
A	SPIN-ORBIT COUPLING WITH LCAO BASIS	62
B	LANCZOS ALGORITHM	65
	References	67

LIST OF FIGURES

4.1	Bandstructure of Ni (FCC) calculated with our real-space method.	53
4.2	Bandstructure of Fe (BCC) calculated with our real-space method.	54
4.3	FBZs of FCC and BCC lattices.	54
4.4	Non-collinear spin texture of Cr_3 (a) and Fe_5 (b) clusters.	55
4.5	The conventional cell and the corresponding first Brillouin zone of SnTe.	57
4.6	Body-centered tetragonal structure of TaAs and its first Brillouin zone. Figure from Ref.[1].	57
4.7	Bandstructure of SnTe calculated with our real-space method and with PAW. . . .	58
4.8	Bandstructure of TaAs calculated with our real-space method and with PAW. . . .	58

LIST OF TABLES

4.1	Magnetic moment per atom $M_{at}(\mu_B)$ from experiment and calculated by different DFT methods.	53
4.2	The calculated and previously reported atomic magnetic moment (M_{at}, θ) of each atom in the Fe ₅ pentamer. The numbers are ordered according to atom 1-5 (Fig.4.4b)	56

Abstract

Phenomena related to the spin degree of freedom of electrons are an important aspect in material physics. They can play significant roles in both magnetic and nonmagnetic materials. It is thus desirable to have a material simulation tool that is capable of computing spin related quantities from first principles.

In this thesis we report our implementation of spin related calculations in the RESCU real-space DFT code[2], including collinear and noncollinear spin-polarized calculations, and spin-orbit coupling (SOC) calculation. The main eigensolver of RESCU is based on the Chebyshev-filtered subspace iteration method proposed in Ref.[3]. In this work the filter operator is extended as the Kohn-Sham effective Hamiltonian is spanned in the spin-space. As the inclusion of SOC makes spin no longer a good quantum number, it is implemented within the noncollinear spin framework. To account for SOC we follow the relativistic pseudopotential scheme proposed in Ref.[4], where the SOC term is explicitly separated from the scalar term. As usual the resulting pseudopotential is of the fully separable form and hence is favorable for numerical implementations.

To test our spin-polarized implementation, we have performed simulations on Fe and Ni bulks, and Fe₅ and Cr₃ clusters. We have compared our results with those reported before and found good agreements. To validate our SOC implementation we calculate the bandstructures of SnTe and TaAs and again we find good agreements with results obtained with a state-of-art package VASP[5].

Résumé

Phénomènes liés au degré de spin d'électrons sont un aspect important en la physique des matériaux. Ils peuvent jouer un rôle important dans les deux matériaux magnétiques et non magnétiques. Il est donc souhaitable de disposer d'un outil de simulation matériau qui est capable de calculer les quantités de spin liées à partir des lois fondamentales de la physique.

Dans cette thèse, nous rapportons notre implémentation des spin calculs connexes dans le code RESCUE[2], incluant le collinear et le noncollinear spin polarisés calculs, et le couplage spin-orbite (SOC) calcul. La eigensolver principale est basé sur le "Chebyshev-filtered subspace iteration" méthode premièrement proposé dans Ref.[3]. Dans ce travail, l'opérateur de filtre est prolongé en fait que l'Hamiltonien effectif du Kohn-Sham est enjambée dans l'espace de spin. Comme l'inclusion de SOC fait le spin non pas plus un bon nombre quantique, il est implémenté dans le cadre de spin noncollinear. Pour considérer le SOC nous suivons le pseudopotential schéma relativiste, où le terme du SOC est explicitement séparé du terme scalaire. Comme d'habitude le pseudopotential résultant est de la forme entièrement séparable et est favorable à la mise en oeuvre numérique donc.

Pour tester notre implémentation polarisé en spin, nous avons effectué des simulations sur Fe et Ni vracs, et les grappes Fe₅ et Cr₃. Nous avons comparé nos résultats avec ceux rapportés avant et trouvé de bons accords. Pour valider notre implémentation du SOC, nous calculons les structures de la bande de SnTe et TaAs et encore, nous trouvons de bons accords avec les résultats obtenus avec un package state-of-art[5].

Statement of Originality

The work of this thesis consists in the implementation of spin polarized density functional theory, based on the recently developed real-space DFT code RESCU[2]. My contributions include:

- The implementation of collinear-spin, noncollinear-spin, and spin-orbit coupling calculations within the RESCU framework. The implementation is done both with the LCAO basis set and directly on the real-space grid.
- A post-analysis code (based on the output of the NANOBASE software) for generating the j -dependent pseudopotentials to be used in spin-orbit coupling calculations.
- Routines for operating complex matrices are added to the MATLAB-ScaLAPACK interface used by RESCU.
- Applications and tests of the written code that show good agreement with results reported before.

Acknowledgments

First of all, I would like to thank my supervisor, Professor Hong Guo, for giving me this research opportunity. Without his guidance and support, life would be much tougher. I thank Vincent Michaud-Rioux and Dr. Zhang Lei for helping me understand the real-space DFT method and the polynomial-filer algorithms. I have taken a lot of benefits from their codes. I also owe thanks to Dr. Xiao Yang for many helpful discussions on the topic of non-collinear spin theories.

List of Abbreviations

DFT	Density Functional Theory
GGA	Generalized Gradient Approximation
HF	Hartree-Fock
KS	Kohn-Sham
LCAO	Linear Combination of Atomic Orbitals
L(S)DA	Local (Spin) Density Approximation
PS	Pseudopotential
SC	Self-Consistent
SOC	Spin-Orbit Coupling
XC	Exchange-Correlation

Introduction

With the highly advanced performances of today's computation technologies, we are in a good position to compute various properties of materials from first-principles, that is, based on the fundamental physics at the atomistic level without resorting to any free parameters. This ambitious goal could not be achieved without the help of the density functional theory (DFT)[6, 7], which casts the intractable interacting many-electron problem into an effective noninteracting problem while without much loss of accuracy and generality.

In a conventional DFT method, one expands the Kohn-Sham Hamiltonian in terms of a chosen basis set such as planewaves or atomic centered orbitals, and the Kohn-Sham equation is thus transformed into a matrix eigenpair problem within the subspace spanned by the basis functions. The existent state-of-art DFT simulation packages are mostly planewave based, but the planewave basis is often unfavorable for simulating finite systems such as clusters or big polymers, because the planewave approach imposes the periodic boundary condition on the simulation box, that is, any system is artificially extended in each dimension periodically. For the same reason, one often has a hard time using the planewave basis to simulate charged systems, where a compensating background charge is demanded in order to prevent the total energy divergence. From the numerical implementation point of view, planewave DFT often requires massive global communications among processors due to the Fast Fourier Transform operations.

As pointed out by Chelikowsky *et. al.*, all the aforementioned disadvantages about planewave DFT can be circumvented if one chooses to solve the Kohn-Sham equation directly on a real-space grid using some high order finite-difference method[8]. Such kind of real-space DFT method has received successive interests during the past decade. Recently an efficient real-space eigensolver has been proposed based on the Chebyshev polynomial-filtering method[3] and further improvements (in both algorithm and coding aspects) have been achieved to make this method particularly suitable for simulating large scale systems[2]. Successful numerical simulations of supercells containing more than 5,000 atoms have been reported[2] and therein the algorithm shows a promising overall scaling.

Among all the applications of DFT, one finds it very useful for studying the magnetism related phenomena. The spontaneous magnetization is a true quantum many-body effect: in some systems electrons can lower their total energy by getting localized and forming spin alignments or anti-alignments; this occurs when the resulting quantum exchange effect due to Pauli's exclusion principle overplays the localization effect which leads to the increase of kinetic energy due to Heisenberg's uncertainty principle (c.f. Eq.2.11). Therefore DFT becomes a powerful tool for this kind of issue since it well accounts for the underlying many-body effects. The tendency toward magnetism is considerably enhanced in lower-dimensional systems such as metallic surfaces and interfaces, films, wires, and clusters. These magnetic systems have received considerable experimental and theoretical attention in recent years and have been extensively studied by spin-polarized DFT calculations. One of the major tasks of this work is to implement the spin-polarized calculation within the recently developed real-space DFT framework[2].

Another topic we are interested in is the material properties associated with the spin-orbit coupling (SOC) effect. SOC is a relativistic effect which tends to be stronger as the atomic number increases. The SOC related physics has become a hotspot in today's material science, especially in novel optoelectronic materials[9], thermoelec-

tric materials[10], and topological insulators or semimetals[11, 12]. SOC effects are also of great importance in regular magnetic systems. For example, the favorable magnetization direction or spin configuration of a magnetic material is determined by the SOC part of the total energy of such configuration.

At the first glance it might be not so obvious how to describe SOC within DFT since the standard Kohn-Sham equations are nonrelativistic. However, as relativistic effects mainly enter the picture through the high velocity core electrons, they can be taken care of by the ionic pseudopotentials which Kohn-Sham equations are often implemented in conjunction with[13]. It is also worth mentioning that once SOC is taken into account, spin is no longer a good quantum number; therefore the SOC calculation should be implemented within the noncollinear spin framework.

The thesis is organized as follows. Chapter 2 contains a very brief review of the fundamental many-body physics and then the Kohn-Sham density functional theory is introduced. The remaining of chapter 2 is devoted to completing the Kohn-Sham DFT numerical scheme, including the method for approximating the exchange and correlation functionals, the mixing methods, and the pseudopotential methods. In chapter 3 we demonstrate how to discretize Kohn-Sham equations in the real-space. The eigensolver numerical recipe is also presented thereafter. A few tests and applications of our implementation of spin-polarized DFT are presented in chapter 4.

Theories of Electronic Structure Calculations

2.1 *First Principles Theories*

In what follows we first introduce the basic language (Green functions) of the many-body physics based on the most fundamental quantum mechanics. This makes clear what are the main tasks in solving a many-body problem and how difficulties arise. Then we derive the Hartree-Fock method using the Green functions, which makes us aware of the two different many-body effects, i.e. exchange and correlation, and why both of them are important in describing electrons in solid-state systems. Finally we present the Density Functional Theory (DFT) which takes a different track of formulating the many-body problem. The connection between DFT and the Green function method is discussed briefly.

2.1.1 *Basics of the Interacting Many-Body Problem*

The basic interacting many-electron Hamiltonian that we are trying to address reads[14, 15] ($\hbar = e = m_e = 1$)

$$H = H_0 + V_{ee}$$

$$H_0 = -\frac{1}{2} \sum_{\sigma} \int d^3r d^3r' \hat{\Psi}_{\sigma}^{\dagger}(\mathbf{r}) \nabla^2 \hat{\Psi}_{\sigma}(\mathbf{r}') + \sum_{\sigma} \int d^3r (V_H(\mathbf{r}) + V_{\text{ext}}(\mathbf{r})) \hat{\Psi}_{\sigma}^{\dagger}(\mathbf{r}) \hat{\Psi}_{\sigma}(\mathbf{r})$$

$$\begin{aligned}
V_{ee} &= \frac{1}{2} \sum_{\sigma_1, \sigma_2} \int d^3 r_1 d^3 r_2 \hat{\Psi}_{\sigma_1}^\dagger(\mathbf{r}_1) \hat{\Psi}_{\sigma_2}^\dagger(\mathbf{r}_2) v(\mathbf{r}_2 - \mathbf{r}_1) \hat{\Psi}_{\sigma_2}(\mathbf{r}_2) \hat{\Psi}_{\sigma_1}(\mathbf{r}_1) \\
&\quad - \sum_{\sigma} \int d^3 r V_H(\mathbf{r}) \hat{\Psi}_{\sigma}^\dagger(\mathbf{r}) \hat{\Psi}_{\sigma}(\mathbf{r}) \\
V_H(\mathbf{r}) &= \int d^3 r' v(\mathbf{r} - \mathbf{r}') \sum_{\sigma} \rho_{\sigma}(\mathbf{r}') , \quad v(\mathbf{r} - \mathbf{r}') \equiv \frac{1}{|\mathbf{r} - \mathbf{r}'|},
\end{aligned}$$

where ρ denotes electron density, and $\hat{\Psi}_{\sigma}^\dagger(\mathbf{r})$ ($\hat{\Psi}_{\sigma}(\mathbf{r})$) creates (annihilates) an electron with spin σ at real-space point $\mathbf{r}$. The unperturbed Hamiltonian H_0 is composed of the kinetic energy and the Hartree term. The perturbation V_{ee} is due to the two-body coulomb interaction (with the Hartree contribution deducted). We place the Hartree interaction $\hat{\Psi}_{\sigma}^\dagger(\mathbf{r}) V_H(\mathbf{r}) \hat{\Psi}_{\sigma}(\mathbf{r})$ in H_0 instead of V_{ee} because it only gives a static mean-field contribution and is easy to handle. The external potential $V_{\text{ext}}(\mathbf{r})$ comes from the nuclei in the system. In this section we do not include the relativistic effects, and the Zeeman-like transverse interaction is out of the scope of usual many-electron problems[14]. Therefore, there is no spin flipping mechanism in H .

As we are interested in the ground state properties of the system, it is suitable to work in the zero-temperature formalism[16] and the central quantity to compute is the Green function

$$G(x, x') = \begin{cases} -i \langle N | \hat{\Psi}(x) \hat{\Psi}^\dagger(x') | N \rangle & \text{for } t > t' \\ i \langle N | \hat{\Psi}^\dagger(x') \hat{\Psi}(x) | N \rangle & \text{for } t < t' \end{cases} \quad (2.1)$$

where $\hat{\Psi}_{\sigma}^\dagger(x)$ and $\hat{\Psi}_{\sigma}(x)$ are operators in the Heisenberg picture, and $x \equiv (\mathbf{r}, t, \sigma)$ for notation brevity. Here $|N\rangle$ represents the *exact* N-electron ground state of the full Hamiltonian H . The physical meaning of the Green function is clear: for $t' > t$ it is the probability that a hole created at x will propagate to x' and for $t > t'$ it is the probability that an electron added at x' will propagate to x . Therefore, a lot of information can be extracted from the Green function, such as the one-electron excitation spectrum, particle density distribution, and ground state energy $\langle N | H | N \rangle$ [14, 16, 15]. Since spin is a good quantum number here, in what follows we will assign

a spin index to the spin related quantities.

As one of the standard procedures to set up equations for Green functions, one calculates the time derivative $i\frac{\partial}{\partial t}G_\sigma(x, x')$. This is done in a brute force way using the definition Eq.(2.1) and the Heisenberg equation of motion $i\frac{\partial}{\partial t}\hat{\Psi} = [\hat{\Psi}, H]$ repeatedly. This procedure leads to the equation of motion of Green functions:[15, 17]

$$i\frac{\partial}{\partial t}G_\sigma(x, x') = \delta(x - x') + \int d^3x'' [H_0(x, x'') + \Sigma_\sigma(x, x'')] G_\sigma(x'', x'). \quad (2.2)$$

Here the self-energy function $\Sigma_\sigma(x, x'')$ is introduced. We don't wish to delve in the formal derivation of Eq.(2.2) and how to calculate $\Sigma_\sigma(x, x'')$; here we only want to introduce the mathematical structure of Eq.(2.2) (and subsequently Eq.2.5) and to compare it with the Hartree-Fock and Kohn-Sham equations which we shall talk later. As our system is in equilibrium, the time variable should bear translational symmetry. We can thus Fourier transform Eq.(2.2) with respect to $(t - t')$, which yields

$$\int d^3r'' [\omega\delta(\mathbf{r} - \mathbf{r}'') - H_0(\mathbf{r}, \mathbf{r}'') - \Sigma_\sigma(\mathbf{r}, \mathbf{r}'', \omega)] G_\sigma(\mathbf{r}'', \mathbf{r}', \omega) = \delta(\mathbf{r} - \mathbf{r}'). \quad (2.3)$$

Trivially, if $\Sigma_\sigma(\mathbf{r}, \mathbf{r}', \omega)$ is known the Green function can be solved exactly from the following spectral representation

$$G_\sigma(\mathbf{r}, \mathbf{r}', \omega) = \sum_j \frac{\psi_{j,\sigma}(\mathbf{r}, \omega)\psi_{j,\sigma}^\dagger(\mathbf{r}', \omega)}{\omega - E_{j,\sigma}(\omega)}, \quad (2.4)$$

where $\psi_{j,\sigma}(\mathbf{r}, \omega)$'s are solutions to the quasiparticle equation

$$\underbrace{\left[-\frac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{\text{ext}}(\mathbf{r})\right]}_{\hat{H}_0} \psi_{j,\sigma}(\mathbf{r}, \omega) + \int d^3r' \Sigma_\sigma(\mathbf{r}, \mathbf{r}', \omega)\psi_{j,\sigma}(\mathbf{r}', \omega) = E_{j,\sigma}(\omega)\psi_{j,\sigma}(\mathbf{r}, \omega). \quad (2.5)$$

At each given ω this equation resembles the usual one-particle Schrödinger equation, except that in Eq.(2.5) the effective Hamiltonian $(H_0 + \Sigma_\sigma(\omega))$ is frequency dependent and is in general not Hermitian. Because of the latter fact, the eigenvalue $E_{j,\sigma}(\omega)$

can be complex and hence $\psi_{j,\sigma}$ does not mean any real electron state.

The significance of Eq.(2.5) is that it shows how a many-body interacting problem can be formally transformed into a set of one-particle Schrodinger-like equations, providing the self-energy $\Sigma_\sigma(\mathbf{r}, \mathbf{r}', \omega)$ is known beforehand. However, the reality is that there is no way to get the exact self-energy because of its enormous complexity. Therefore, many of the first-principles theories amount to using some simpler functions (functionals) to approximate the self-energy[18].

Another useful theorem (for finite systems only) is the Lehmann representation of the Green function:[17]

$$G_\sigma(\mathbf{r}, \mathbf{r}', \omega) = \sum_i \frac{h_{i,\sigma}(\mathbf{r})h_{i,\sigma}^*(\mathbf{r}')}{\omega - \mu + e(N-1, i) - i0^+} + \sum_i \frac{p_{i,\sigma}^*(\mathbf{r})p_{i,\sigma}(\mathbf{r}')}{\omega - \mu - e(N+1, i) + i0^+} \quad (2.6)$$

$$\begin{aligned} G_\sigma(\mathbf{r}, t; \mathbf{r}', t') = & i\theta(t' - t) \sum_i h_{i,\sigma}(\mathbf{r})h_{i,\sigma}^*(\mathbf{r}')e^{i[\mu - e(N-1, i)](t' - t)} \\ & - i\theta(t - t') \sum_i p_{i,\sigma}^*(\mathbf{r})p_{i,\sigma}(\mathbf{r}')e^{-i[\mu + e(N+1, i)](t - t')} \end{aligned} \quad (2.7)$$

in frequency and time domain respectively, where

$$h_{i,\sigma}(\mathbf{r}) = \langle N-1, i | \hat{\Psi}_\sigma(\mathbf{r}, t=0) | N \rangle, \quad p_{i,\sigma}(\mathbf{r}) = \langle N+1, i | \hat{\Psi}_\sigma^\dagger(\mathbf{r}, t=0) | N \rangle \quad (2.8)$$

$|N \pm 1, i\rangle$ is the i th eigenstate of the $N \pm 1$ electron system with an excitation energy $e(N \pm 1, i) = E(N \pm 1, i) - E(N \pm 1)$, which is positive and $E(N \pm 1)$ is the ground state energy of the $N \pm 1$ electron system. μ is the chemical potential $\mu = E(N+1) - E(N) = E(N) - E(N-1) + O(1/N)$. From Eq.(2.6) we see that the poles of the Green function equal the exact (one-particle) excitation energies. Besides, from Eq.(2.7) we find a suggestive expression for the charge density, in terms of the

quasi-particle orbitals:

$$\rho_\sigma(\mathbf{r}) \equiv -iG_\sigma(\mathbf{r}, t; \mathbf{r}, t^+) = \sum_i |h_{i,\sigma}(\mathbf{r})|^2 \quad (2.9)$$

as if $h_{i,\sigma}(\mathbf{r})$'s were occupied one-particle states.

All the theories and equations presented so far are exact. Eqs.(2.5), (2.6), and (2.9) are the key points of this subsection.

2.1.2 Hartree-Fock Approximation

The Hartree-Fock method can be viewed as a precursor before the DFT emerged. This method can be derived in several different yet equivalent ways, for example, starting with a mean-field effective Hamiltonian[19] or assuming the ground-state wavefunction to be a single Slater-determinant[20]. Here we choose to derive the Hartree-Fock equation under the Green function formalism.

The Hartree-Fock approximation consists in approximating the self-energy $\Sigma_\sigma(\mathbf{r}, \mathbf{r}', \omega)$ by its lowest order term[15]

$$\begin{aligned} \Sigma_\sigma^{\text{HF}}(\mathbf{r}, \mathbf{r}', \omega) &= i \int \frac{d\omega'}{2\pi} G_\sigma^{\text{HF}}(\mathbf{r}, \mathbf{r}', \omega') v(\mathbf{r} - \mathbf{r}') e^{i\omega'0^+} \\ &= - \sum_i \frac{h_{i,\sigma}(\mathbf{r}) h_{i,\sigma}^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \end{aligned} \quad (2.10)$$

where the second equality comes from the Lehmann representation Eq.(2.7). Since $\Sigma_\sigma^{\text{HF}}$ is Hermitian and frequency *independent* (static mean-field), Eq.(2.5) reduces to a Schrödinger equation and hence Eq.(2.4) should coincide with Eq.(2.6). We thus plug Eq.(2.10) in Eq.(2.5):

$$\left[-\frac{1}{2} \nabla^2 + V_{\text{H}}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r}) \right] h_{j,\sigma}(\mathbf{r}) + \int d^3r' \Sigma_\sigma^{\text{HF}}(\mathbf{r}, \mathbf{r}') h_{j,\sigma}(\mathbf{r}') = E_{j,\sigma} h_{j,\sigma}(\mathbf{r})$$

and then arrive at the quasi-particle Schrödinger equation:

$$\left[-\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) \right] h_{j,\sigma}(\mathbf{r}) + \int d^3r' \sum_{\sigma'} \sum_i^N \frac{|h_{i,\sigma'}(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} h_{j,\sigma}(\mathbf{r}) - \int d^3r' \sum_i^N \frac{h_{i,\sigma}^*(\mathbf{r}') h_{i,\sigma}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} h_{j,\sigma}(\mathbf{r}) = E_{j,\sigma} h_{j,\sigma}(\mathbf{r}). \quad (2.11)$$

This is the well-known Hartree-Fock equation. Upon knowing the orbital number N , this equation can be solved iteratively starting from an initial guess for each $h_{i,\sigma}(\mathbf{r})$; apparently N is just the particle number because $h_{i,\sigma}(\mathbf{r})$'s are orthonormalized and satisfy Eq.(2.9).

The last term on L.H.S of Eq.(2.11) is of great importance: it is termed as the exchange interaction in that it reflects the Pauli's exclusion principle[20], which reduces the possibility of spin-aligned electrons getting too close. Therefore the exchange effect results in a deduction from the bare coulomb interaction. This effect also implies that electrons might be able to lower the total energy by aligning their spins, which is important for understanding the magnetic phenomena.

The spatially nonlocal feature of the exchange interaction manifests its quantum nature. However, as $\Sigma_{\sigma}^{\text{HF}}$ is frequency independent, the Hartree-Fock approximation completely discards the interaction effects that are nonlocal in time, usually referred as correlation effects in addition to the exchange effect. One important correlation effect is the so-called “dynamical screening” effect: even in an equilibrium system there are electron density fluctuations taking place from time to time, which modulates the electron-electron interaction as if the bare coulomb potential were somewhat “screened” to become short-ranged[15]. This effect turns out to be very important in large-scale systems where the long-range feature of the coulomb interaction plays out. This is partly why the Hartree-Fock method often fails for periodic solid-state systems.

For practical reasons what we need is a first-principles method that can take corre-

lation effects into account while keeping the computational cost affordable. However, going beyond Hartree-Fock in the Green function formalism, though systematic, yet is known to be very difficult and computationally demanding[17].

2.1.3 Kohn-Sham DFT Formalism

Now let us turn into another route to solving the many-body problem. Instead of manipulating the Green functions, DFT starts with investigating the formally exact ground-state many-body wavefunction, which is the solution to the many-body Schrödinger equation. One then obtains the cornerstone of the entire DFT theory – Hohenberg-Kohn Theorem[21] (the derivation can be found in any DFT textbook such as Ref.[22]). The Hohenberg-Kohn Theorem is essentially a variational principle, which states that

For a system of interacting electrons in an external potential $V_{\text{ext}}(\mathbf{r})$, the total energy E_T in the ground state is a universal functional of the ground state particle density $\rho(\mathbf{r})$. Besides, this energy functional $E_T[\rho]$ approaches its global minimum value varying the density $\rho(\mathbf{r})$ towards the true ground-state density.

Though the Hohenberg-Kohn theorem is theoretically exact, one might find it useless since it says nothing about how to construct the density functional $E_T[\rho]$, which however appears to be the heart of this theory.

From our basic knowledge we know that the total energy must consist of four parts. Namely, the kinetic energy, exchange-correlation (XC) energy, Hartree energy, and the energy induced by the external electrical field. Among these four, the density functionals for the first two are yet unknown, but the last two are just electrostatic energies which can be formulated exactly:

$$E_H[\rho] = \frac{1}{2} \int d^3r \int d^3r' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (2.12)$$

$$E_{\text{ext}}[\rho] = \int d^3r \rho(\mathbf{r}) \cdot V_{\text{ext}}(\mathbf{r}) \quad (2.13)$$

Thus the difficulty of expressing $E_{\text{T}}[\rho]$ consists in finding the density functionals for the kinetic and XC energies. To proceed one resorts to the Kohn-Sham (KS) Ansatz[6].

The ansatz consists in assuming that the total particle density can still be expressed in the general form (c.f. Eq.2.9)

$$\rho(\mathbf{r}) = \sum_i^{\text{occ}} |\psi_i(\mathbf{r})|^2 \quad (2.14)$$

with some quasiparticle orbitals $\psi_i(\mathbf{r})$'s to be determined. As part of the ansatz the $\psi_i(\mathbf{r})$'s are required to be orthonormalized: $\langle \psi_i | \psi_j \rangle = \delta_{ij}$. Then the Kohn-Sham kinetic energy is introduced:

$$T_{\text{KS}}[\rho] = \sum_i \langle \psi_i | -\frac{1}{2} \nabla^2 | \psi_i \rangle, \quad (2.15)$$

which is of course different from the true kinetic energy, but this difference can be absorbed into the XC functional $E_{\text{xc}}[\rho]$ since we do not know the exact expression of the latter anyway. This process might seem rather irresponsible, but as we shall see later in section 2.2 there exist well developed methods to approximate $E_{\text{xc}}[\rho]$ and these methods can work sufficiently well for lots of real problems.

The partition of $E_{\text{T}}[\rho]$ within the Kohn-Sham formalism is summarized as

$$E_{\text{T}} = T_{\text{KS}} + E_{\text{ext}} + E_{\text{H}} + E_{\text{xc}}.$$

The genius of this partition is that by explicitly separating out the long-range electrostatic terms and the kinetic energy term which bears an independent-particle form, the remaining E_{xc} can reasonably be approximated as a local or nearly local density

functional[22], which can be expressed in the form

$$E_{\text{xc}}[\rho] = \int d^3r \varepsilon_{\text{xc}}[\rho](\mathbf{r}) \cdot \rho(\mathbf{r}), \quad (2.16)$$

where the functional $\varepsilon_{\text{xc}}[\rho]$ is often called the XC energy density. For clarity we write down the full expression of $E_{\text{T}}[\rho]$:

$$\begin{aligned} E_{\text{T}} &= T_{\text{KS}} + E_{\text{ext}} + E_{\text{H}} + E_{\text{xc}} \\ &= \sum_i \langle \psi_i | -\frac{1}{2} \nabla^2 | \psi_i \rangle + \int d^3r \rho(\mathbf{r}) \cdot V_{\text{ext}}(\mathbf{r}) \\ &\quad + \frac{1}{2} \int d^3r \int d^3r' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \int d^3r \varepsilon_{\text{xc}}[\rho](\mathbf{r}) \cdot \rho(\mathbf{r}) \end{aligned} \quad (2.17)$$

So far we still have not obtained any practical scheme to actually calculate something. To proceed we shall utilize the variational principle proposed in the Hohenberg-Kohn theorem: $\delta E_{\text{T}}[\rho]/\delta \rho|_{\rho=\rho_0} = 0$. With the help of Kohn-Sham ansatz, variations of $E_{\text{T}}[\rho]$ with respect to $\rho(\mathbf{r})$ can now be carried out with respect to $\psi_i(\mathbf{r})$'s instead, under the constraint $\langle \psi_i | \psi_j \rangle = \delta_{ij}$. Thus we have

$$\delta \left\{ E_{\text{T}}[\rho] - \sum_i \varepsilon_i \cdot [\langle \psi_i | \psi_i \rangle - 1] \right\} / \delta \langle \psi_i | = 0 \quad (2.18)$$

Here the Lagrange multiplier ε_i is introduced. (Strictly speaking, the multipliers should form a matrix ε_{ij} rather than just a vector ε_i , but one can always diagonalize ε_{ij} at the end, which is equivalent to performing a unitary transformation about $|\psi_i\rangle$'s [20].) Eq.(2.18) immediately leads to the Kohn-Sham equation

$$\left(-\frac{1}{2} \nabla^2 + V_{\text{ext}} + \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\text{xc}}[\rho] \right) |\psi_i\rangle = \varepsilon_i |\psi_i\rangle \quad (2.19)$$

where the local XC potential $V_{\text{xc}}[\rho]$ is defined as $V_{\text{xc}}[\rho](\mathbf{r}) = \delta E_{\text{xc}}[\rho(\mathbf{r})]/\delta \rho(\mathbf{r})$. This equation resembles Eq.(2.5) again and the interpretation is similar. The Kohn-Sham equation can be easily generalized for spin-polarized systems by using a spin-density

dependent XC functional $E_{\text{xc}}[\rho^\uparrow, \rho^\downarrow]$ (see Eq.2.23 below); here we continue suppressing the spin indices for demonstration convenience. It is worth stressing that the density functional $V_{\text{xc}}[\rho]$ is still unknown at this point.

Since we assume in DFT that the density calculated from

$$\rho(\mathbf{r}) \equiv \sum_{i=1}^N |\psi_i(\mathbf{r})|^2 = -iG_{\text{KS}}(\mathbf{r}, t; \mathbf{r}, t^+)$$

equals the true groundstate density $\rho_0(\mathbf{r}) \equiv -iG(\mathbf{r}, t; \mathbf{r}, t^+)$, by using Eqs.(2.3), (2.4), and (2.19), the relation between $V_{\text{xc}}[\rho]$ and the self-energy is found to be[18]:

$$\begin{aligned} & \int d^3r' V_{\text{xc}}(\mathbf{r}') \int \frac{d\omega}{2\pi} G_{\text{KS}}(\mathbf{r}, \mathbf{r}', \omega) G(\mathbf{r}', \mathbf{r}, \omega) \\ &= \int d^3r_1 \int d^3r_2 \int \frac{d\omega}{2\pi} G_{\text{KS}}(\mathbf{r}, \mathbf{r}_1, \omega) \Sigma(\mathbf{r}_1, \mathbf{r}_2, \omega) G(\mathbf{r}_2, \mathbf{r}, \omega), \end{aligned} \quad (2.20)$$

where G denotes the exact Green function and G_{KS} is defined using Eq.(2.4) with exact Kohn-Sham orbitals $\psi_i(\mathbf{r})$'s. From Eq.(2.20) it becomes clear that V_{xc} is definitely not just a simple static local potential because it should incorporate complicated many-body effects contained in the exact self-energy, which is nonlocal in both real-space and time domain. Approximate expressions for V_{xc} will be discussed later in section 2.2.

Now we see that the Kohn-Sham scheme maps the interacting many-body system down to an effective noninteracting system, where we only need to solve Eq.(2.19) which is just a single-particle Schrödinger equation with the effective Hamiltonian

$$H_{\text{eff}} = -\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r})$$

$$V_{\text{eff}} = V_{\text{ext}} + V_{\text{H}} + V_{\text{xc}}.$$

Once Eq.(2.19) is solved the density can be readily calculated from Eq.(2.14). However, in this approach one tricky thing is that before knowing the ground-state density

$\rho_0(\mathbf{r})$ the correct H_{eff} remains unknown because V_{xc} and V_{H} explicitly depend on $\rho_0(\mathbf{r})$, while $\rho_0(\mathbf{r})$ is in turn related to the eigen-wavefunctions of H_{eff} via Eq.(2.14). Thus an iteration scheme is required in order to obtain the correct $\rho_0(\mathbf{r})$ and H_{eff} simultaneously, which is usually referred as the self-consistency iteration. Remarkably, in each iteration only the density $\rho(\mathbf{r})$ needs to be updated. This makes the numerical implementation a lot easier, compared to other first principles methods such as the Hartree-Fock method, where all quasiparticle orbitals need to be updated at each iteration.

Before entering into the next section, let us discuss what is the proper interpretation for the solutions to the Kohn-Sham equation. The first thing is the Kohn-Sham orbitals $\psi_i(\mathbf{r})$'s. Although they are commonly used to depict the probability distribution of single electrons in the system, strictly speaking there is no such direct relation: $\psi_i(\mathbf{r})$'s are just some variables introduced in the DFT variational procedure from mathematical point of view, or just some quasi-particle orbitals in the spectral representation of Green functions. The second thing is the eigenvalues ε_i 's. Again, in practice they are commonly accepted to represent the electronic structure of the system. However, generally speaking, ε_i cannot be interpreted as the energy to add (affiliation) or subtract (ionization) an electron to/from the interacting many-body system; there is only one exception though[23]: if the system is finite, the eigen-energy of the highest occupied state should equal the ionization energy (up to a minus sign). Another way to investigate the meanings of ε_i 's is to compare them with the total energy. To this end we apply $\langle\psi_i|$ on both sides of Eq.(2.19) and then sum them up:

$$\begin{aligned} \sum_{i=1}^N \varepsilon_i &= \sum_i \langle\psi_i| -\frac{1}{2}\nabla^2 |\psi_i\rangle + \int d^3r V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) \\ &\quad + \int d^3r \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\rho(\mathbf{r}) + \int d^3r V_{\text{xc}}(\mathbf{r})\rho(\mathbf{r}) \end{aligned} \quad (2.21)$$

By comparing Eq.(2.21) to Eq.(2.17) we find

$$E_T = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \int d^3r \int d^3r' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \int d^3r \rho(\mathbf{r}) \cdot [\varepsilon_{xc}(\mathbf{r}) - V_{xc}(\mathbf{r})] \quad (2.22)$$

As can be seen there is no direct relation between E_T and ε_i either. In fact the correct relation between E_T and ε_i is given by the Slater-Janak theorem[24, 7] instead: $\partial E_T / \partial n_i = \varepsilon_i$, where n_i is the occupation number (fractional) of the i th state. As a byproduct, one often finds Eq.(2.22) a convenient formula for calculating the groundstate total energy.

2.2 Exchange-Correlation Functionals

Now the only barrier we are faced with toward solving the Kohn-Sham equation is the unknown functional $V_{xc}[\rho]$. Although the existence of such functional is definite, getting its exact form means the full solution to the groundstate many-electron problem, which is still impossible at the present stage. One thus has to resort to various approximations in order to make DFT a numerically practical scheme for real material simulations.

In this work we are confined to the level of local spin density approximation (LSDA) and the generalized gradient approximation (GGA) for the XC functional. In LSDA or GGA the exchange effect along with the correlation effect is calculated only approximately, which could in principle be done accurately as in the Hartree-Fock method, but for solid state systems LSDA and GGA actually work much better than Hartree-Fock. It seems that the inclusion of the correlation effect, though “coarsely” approximated, is much more important than just calculating the exchange effect.

2.2.1 XC Potentials for Collinear Spin

The term “collinear spin” refers to the situation where the axis of the spin quantization is the same at all points in space. In this case spin remains a good quantum number and the Kohn-Sham equations can thus be separated into two uncoupled sets:

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{ext}} + \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\text{xc}}^\sigma \right) |\psi_i^\sigma\rangle = \varepsilon_i^\sigma |\psi_i^\sigma\rangle, \quad (2.23)$$

where $\sigma = \uparrow, \downarrow$. As can be seen the only factor to distinguish the spin-up part from the down part is the spin dependence in the XC potential V_{xc}^σ , which is obtained from varying the XC energy functional with respect to the corresponding spin-resolved charge density: $V_{\text{xc}}^\sigma(\mathbf{r}) = \delta E_{\text{xc}} / \delta \rho^\sigma(\mathbf{r})$.

On the LSDA level the XC energy functional is simply formulated as an integral over all space with the XC energy density at each point assumed to be the same as in a homogeneous electron gas with that density:

$$E_{\text{xc}}^{\text{LSDA}}[\rho^\uparrow, \rho^\downarrow] = \int d^3r \rho(\mathbf{r}) \varepsilon_{\text{xc}}^{\text{hom}}[\rho^\uparrow(\mathbf{r}), \rho^\downarrow(\mathbf{r})].$$

The expression of the XC potential V_{xc}^σ is thus

$$V_{\text{xc}}^\sigma(\mathbf{r}) = \varepsilon_{\text{xc}}^{\text{hom}} + \rho \frac{\partial \varepsilon_{\text{xc}}^{\text{hom}}}{\partial \rho^\sigma}. \quad (2.24)$$

The homogeneous XC energy density $\varepsilon_{\text{xc}}^{\text{hom}}$ can be split into two parts, $\varepsilon_{\text{xc}}^{\text{hom}} = \varepsilon_{\text{x}}^{\text{hom}} + \varepsilon_{\text{c}}^{\text{hom}}$, where the exchange term $\varepsilon_{\text{x}}^{\text{hom}}[\rho^\sigma]$ simply scales as $(\rho^\sigma)^{-1/3}$. Because of this, it follows that the exchange part of V_{xc}^σ is particularly simple:

$$V_{\text{x}}^\sigma(\mathbf{r}) = \frac{4}{3} \varepsilon_{\text{x}}^{\text{hom}}[\rho^\sigma(\mathbf{r})],$$

while the correlation part depends on the specific form to be assumed. In fact, no analytical expression is available for the exact correlation energy of a homogeneous electron gas. However it can be calculated from stochastic methods[25] based on

exact expressions for the low and high density asymptotic behaviors of the uniform electron gas[26]. Other approaches consist in using analytic expressions that give the correct limiting behaviors in agreement with the data interpolation[27, 28, 29].

The success of LSDA has led to the development of various generalized-gradient approximations (GGAs) with marked improvements over LSDA for many cases. The first step to go beyond local approximations is to include the magnitude of the density gradient $|\nabla\rho^\sigma|$ as another argument of the XC functional. A spin-polarized GGA functional has the generalized form:[30]

$$\begin{aligned} E_{\text{xc}}^{\text{GGA}}[\rho^\uparrow, \rho^\downarrow] &= \int d^3r \rho(\mathbf{r}) \varepsilon_{\text{xc}}[\rho^\uparrow, \rho^\downarrow, |\nabla\rho^\uparrow|, |\nabla\rho^\downarrow|] \\ &= \int d^3r \rho(\mathbf{r}) \varepsilon_{\text{x}}^{\text{hom}}[\rho] F_{\text{xc}}[\rho^\uparrow, \rho^\downarrow, |\nabla\rho^\uparrow|, |\nabla\rho^\downarrow|], \end{aligned}$$

where F_{xc} is a dimensionless function and $\varepsilon_{\text{x}}^{\text{hom}}$ is the exchange energy per particle in the unpolarized homogeneous electron gas. In GGA one can derive the XC potential by finding the change $\delta E_{\text{xc}}^{\text{GGA}}$ to linear order in $\delta\rho^\sigma$:

$$\delta E_{\text{xc}}^{\text{GGA}} = \sum_{\sigma} \int d^3r \left[\varepsilon_{\text{xc}} + \rho(\mathbf{r}) \frac{\partial \varepsilon_{\text{xc}}}{\partial \rho^\sigma} + \rho(\mathbf{r}) \frac{\partial \varepsilon_{\text{xc}}}{\partial \nabla \rho^\sigma} \nabla \right] \cdot \delta \rho^\sigma(\mathbf{r})$$

The term in square brackets might thus be identified as the XC potential; however, it does not have the form of a local potential because of the last term, which is a differential operator. One approach to dealing with the last term is to perform a partial integration. This yields[22, 31]

$$V_{\text{xc}}^\sigma(\mathbf{r}) = \varepsilon_{\text{xc}} + \rho(\mathbf{r}) \frac{\partial \varepsilon_{\text{xc}}}{\partial \rho^\sigma} - \nabla \cdot \left(\rho(\mathbf{r}) \frac{\partial \varepsilon_{\text{xc}}}{\partial \nabla \rho^\sigma} \right).$$

This is the form most commonly used.

In this thesis we do not delve into the details of constructing XC functionals. In the RESCU code the Libxc library[32] with a great number of functionals built inside can be called whenever the XC potential is to be calculated.

Lastly, for collinear-spin the total energy should be calculated as

$$E_T = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \int d^3r (\rho^\uparrow + \rho^\downarrow) V_H + \sum_{\sigma=\uparrow,\downarrow} \int d^3r \rho^\sigma \cdot [\varepsilon_{xc}^\sigma - V_{xc}^\sigma] \quad (2.25)$$

2.2.2 XC Potentials for Non-Collinear Spin

In the collinear spin case, there are only two spin-resolved densities $[\rho^\uparrow, \rho^\downarrow]$ and accordingly two potentials $[V_{xc}^\uparrow, V_{xc}^\downarrow]$. This is, however, not the most general case since the spin axis can vary in space (non-collinear spin configuration). When spin is no longer a good quantum number, any electronic state $|\psi\rangle$ should be expressed as a two component spinor, a direct product of the real-space and the spin wavefunctions:

$$|\psi\rangle = |\varphi^\uparrow(\mathbf{r})\rangle \otimes |\uparrow\rangle + |\varphi^\downarrow(\mathbf{r})\rangle \otimes |\downarrow\rangle.$$

Accordingly, the effective one-particle potential $\hat{V}^{\alpha\beta}$ to act on such a spinor should be expressed as

$$\hat{V}^{\alpha\beta} |\psi\rangle = \begin{pmatrix} \hat{V}^{\uparrow\uparrow} & \hat{V}^{\uparrow\downarrow} \\ \hat{V}^{\downarrow\uparrow} & \hat{V}^{\downarrow\downarrow} \end{pmatrix} \begin{pmatrix} \varphi^\uparrow \\ \varphi^\downarrow \end{pmatrix}.$$

Note that here $\hat{V}^{\alpha\beta}$ has nonzero off-diagonal elements, unlike in the collinear spin case where $\hat{V}^{\alpha\beta}$ is diagonal in the spin space. There are two sources for these off-diagonal elements. One is the off-diagonal terms from the non-collinear XC potential and the other is the j -dependent pseudopotentials if the spin-orbit coupling effect is considered [see Eq.(2.45)]. In this subsection we focus on the former.

To get the non-collinear spin XC potential, one first defines a local spin density matrix

$$\rho^{\alpha\beta}(\mathbf{r}) = \sum_{\nu}^{occ} f_{\nu} \varphi_{\nu}^{\alpha}(\mathbf{r}) \varphi_{\nu}^{\beta}(\mathbf{r})^* \rightarrow \begin{pmatrix} \rho^{\uparrow\uparrow} & \rho^{\uparrow\downarrow} \\ \rho^{\downarrow\uparrow} & \rho^{\downarrow\downarrow} \end{pmatrix}, \quad (2.26)$$

where f_{ν} is the occupation number of the ν th eigen-state φ_{ν} . To get a geometric picture of the local spin configuration, one can cast the density matrix into the following

form[33]:

$$\rho^{\alpha\beta}(\mathbf{r}) = \frac{1}{2}N \cdot \sigma_0 + \frac{1}{2}\vec{m} \cdot \vec{\sigma}$$

with

$$\begin{aligned} N &= \rho^{\uparrow\uparrow} + \rho^{\downarrow\downarrow} \\ m_x &= 2\text{Re}[\rho^{\uparrow\downarrow}] \\ m_y &= -2\text{Im}[\rho^{\uparrow\downarrow}] \\ m_z &= \rho^{\uparrow\uparrow} - \rho^{\downarrow\downarrow} \end{aligned} \tag{2.27}$$

and σ_0 being the 2 by 2 unity matrix, $\vec{\sigma} = [\sigma_x, \sigma_y, \sigma_z]$ being the Pauli matrices. The vector $\vec{m}(\mathbf{r})$ defined above represents the local magnetization and the quantity N is just the local electron density. The direction of $\vec{m}(\mathbf{r})$ can be represented by the Euler angle (θ, ϕ) specified at each point. Using the Euler angle we can define a *local* unitary transformation matrix[34, 35]

$$U = \begin{pmatrix} e^{i\phi/2} \cos(\theta/2) & e^{-i\phi/2} \sin(\theta/2) \\ -e^{i\phi/2} \sin(\theta/2) & e^{-i\phi/2} \cos(\theta/2) \end{pmatrix},$$

which can be shown to diagonalize $\rho^{\alpha\beta}(\mathbf{r})$:

$$U \begin{pmatrix} \rho^{\uparrow\uparrow} & \rho^{\uparrow\downarrow} \\ \rho^{\downarrow\uparrow} & \rho^{\downarrow\downarrow} \end{pmatrix} U^\dagger = \begin{pmatrix} \rho'^{\uparrow} & 0 \\ 0 & \rho'^{\downarrow} \end{pmatrix},$$

where $\rho'^{\uparrow,\downarrow} = \frac{1}{2}(N \pm |m|)$ are *local* spin-up and spin-down densities. Note that under this definition the local spin-up component is always greater than the spin-down component, which should have no problem within LSDA but will cause some ambiguity in GGA; I will discuss this later. Now we turn to demonstrate how the XC potential should be generated after $\rho'^{\uparrow,\downarrow}(\mathbf{r})$ is obtained.

In LSDA it is simply assumed that

$$V_{xc}^{\alpha\beta} = \begin{pmatrix} V_{xc}^{\uparrow\uparrow} & V_{xc}^{\uparrow\downarrow} \\ V_{xc}^{\downarrow\uparrow} & V_{xc}^{\downarrow\downarrow} \end{pmatrix} = U^\dagger \begin{pmatrix} V_{xc}'^{\uparrow} & 0 \\ 0 & V_{xc}'^{\downarrow} \end{pmatrix} U, \tag{2.28}$$

where $V'_{xc}{}^\sigma$ is obtained using Eq.(2.24) with $\varepsilon_{xc}^{\text{hom}}[\rho'^\sigma(\mathbf{r})]$ defined in the collinear spin case. Note that in the LSDA case, as generating the XC potential is done locally, it doesn't matter which of the $\frac{1}{2}(N \pm |m|)$ is viewed to be of spin-up and the other to be of spin-down; either way results in a same $V_{xc}^{\alpha\beta}$ as long as the U-matrix is defined consistently.

In the GGA case we could also follow the above procedure to generate $V_{xc}^{\alpha\beta}$ by utilizing the GGA collinear spin functionals (this is exactly what we decide to do in this work). However, there is one subtle thing hidden behind. In GGA, spin density gradients are always calculated *separately* for spin up and down components. Therefore we need to specify on each real-space point which of the $\frac{1}{2}(N \pm |m|)$ is spin-up, which is spin-down, without ambiguity. But in the non-collinear case there is no global spin axis and hence the “up” or “down” direction is not well defined *per se*. Therefore we adopt the convention that the greater component $\frac{1}{2}(N + |m|)$ is always taken as spin-up at all points, which is apparently not always true (just think of an antiferromagnetic spin configuration).

Lastly, similar to Eq.(2.25), the formula for calculating the total energy in the non-collinear spin case is

$$E_T = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \int d^3r \left(\rho'^\uparrow + \rho'^\downarrow \right) V_H + \sum_{\sigma=\uparrow,\downarrow} \int d^3r \rho'^\sigma \cdot \left[\varepsilon'_{xc}{}^\sigma - V'_{xc}{}^\sigma \right].$$

2.3 Mixing Techniques - Achieving Self-Consistency

Now that we have clarified how to compute $V_{xc}[\rho]$ in different situations, we are ready to make up a scheme for solving the Kohn-Sham equation. The scheme is sketched as follows:

1. Starting from a prediction of the density $\rho_{\text{in}}^{(\alpha\beta)}(\mathbf{r})$ as an input, calculate the corresponding effective potential $V_{\text{eff}}[\rho_{\text{in}}^{(\alpha\beta)}]$.

2. Plug $V_{\text{eff}}[\rho_{\text{in}}^{(\alpha\beta)}]$ in Kohn-Sham equation and solve it on a chosen basis set.
3. Determine the Fermi energy and compute the new density $\rho_{\text{out}}^{(\alpha\beta)}(\mathbf{r})$ using Eq.(2.14) or (2.26).
4. If the difference between $\rho_{\text{in}}^{(\alpha\beta)}$ and $\rho_{\text{out}}^{(\alpha\beta)}$ is within some prescribed tolerance, then stop; if not, make a new prediction (update) for $\rho_{\text{in}}^{(\alpha\beta)}$ and go back to step1.

As the Kohn-Sham equation can only be solved self-consistently, a key problem is how to make predictions for the input density $\rho_{\text{in}}^{(\alpha\beta)}$. When carrying out the Kohn-Sham self-consistency calculation, one should bear in mind that there is no guarantee the new $\rho_{\text{out}}^{(\alpha\beta)}$ calculated at the end of the last loop would be closer to the true solution than the ones calculated before. Therefore some algorithms are wanted to improve the convergence performance. Developing such algorithms is not only a mathematical problem but should also involve our understandings of material physics, because they are required to be adaptive and transferable for many different materials.

The most well developed method for speeding up convergence in electronic structure calculations is the “mixer” algorithm, which generates a new input density $\rho_{\text{in}}^{(\alpha\beta)}$ for the next loop by mixing the newly calculated $\rho_{\text{out}}^{(\alpha\beta)}$ with the ones calculated in previous iterations. The better the mixer quality, the faster the convergence will be achieved. In the remaining of this section we demonstrate the basic ideas of the mixer machinery; we suppress the spin indices because different spin components are treated similarly in a mixer.

Suppose at the beginning of a certain iteration our input ρ_{in} is already quite close to the self-consistent density ρ_{sc} , then the deviation in the output density ρ_{out} to linear order in the deviation in ρ_{in} is given by

$$\rho_{\text{out}} - \rho_{\text{sc}} = \frac{\delta\rho_{\text{out}}}{\delta V_{\text{in}}} \frac{\delta V_{\text{in}}}{\delta\rho_{\text{in}}} (\rho_{\text{in}} - \rho_{\text{sc}}). \quad (2.29)$$

If we denote

$$J \equiv 1 - \frac{\delta \rho_{\text{out}}}{\delta V_{\text{in}}} \frac{\delta V_{\text{in}}}{\delta \rho_{\text{in}}}$$

then it is easy to show that Eq.(2.29) can be rearranged into

$$\rho_{\text{sc}} = \rho_{\text{in}} + J^{-1} (\rho_{\text{out}} - \rho_{\text{in}}).$$

Therefore, suppose we are at the end of the m th iteration loop, then the optimized input density for the $(m+1)$ th iteration would be

$$\rho_{\text{in}}^{m+1} = \rho_{\text{in}}^m + J^{-1} (\rho_{\text{out}}^m - \rho_{\text{in}}^m).$$

If ρ_{in}^m is already very close to ρ_{sc} and assuming we are able to evaluate J^{-1} accurately, then we would immediately have $\rho_{\text{in}}^{m+1} \approx \rho_{\text{sc}}$ (problem solved). But unfortunately the calculation of J^{-1} is rather complicated, because physically speaking J is actually the dielectric matrix[22, 36] of the system which should be calculated as

$$J(\mathbf{r}_1, \mathbf{r}_2) = \delta(\mathbf{r}_1, \mathbf{r}_2) - \int d\mathbf{r}'^3 \chi(\mathbf{r}_1, \mathbf{r}') K(\mathbf{r}', \mathbf{r}_2)$$

with

$$K(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r}) \delta \rho(\mathbf{r}')},$$

$$\chi(\mathbf{r}, \mathbf{r}') = 2 \sum_{i=1}^{\text{occ}} \sum_j^{\text{empty}} \frac{\psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) \psi_j^*(\mathbf{r}') \psi_i(\mathbf{r}')}{\varepsilon_i - \varepsilon_j},$$

where ψ 's are eigen-wavefunctions of the Kohn-Sham equation. As directly calculating J is nearly unrealistic, every mixing algorithm amounts to finding an appropriate approximation for J^{-1} .

A simple but efficient mixing scheme for bulk systems was proposed by Kerker[37], where the dielectric matrix of simple metallic systems is utilized to approximate J in the momentum space:

$$J^{-1} \approx \mathbf{G}_q^1 = A \frac{q^2}{q^2 + q_0^2}.$$

Considering the most general case Eq.(2.27), the quantity to be mixed can have as many as four components, namely N, m_x, m_y, m_z , then the parameters A and q_0 should be defined separately for each component.

The Kerker mixing scheme bears the feature that for large wave vectors it reduces to a simple linear mixing with the linear mixing parameter A , i.e.

$$\rho_{\text{in}}^{m+1}(q) = \rho_{\text{in}}^m(q) + A[\rho_{\text{out}}^m(q) - \rho_{\text{in}}^m(q)] \quad \text{for large } q,$$

while for small q the function behaves like Aq^2/q_0^2 and thus mixes only a small amount of $\rho_{\text{out}}^m(q)$ into $\rho_{\text{in}}^m(q)$. This last feature is essential because it helps to suppress the notorious “charge-sloshing” effect that could largely lower the convergence speed for metallic systems. The “charge-sloshing” problem occurs because the change in Hartree potential is much more sensitive to small q components than to the large ones. Because the Hartree potential is a major part in V_{eff} , if we don’t damp the change in the small q components, the new V_{eff} generated by ρ_{in}^{m+1} might change noticeably, which is against getting convergence. An even worse thing that could happen is that $\{\rho_{\text{out}}^m, \rho_{\text{out}}^{m+1}, \rho_{\text{out}}^{m+2} \dots\}$ keep oscillating around ρ_{sc} and just don’t converge to it.

Another widely used algorithm is the Pulay mixer[5, 36, 38], which is more advanced and robust than the Kerker’s scheme. In Pulay scheme the $\mathbf{G}_q$ matrix is updated in each iteration:

$$J^{-1} \approx \mathbf{G}^m = \mathbf{G}^1 - \sum_{k,n=1}^{m-1} \bar{A}_{kn}^{-1} (\mathbf{G}^1 |\Delta R^n\rangle + |\Delta \rho^n\rangle) \langle \Delta R^k|$$

with

$$\mathbf{G}_q^1 = A \frac{q^2}{q^2 + q_0^2}$$

$$\Delta \rho^n = \rho_{\text{in}}^{n+1} - \rho_{\text{in}}^n,$$

$$\bar{A}_{kn} = \langle \Delta R^n | \Delta R^k \rangle,$$

$$\Delta R^n = (\rho_{\text{out}}^{n+1} - \rho_{\text{in}}^{n+1}) - (\rho_{\text{out}}^n - \rho_{\text{in}}^n).$$

The inner product is defined as

$$\langle A | B \rangle = \sum_q f_q A_q^* B_q \quad (2.30)$$

with

$$f_q = \frac{q^2 + q_1^2}{q^2}.$$

The inclusion of the weighting factor f_q turns out to be very important for suppressing the “charge-sloshing” effect: in calculating Eq.(2.30) more weight is put on the small q components to force the mixing algorithm to first converge the components with long-range oscillations[5, 36].

2.4 Pseudopotential Method

In solving the Schrödinger equation for condensed aggregates of atoms, space can be divided into two regions with quite different properties. The region near the nuclei, the core region, is composed primarily of tightly bound core electrons which respond very little to the presence of neighboring atoms. The remaining volume, the interstitial region, contains valence electrons which are involved in chemical bonds. It is apparently very redundant to carry along the core electron states in all calculations. Based on this observation, pseudopotentials are introduced to construct an effective ion potential to describe the interaction between valence-electrons and core-electrons. Another common difficulty in electronic structure calculations is that due to the requirement of orthogonalization the valence-electron wavefunctions usually bear a high-frequency oscillation feature in the core region, which is rather unfavorable for numerical calculations. To get over this problem, the pseudopotential is also devised to smoothen the (pesudo)valence-electron wavefunctions as well as the effective atomic potential in the core region.

In this section we will go over the basic pseudopotential theory and demonstrate the procedures to generate l - and j -dependent pseudopotentials. The latter is relevant to our implementation of spin-orbit coupling (SOC) calculations.

2.4.1 Construction of l -Dependent Pseudopotentials

Pseudopotentials are constructed element by element. The procedure always starts with an “all-electron (AE)” (involving every single electron in the system) DFT calculation on a single atom at a given reference electronic state. Note that the atomic reference state does not always have to be the ground state; in some situations excited states or even ionized states are more favorable[39].

In the absence of SOC, the angular momentum number l remains to be a good quantum number. The Kohn-Sham equation thus reduces to a set of radial Schrödinger equations as follows:

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V_{\text{eff}}^{\text{AE}}[\alpha(r)] \right] r R_{nl}^{\text{AE}}(r) = \varepsilon_{nl} r R_{nl}^{\text{AE}}(r), \quad (2.31)$$

where n is the principle quantum number, R is the radial part of the Kohn-Sham orbital, V^{AE} is the DFT effective potential, and α is the radial density distribution

$$\alpha(r) = \int \rho(r) r^2 d\phi d\cos\theta = \sum_{nlm} f(nlm) \cdot |r R_{nl}^{\text{AE}}(r)|^2, \quad (2.32)$$

where f represents the occupation number which can be fractional. Here the total density ρ is assumed to be isotropic[40]. Note that when generating pseudopotentials it is important to use the same XC functional $V_{xc}[\rho]$ as the one to be used in later simulations.

As mentioned above, the wavefunctions $R_{nl}^{\text{AE}}(r)$ solved from Eq.(2.31) usually oscillate strongly in the core region, which is unfavorable for numerical calculations. One thus wishes to replace $R_{nl}^{\text{AE}}(r)$ with a much smoother pseudo-wavefunction $R_{nl}^{\text{PS}}(r)$

and meanwhile $V_{\text{eff}}^{\text{AE}}(r)$ by a flatter pseudopotential $V_{\text{eff}}^{\text{PS}}(r)$. It is easy to see that the pseudopotential $V_{\text{eff}}^{\text{PS}}(r)$ should be l -dependent as well and is related to $R_{nl}^{\text{PS}}(r)$ through

$$V_{\text{scr},l}^{\text{PS}}(r) = \varepsilon_l - \frac{l(l+1)}{2r^2} + \frac{1}{2rR_l^{\text{PS}}(r)} \frac{d^2}{dr^2} [rR_l^{\text{PS}}(r)]. \quad (2.33)$$

Here the n quantum number has been omitted since we are only interested in the pseudopotentials for valence-electrons. Also we have added a subscript “scr” (screened) in $V_{\text{scr},l}^{\text{PS}}$ to indicate that this pseudopotential does not really represent the effective potential from a bare ion but still contains the effect of valence-electrons interacting with themselves; as will be seen later, to obtain a transferable pseudopotential the $V_{\text{scr},l}^{\text{PS}}(r)$ is subject to an “unscreening” process.

The philosophy underlying the pseudopotential method is that one hopes the solution to the Kohn-Sham equation of the many-atom system with pseudopotential inserted agrees with the all-electron calculation. So far what we have done is just rearranging the radial Schrödinger equation, but how can this make things work? It turns out that this really works if the following requirements are met[41]:

1. All-electron and pseudo valence eigenvalues agree for the chosen atomic reference configuration.
2. All-electron and pseudo valence wavefunctions exactly agree beyond a chosen core radius r_c .
3. The logarithmic derivatives of the all-electron and pseudo wavefunctions agree at r_c .
4. The integrated charge inside r_c for each wavefunction agrees (norm-conservation condition).

The first two requirements are a must so that the same scattering properties are held by pseudo-electrons. The third requirement follows from the fact that the wavefunction and its first derivative must be continuous at any point for a smooth potential.

Point 4 requires that the integrated charge within the core region

$$Q_l = \int_0^{r_c} dr r^2 |R_l^{\text{PS}}(r)|^2 = \int_0^{r_c} dr r^2 |R_l^{\text{AE}}(r)|^2$$

is the same for $R_l^{\text{PS}}(r)$ as for the all-electron counterpart $R_l^{\text{AE}}(r)$. Because of this last requirement, pseudopotentials determined by these four criteria are often called norm-conserving pseudopotentials (NCP). The norm conservation ensures the correct total charge. It also implies that the first energy derivative of the logarithmic derivatives of the all-electron and pseudo wavefunctions agree at r_c , which guarantees the preservation of low energy scattering properties as well as the transferability of the pseudopotential[22].

The cutoff radius r_c is a crucial parameter in generating the pseudopotential. The smaller r_c the better transferability but resulting in a less flat pseudo wavefunction or potential, and vice versa. When setting r_c one should always consider the balance between the transferability of the pseudopotential and the computational cost of the following material simulations.

Although throughout this work we only use the NCP scheme, it is worth mentioning that in some modern pseudopotential methods[22, 42, 43, 44], the norm-conservation requirement (point 4) can be removed so as to make the pseudo wavefunctions even smoother. What's more, in the widely used PAW method (projector augmented wave)[43, 44, 45] multipole moments of the core *all-electron* density are taken into account when calculating the electrostatic potential, while in our NCP approach the core charge density is assumed to be spherically symmetric.

As can be seen from above, the central task of generating a pseudopotential is to choose a proper function to define the pseudo wavefunction within the cutoff radius r_c . There are various ways to accomplish this[46, 47]. In this work we adopt the NCP generation scheme proposed by Troullier and Martins in Ref.[46], which has been implemented in the NanoBase code[48]. In the Troullier-Martins approach the

pseudo wavefunction is defined as follows[46]:

$$R_l^{\text{PS}}(r) = \begin{cases} R_l^{\text{AE}}(r) & \text{if } r \geq r_c \\ r^l \exp[p(r)] & \text{if } r \leq r_c, \end{cases}$$

where $p(r)$ is a polynomial:

$$p(r) = c_0 + c_2 r^2 + c_4 r^4 + c_6 r^6 + c_8 r^8 + c_{10} r^{10} + c_{12} r^{12}.$$

In this series, the c_1 term is removed in order to avoid a singularity of the screened pseudopotential $V_{\text{scr},l}^{\text{PS}}(r)$ at $r = 0$. The seven coefficients in the polynomial are then determined from the four NCPP requirements plus three extra restrictions: the continuity of the third and fourth derivatives of the pseudo wavefunction and the zero curvature of the screened pseudopotential at origin i.e. $V_{\text{scr},l}''(0) = 0$. Once $R_l^{\text{PS}}(r)$ is determined, the screened pseudopotential $V_{\text{scr},l}^{\text{PS}}(r)$ is obtained from Eq.(2.33), which can be explicitly written as[46]

$$V_{\text{scr},l}^{\text{PS}}(r) = \begin{cases} V_l^{\text{AE}}(r) & \text{if } r \geq r_c \\ \varepsilon_{nl} + \frac{l+1}{r} p'(r) + \frac{p''(r) + [p'(r)]^2}{2} & \text{if } r \leq r_c. \end{cases}$$

To unscreen $V_{\text{scr},l}^{\text{PS}}(r)$ and get the bare ion pseudopotential, i.e. to exclude the self-interaction between valence-electrons themselves, we need to subtract their XC and Hartree contributions from $V_{\text{scr},l}^{\text{PS}}(r)$:

$$V_{\text{ion},l}^{\text{PS}}(r) = V_{\text{scr},l}^{\text{PS}}(r) - V_H([\rho^{\text{PS}}], r) - V_{\text{xc}}([\rho^{\text{PS}}], r) \quad (2.34)$$

This gives an l -dependent ionic pseudopotential which represents the effective potential that the l th component of a propagating electron will “see” from the local ion.

Readers might have found questionable the subtraction done in Eq.(2.34) in terms of the XC part. Indeed, unlike the Hartree potential, in the XC potential it is impos-

sible to separate apart the valence electron contribution, because the XC potential is not a linear function of density. If we simply use the $V_{\text{ion},l}^{\text{PS}}$ generated from Eq.(2.34) for following material simulations, problems would arise when updating the XC potential in each self-consistency iteration. This problem can be remedied by adding a “nonlinear core correction” to the pseudo electron density, together as the input for the XC functional[22]. This correction is particularly important for transition metals where the inner 3p states strongly interact with the outer 3d states. Here we won’t discuss this effect any further; readers with interest in this topic are referred to Refs.[22, 49, 50].

Since the ionic pseudopotential $V_{\text{ion},l}^{\text{PS}}$ decays as Z/r when $r \rightarrow \infty$, Z being the positive charge carried by the ion, we usually choose to partition $V_{\text{ion},l}$ into its long- and short-range components separately:

$$\begin{aligned}\hat{V}_{\text{ion}}^{\text{PS}} &= \sum_{lm} |lm\rangle V_{\text{ion},l}^{\text{PS}}(r) \langle lm| \\ &= V_{\text{ion,local}}^{\text{PS}}(r) + \sum_{lm} |lm\rangle V_{\text{NL},l}(r) \langle lm|,\end{aligned}\quad (2.35)$$

where $V_{\text{ion,local}}^{\text{PS}}(r)$ is the local potential which decays as Z/r at long-range and,

$$V_{\text{NL},l}(r) = V_{\text{ion},l}^{\text{PS}}(r) - V_{\text{ion,local}}^{\text{PS}}(r)$$

is the radial part of the nonlocal (NL) or more precisely the semilocal potential for the l component. Note that according to its definition $V_{\text{NL},l}(r)$ should vanish beyond r_c . The local potential $V_{\text{ion,local}}^{\text{PS}}(r)$ can, in principle, be arbitrarily chosen, and it is widely adopted to make this local potential equal to the potential created by a positive charge distribution of the form[51]

$$\rho(r) \propto \exp \left[-(\sinh(abr)/\sinh(b))^2 \right],$$

where a and b are chosen to provide simultaneously optimal real-space localization

and reciprocal-space convergence[52]. Besides, because the long range feature of $V_{\text{ion,local}}^{\text{PS}}(r)$ will make it difficult to calculate the Madelung energy (ion-ion interaction), it is often desired to screen $V_{\text{ion,local}}^{\text{PS}}(r)$ by an electron cloud with the same amount of charge as the ion has, and as a compensation the same amount of charge will be deducted from valence-electrons when calculating the Hartree potential. The resulting “neutral-atom” (NA) potential is thus short ranged. To be concrete, now let $\delta\rho$ be the difference between the screening cloud density and the valence-electron density ρ , and let δV_{H} be the electrostatic potential generated by $\delta\rho$. Then the total Hamiltonian can be rewritten as

$$H = T + V_{\text{NL}} + V_{\text{NA}} + \delta V_{\text{H}} + V_{\text{xc}}. \quad (2.36)$$

Now we turn to look at the second part of the pseudopotential operator [Eq.(2.35)], $\sum_{lm} |lm\rangle V_{\text{NL},l}(r) \langle lm|$, which is often termed to be “semilocal” in that its radial part is local while its angular part is not. Consider $|lm\rangle V_{\text{NL},l}(r) \langle lm|$ acting on an arbitrary wavefunction $\varphi(r)$:

$$\langle \vec{r}(r, \theta, \phi) | lm \rangle V_{\text{NL},l}(r) \langle lm | \varphi \rangle = Y_l^m(\theta, \phi) \cdot V_{\text{NL},l}(r) \cdot \left\{ \int d\Omega' [Y_l^m(\theta', \phi')]^* \varphi(r, \theta', \phi') \right\},$$

where one needs to carry out the angular integral in the curly bracket at each r . This is apparently cumbersome for numerical calculations. To simplify calculations Kleinman and Bylander (KB)[53] came up with a fully nonlocal (or separable) form of the pseudopotential projector, which reads:

$$\hat{V}_{\text{NL},lm}^{\text{KB}} = \frac{|V_{\text{NL},l}(r) R_l^{\text{PS}}(r) Y_l^m(\theta, \phi)\rangle \langle V_{\text{NL},l}(r) R_l^{\text{PS}}(r) Y_l^m(\theta, \phi)|}{\langle r R_l^{\text{PS}}(r) | V_{\text{NL},l}(r) | r R_l^{\text{PS}}(r) \rangle}, \quad (2.37)$$

where the pseudo wavefunctions are also utilized. The radial part $V_{\text{NL},l}(r) R_l^{\text{PS}}(r)$ is often called the KB-orbital. One can justify this approximation simply by considering

$\hat{V}_{NL,lm}^{\text{KB}}$ acting on its corresponding pseudo wavefunction:

$$\begin{aligned}\hat{V}_{NL,lm}^{\text{KB}} [R_l^{\text{PS}}(r)Y_l^m(\theta, \phi)] &= \frac{|V_{NL,l}(r)R_l^{\text{PS}}(r)Y_l^m(\theta, \phi)\rangle}{\langle rR_l^{\text{PS}}(r)|V_{NL,l}(r)|rR_l^{\text{PS}}(r)\rangle} \int dr r^2 V_{NL,l}(r) |R_l^{\text{PS}}(r)|^2 \\ &= V_{NL,l}(r)R_l^{\text{PS}}(r)Y_l^m(\theta, \phi).\end{aligned}$$

As can be readily checked, this result is identical to that obtained from $|lm\rangle V_{NL,l}(r) \langle lm|$ acting on the same wavefunction. Substantial savings in both computation time and storage can be achieved using this separable expression[53]. However, one notorious drawback of the KB approximation is the possible existence of “ghost states”. This is because under this approximation the bound states of the radial Schrödinger equation are not necessarily ordered in energy by their number of nodes; spurious bound states with more nodes can appear below or between the zero- and one-node states. Therefore, careful examinations are necessary in generating the separable pseudopotentials (see discussions in Refs.[54, 55]).

2.4.2 Construction of j -Dependent Relativistic Pseudopotentials

Relativistic effects contain two parts, the scalar-relativistic effects and the spin-orbit coupling (SOC) effect. The former consists of the kinematic relativistic effects (mass-velocity and Darwin term) and does not mix different spin components. Consequently it does not cause any relativistic splittings in the resulting band structures. For cases where relativistic band splittings are supposed to exist and play a major role, it is mandatory to include SOC explicitly. For periodic systems the spin-orbit interaction is appropriately written in the form[56]

$$H_{SO} = \frac{\hbar}{4m_0^2c^2} (\nabla V \times \vec{p}) \cdot \vec{\sigma},$$

where V is the electrostatic potential and $\vec{\sigma}$ is the Pauli matrix. SOC is strongest close to the ion cores where ∇V is largest. However, this is the region covered by the pseudopotentials. Thus it is impossible to add the SOC effect *a posteriori* to the

scalar-relativistic pseudopotential results; instead the pseudopotentials themselves must account for SOC.

The pseudopotential generating procedure again starts with an all-electron DFT calculation on a single atomic configuration. But to fully include the relativistic effects it is the Dirac equation that needs to be solved[39, 48, 13]:

$$\frac{dG_{nlj}}{dr} + \frac{\kappa}{r}G_{nlj} - \alpha \left[\frac{2}{\alpha^2} + \varepsilon_{nlj} - V(r) \right] F_{nlj} = 0 \quad (2.38)$$

$$\frac{dF_{nlj}}{dr} - \frac{\kappa}{r}F_{nlj} + \alpha [\varepsilon_{nlj} - V(r)] G_{nlj} = 0, \quad (2.39)$$

where $\alpha = 1/137.036$ is the fine structure constant and $\varepsilon = E - \alpha^{-2}$, E being the relativistic energy. G and F are the majority and minority components of the radial wavefunction. $\kappa = l$ for $j = l - 1/2$ and $\kappa = -l - 1$ for $j = l + 1/2$. The Dirac equations are solved for a chosen atomic reference state[39] and the output is a set of one-electron eigenvalues (for the valence states) and radial wavefunctions (G and F) as well as the self-consistent effective potential $V(r)$. To get a formula favorable for pseudopotential constructions, we follow Kleinman's argument[13]: for valence electrons (usually with small ε and feel small V) Eq.(2.39) can be used to replace F by

$$F_{nlj} = \frac{\alpha}{2} \left[\frac{dG_{nlj}}{dr} + \frac{\kappa}{r}G_{nlj} \right] \quad (2.40)$$

Substituting Eq.(2.40) into Eq.(2.38) we arrive at

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{\kappa(\kappa+1)}{2r^2} + V(r) \right] G_{nlj}(r) = \varepsilon_{nlj} G_{nlj}(r), \quad (2.41)$$

which is exactly of the same form as in Eq.(2.31). Thus the regular Troullier-Martins scheme can be applied to generate pseudopotentials for each pair of l and j independently. Note that in some of the modern DFT codes the Kleinman's approximation Eq.(2.40) is not adopted and consequently a much more complicated differential equation is to be solved[35, 57]; the Kleinman's approximation is correct up to terms of order α^2 .

In the relativistic case the ionic pseudopotential operator should be written in the following general form (corresponding to the first line of Eq.(2.35)):

$$\begin{aligned}
\hat{V}_{\text{ion}}^{\text{PS}} &= \sum_{ljM} |ljM\rangle \hat{V}_{lj}^{\text{PS}} \langle ljM| \\
&= \sum_{l,M} |l+1/2, M\rangle \hat{V}_{l,l+1/2}^{\text{PS}} \langle l+1/2, M| \\
&\quad + \sum_{l,M'} |l-1/2, M'\rangle \hat{V}_{l,l-1/2}^{\text{PS}} \langle l-1/2, M'|,
\end{aligned} \tag{2.42}$$

where for $j = l + 1/2, M = m + 1/2$

$$|l+1/2, M\rangle = \left(\frac{l+m+1}{2l+1}\right)^{1/2} Y_l^m |\uparrow\rangle + \left(\frac{l-m}{2l+1}\right)^{1/2} Y_l^{m+1} |\downarrow\rangle$$

and for $j = l - 1/2, M' = m - 1/2$

$$|l-1/2, M'\rangle = \left(\frac{l-m+1}{2l+1}\right)^{1/2} Y_l^{m-1} |\uparrow\rangle - \left(\frac{l+m}{2l+1}\right)^{1/2} Y_l^m |\downarrow\rangle$$

and $V_{l,l\pm 1/2}^{\text{PS}}(r)$ is the radial part of the pseudopotential generated from the Troullier-Martins scheme. Here I would like to point out that the Dirac equation solver has already been built in the NanoBase code[48] and in its output file both the screened and the unscreened version of $V_{l,l\pm 1/2}^{\text{PS}}(r)$ can be found; this is the starting point of our SOC implementation.

Next, we look to transform the semi-local pseudopotentials in Eq.(2.42) into the separable KB form. This can be done in a number of different ways[35, 4, 58, 59, 60]. For implementation convenience we choose to follow the approach proposed in Ref.[4].

First, Eq.(2.42) is rewritten as (we indicate pseudopotentials with an overline)

$$\begin{aligned}\hat{V}_{\text{ion}}^{\text{PS}} = & \sum_{lm} |lm\rangle \bar{V}_l^{\text{SC}} \langle lm| \\ & + \sum_{l,M} \left| l + \frac{1}{2}, M \right\rangle \frac{l}{2} \bar{V}_l^{\text{SO}} \left\langle l + \frac{1}{2}, M \right| \\ & - \sum_{l,M'} \left| l - \frac{1}{2}, M' \right\rangle \frac{l+1}{2} \bar{V}_l^{\text{SO}} \left\langle l - \frac{1}{2}, M' \right| \end{aligned} \quad (2.43)$$

by introducing

$$\begin{aligned}\bar{V}_l^{\text{SC}}(r) &= \frac{1}{2l+1} [lV_{l,l-1/2}^{\text{PS}} + (l+1)V_{l,l+1/2}^{\text{PS}}] \\ \bar{V}_l^{\text{SO}}(r) &= \frac{2}{2l+1} [V_{l,l+1/2}^{\text{PS}} - V_{l,l-1/2}^{\text{PS}}] \end{aligned} \quad (2.44)$$

The last two terms in Eq.(2.43), the SOC terms, mix up different spin components of the wavefunction, and it is these two terms that account for the relativistic energy-level splitting. In contrast, the first term, often called the scalar (SC) term, is l -dependent only and thus acts on both spin components in the same manner. In fact, in most modern DFT codes the scalar term is the pseudopotential to be used even when the relativistic effects are not important. The major advantage of the partition performed in Eq.(2.43) is that we can still utilize the original code for the scalar term and the original code will not be changed much when implementing SOC calculations.

Since the long range tails are canceled out in $\bar{V}_l^{\text{SO}}(r)$ (see Eq.2.44), we can apply the KB procedure to SOC terms directly. The resulting separable pseudopotential reads[4]:

$$\begin{aligned}\hat{V}_{PS}^{\text{SO}} = & \sum_l \frac{l}{2} E_{l+1/2} \sum_{M=-l-1/2}^{l+1/2} |\chi_{l+1/2,M}^{\text{SO}}\rangle \langle \chi_{l+1/2,M}^{\text{SO}}| \\ & - \sum_l \frac{l+1}{2} E_{l-1/2} \sum_{M=-l+1/2}^{l-1/2} |\chi_{l-1/2,M}^{\text{SO}}\rangle \langle \chi_{l-1/2,M}^{\text{SO}}| \end{aligned} \quad (2.45)$$

with, to be explicit,

$$\chi_{l+1/2,M}^{\text{SO}}(\vec{r}) = \frac{R_{l+1/2}^{\text{PS}}(r) \bar{V}_l^{\text{SO}}(r)}{\left\langle r R_{l+1/2}^{\text{PS}} \bar{V}_l^{\text{SO}} \mid r R_{l+1/2}^{\text{PS}} \bar{V}_l^{\text{SO}} \right\rangle^{1/2}} \times \left[\left(\frac{l+m+1}{2l+1} \right)^{1/2} Y_l^m \mid \uparrow \rangle + \left(\frac{l-m}{2l+1} \right)^{1/2} Y_l^{m+1} \mid \downarrow \rangle \right] \quad (2.46)$$

$$\chi_{l-1/2,M}^{\text{SO}}(\vec{r}) = \frac{R_{l-1/2}^{\text{PS}}(r) \bar{V}_l^{\text{SO}}(r)}{\left\langle r R_{l-1/2}^{\text{PS}} \bar{V}_l^{\text{SO}} \mid r R_{l-1/2}^{\text{PS}} \bar{V}_l^{\text{SO}} \right\rangle^{1/2}} \times \left[\left(\frac{l-m+1}{2l+1} \right)^{1/2} Y_l^{m-1} \mid \uparrow \rangle - \left(\frac{l+m}{2l+1} \right)^{1/2} Y_l^m \mid \downarrow \rangle \right] \quad (2.47)$$

$$E_{l\pm 1/2} = \frac{\left\langle r R_{l\pm 1/2}^{\text{PS}} \bar{V}_l^{\text{SO}} \mid r R_{l\pm 1/2}^{\text{PS}} \bar{V}_l^{\text{SO}} \right\rangle}{\left\langle r R_{l\pm 1/2}^{\text{PS}} \mid \bar{V}_l^{\text{SO}} \mid r R_{l\pm 1/2}^{\text{PS}} \right\rangle} = \frac{\int dr r^2 \left| R_{l\pm 1/2}^{\text{PS}}(r) \bar{V}_l^{\text{SO}}(r) \right|^2}{\int dr r^2 \bar{V}_l^{\text{SO}}(r) \left| R_{l\pm 1/2}^{\text{PS}}(r) \right|^2}, \quad (2.48)$$

where the radial part $R_{l\pm 1/2}^{\text{PS}}(r)$ of the pseudo wavefunction generated from Eq.(2.41) is used in constructing KB-orbitals.

Once $V_{l,l\pm 1/2}^{\text{PS}}(r)$ (the screened one) is read out from the NanoBase output, $R_{l\pm 1/2}^{\text{PS}}(r)$ can be reproduced by solving Eq.(2.41). Another noticeable thing is that in contrast to the angular part in the scalar relativistic term where the real spherical harmonics can be used thanks to their completeness in the (l, m) subspace, here for SOC terms there seems to be no clear way how to transform the complex spherical harmonics into their real counterparts. Usually the real spherical harmonics are preferred under efficiency and memory concerns, but in the real-space formalism, since it is the wavefunctions that occupy the majority of the memories and they are complex anyway when spin is not a good quantum number, we will just use the complex spherical harmonics for SOC terms.

Lastly, since the RESCU code also performs LCAO calculations, for completeness

SOC calculations have also been implemented with the LCAO basis. For this part we follow the scheme proposed in Ref.[61] where SOC is assumed to be a purely on-site effect (see Appendix A for details).

2.5 *Summary*

We began with presenting the basic formula of the many-electron problem in the Green function formalism. We then derived the Hartree-Fock equation and made concrete the concepts of the basic quantum many-body effects in the problem. We then introduced the DFT and Kohn-Sham scheme which casts the original problem down to a set of single-particle Schrodinger equations. The basics of the LSDA and G-GA XC-functionals were introduced, which are required by all DFT implementations. We also demonstrated how to generalize these functionals for the non-collinear spin configuration, which requires the manipulation of local density matrices. To get rid of the computational efforts in the nuclei core region, DFT is often implemented in conjunction with the pseudopotential method. We walked through the basic procedures of generating a qualified pseudopotential and realized that the atomic Dirac equation needs to be solved in order for us to incorporate relativistic effects such as SOC. This leads to a nonlocal pseudopotential projector with each component depending on the j angular-momentum number.

3

Numerical Methods

In the last chapter we have demonstrated how DFT reduces the intractable many-electron problem down to a static mean-field problem and how to set up the Kohn-Sham effective Hamiltonian, including pseudopotentials. This chapter is devoted to demonstrating how to implement these theories into a practical numerical scheme.

3.1 Real-space Representation of Kohn-Sham Equation

We write down the most general form (with non-collinear spin and SOC) of the Kohn-Sham equation explicitly as follows

$$\left[\begin{array}{c|c} -\frac{\nabla^2}{2} + V_{\text{NA}} + \delta V_{\text{H}} & \hat{V}_{\text{ps}}^{\text{SO},\uparrow\downarrow} + V_{\text{xc}}^{\uparrow\downarrow} \\ \hline +\hat{V}_{\text{ps}}^{\text{SC}} + \hat{V}_{\text{ps}}^{\text{SO},\uparrow\uparrow} + V_{\text{xc}}^{\uparrow\uparrow} & \\ \hline \hat{V}_{\text{ps}}^{\text{SO},\downarrow\uparrow} + V_{\text{xc}}^{\downarrow\uparrow} & -\frac{\nabla^2}{2} + V_{\text{NA}} + \delta V_{\text{H}} \\ & +\hat{V}_{\text{ps}}^{\text{SC}} + \hat{V}_{\text{ps}}^{\text{SO},\downarrow\downarrow} + V_{\text{xc}}^{\downarrow\downarrow} \end{array} \right] \begin{bmatrix} \varphi_{\nu}^{\uparrow} \\ \varphi_{\nu}^{\downarrow} \end{bmatrix} = \varepsilon_{\nu} \begin{bmatrix} \varphi_{\nu}^{\uparrow} \\ \varphi_{\nu}^{\downarrow} \end{bmatrix} \quad (3.1)$$

For collinear spin the off-diagonal entries vanish and the Kohn-Sham equations for different spin components are decoupled.

As the Kohn-Sham equations are Schrodinger-like differential equations, to efficiently solve them one always needs a numerical basis set to represent the Hamiltonian and wavefunctions. This way the original differential equations are transformed into a matrix diagonalization problem, and then by finding eigenpairs of the Hamil-

tonian matrix we obtain the projection coefficients of the eigen-wavefunctions onto the chosen basis set as well as the corresponding eigenvalues. There are many choices for such basis sets, but no matter what basis one chooses, a real-space grid is indispensable anyway because the XC potential is calculated in real-space. Therefore, a question arises naturally: why don't we just express everything on the real-space grid?

In fact the major concern about the real-space method is that if we do so, considering the number of grid points we will need, the resulting Hamiltonian will have such a huge size that diagonalizing it will be almost impossible. However, as will be seen later, in most cases we don't really need to diagonalize the Hamiltonian completely because we are interested in only a small part of its spectrum. Besides, the Hamiltonian matrix does not need to be written down explicitly; in fact only the operation function $\hat{H}\psi$ is needed (view ψ as the input), which is particularly simple to be formulated in real-space. What's more, if we want to improve our simulation accuracy we just increase our real-space mesh density, which is quite convenient compared to the orbital based basis sets such as LCAO[22]. Therefore the real-space method is definitely a practical DFT implementation. In what follows we demonstrate how Eq.(3.1) is realized in the real-space DFT formalism.

3.1.1 Finite-Difference Method

The first term one encounters in expressing the Hamiltonian operator is the Laplacian $-\frac{1}{2}\nabla^2$. The gradient of the wavefunction $\nabla\psi$ is also wanted if the periodic boundary condition is imposed on the simulation box[22]. In addition, the gradient of the density $\nabla\rho$ is needed when we compute GGA functionals. Therefore an efficient technique is wanted in order for us to accurately calculate the derivatives of quantities expressed in real-space. This can be achieved by the high order finite-difference method, which was introduced to real-space DFT by Chelikowsky et al.[8, 62, 63] for the first time. For clarity in what follows we illustrate this method in 1D.

In the finite-difference method the k th derivative of a 1D function $\partial^k u(x)/\partial x^k$ at a certain point x_j can be approximated by a linear combination of the function values at neighboring points:

$$\frac{\partial^k u}{\partial x^k}(x_j) = \sum_i \alpha_i u(x_i). \quad (3.2)$$

By using Taylor theorem

$$u(x) = u(x_0) + u_x(x_0)(x - x_0) + \frac{1}{2!}u_{xx}(x_0)(x - x_0)^2 + \frac{1}{3!}u_{xxx}(x_0)(x - x_0)^3 + \cdots,$$

we get

$$\begin{aligned} \sum_i \alpha_i u(x_i) = & u(x_j) \sum_i \alpha_i + u_x(x_j) \sum_i \alpha_i (x_i - x_j) + \frac{1}{2!}u_{xx}(x_j) \sum_i \alpha_i (x_i - x_j)^2 \\ & + \frac{1}{3!}u_{xxx}(x_j) \sum_i \alpha_i (x_i - x_j)^3 + \cdots \end{aligned}$$

To approximate $\partial^k u(x)/\partial x^k$ using Eq.(3.2) we simply solve the following set of linear equations for the appropriate coefficients α_i 's:

$$\begin{aligned} \sum_i^n \alpha_i &= 0 \\ \sum_i^n \alpha_i (x_i - x_j) &= 0 \\ \frac{1}{2!} \sum_i^n \alpha_i (x_i - x_j)^2 &= 0 \\ &\vdots \\ \frac{1}{k!} \sum_i^n \alpha_i (x_i - x_j)^k &= 1 \\ &\vdots \\ \frac{1}{(n-1)!} \sum_i^n \alpha_i (x_i - x_j)^{n-1} &= 0 \end{aligned}$$

The α_i 's are found by solving a Vandermonde system of equations where the right hand side is a vector of zeros except for the k th entry which is equal to $k!$. The resulting approximation for $\partial^k u(x)/\partial x^k$ is accurate up to the order of $o(h^n)$ at least,

h being the grid spacing. As an example, to compute the α_i 's for $\partial^2 u(x)/\partial x^2$ we solve

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ (-2h) & (-h) & 0 & h & (2h) \\ (-2h)^2 & (-h)^2 & 0 & h^2 & (2h)^2 \\ (-2h)^3 & (-h)^3 & 0 & h^3 & (2h)^3 \\ (-2h)^4 & (-h)^4 & 0 & h^4 & (2h)^4 \end{pmatrix} \begin{pmatrix} \alpha_{k-2} \\ \alpha_{k-1} \\ \alpha_k \\ \alpha_{k+1} \\ \alpha_{k+2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

which yields

$$\begin{pmatrix} \alpha_{k-2} & \alpha_{k-1} & \alpha_k & \alpha_{k+1} & \alpha_{k+2} \end{pmatrix} = \frac{1}{12h^2} \begin{pmatrix} -1 & 16 & -30 & 16 & -1 \end{pmatrix}$$

We can thus calculate $\partial^2 u(x)/\partial x^2$ at x_j as

$$u_{xx}(x_j) = \frac{-u(x_{j-2}) + 16u(x_{j-1}) - 30u(x_j) + 16u(x_{j+1}) - u(x_{j+2}))}{12h^2} + o(h^6)$$

with an error of the order $o(h^6)$. To make things concrete we demonstrate as follows the representation of $-\frac{1}{2}\nabla_x^2 u$ at $n = 3$:

$$-\frac{1}{2}\nabla_x^2 u \approx -\frac{1}{2} \begin{pmatrix} c_0 & c_1 & 0 & \dots & 0 & c_1 e^{-ik_x L_x} \\ c_1 & c_0 & c_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \ddots & c_1 & c_0 & c_1 & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & & & c_1 & c_0 & c_1 \\ c_1 e^{ik_x L_x} & 0 & \dots & 0 & 0 & c_1 & c_0 \end{pmatrix} \begin{pmatrix} u(h_x) \\ u(2h_x) \\ \vdots \\ u(Nh_x - h_x) \\ u(Nh_x) \end{pmatrix}, \quad (3.3)$$

where $c_0 = -2/h_x^2$, $c_1 = 1/h_x^2$ and the periodic boundary condition is imposed, k_x being the Bloch wave number and L_x the length of the simulation box. For a finite system such as an isolated molecule or a cluster we can impose the open boundary condition (wavefunctions vanish outside), then in Eq.(3.3) the upper right and the

lower left elements should vanish.

The finite-difference method can be generalized to 3D straightforwardly if an orthogonal grid is used. In this case the operation $-\frac{1}{2}\nabla^2 u(x, y, z)$ takes the following form[64]

$$-\frac{1}{2}\nabla^2 u(x, y, z) = -\frac{1}{2h^2} \sum_{k,l,m=-M}^M [c_k u(x + kh, y, z) + c_l u(x, y + lh, z) + c_m u(x, y, z + mh)],$$

where the coefficients c_m depend only on the choice of the expansion order, $2M$ (usually $M = 6$ is sufficient for most problems[8, 65]). However, unlike in Eq.(3.3), it is difficult to express the 3D Laplacian operator in a matrix form. In fact, in RESCU the multi-dimensional Laplacians are obtained from the 1D Laplacian by using the Kronecker tensor product recursively[2, 65]: suppose that ∇_n^2 is a Laplacian in n dimension, then the Laplacian in $n + 1$ dimension is obtained from $\nabla_{n+1}^2 = \nabla_n^2 \otimes I_1 + I_n \otimes \nabla_1^2$, where I is the identity matrix; the differential operators never show up as matrices throughout the code.

For a general nonorthogonal real-space grid we should define another coordinate system: say $\hat{u}$, $\hat{v}$ and $\hat{w}$ are unit vectors along the nonorthogonal axes and $\hat{e}_i$'s are unit vectors along the canonical orthogonal axes, then the same vector $\mathbf{R}$ can be expressed as $\mathbf{R} = x\hat{e}_1 + y\hat{e}_2 + z\hat{e}_3$ or $= u\hat{u} + v\hat{v} + w\hat{w}$. Using matrix notation we have the following relation between the two sets of coordinates: $\mathbf{R}_{xyz} = \mathbf{A}\mathbf{R}_{uvw}$ and $\mathbf{R}_{uvw} = \mathbf{B}\mathbf{R}_{xyz}$, where $\mathbf{A} \equiv [\hat{u}, \hat{v}, \hat{w}]$ and $\mathbf{B} \equiv \mathbf{A}^{-1}$. The transformation for the gradient operator is thus $\nabla_{xyz} = \mathbf{B}^T \nabla_{uvw}$ and hence for Laplacian we have $\nabla^2 = \nabla_{xyz}^T \nabla_{xyz} = \nabla_{uvw}^T \mathbf{B}\mathbf{B}^T \nabla_{uvw}$ [66].

3.1.2 Ionic Potential

In the previous chapter, we demonstrated the ionic pseudopotential generation procedure. As discussed above, the ionic potential consists of two parts, one local part (V_{NA}) and one nonlocal part (the KB projectors). We now discuss how to implement these in real-space.

In real-space local potentials take the form of a real diagonal matrix. V_{NA} is isotropic for a single atom and hence we can write $V_{\text{NA}}(\mathbf{r}) = V_{\text{NA}}(r)$, that is V_{NA} only depends on the magnitude of r . We thus compute the distance r from each grid point to the center of the atomic nucleus and evaluate $V_{\text{NA}}(r)$. From the NanoBase output the function $V_{\text{NA}}(r)$ is expressed on a radial grid; therefore an interpolation procedure is needed when generating $V_{\text{NA}}(\mathbf{r})$ on the real-space grid[65].

Preparing for the nonlocal KB projectors is the most time consuming part in the Hamiltonian generation. With the presence of $\hat{V}_{\text{ps}}^{\text{SO}}$ the pseudopotential takes the general form $\sum_{\alpha, m_1, m_2} \sum_I |\chi_{\alpha I}^{m_1}\rangle \langle \chi_{\alpha I}^{m_2}|$ (c.f. Eq.2.45), where $\alpha \equiv (l, j)$, m_1 and m_2 denote different magnetic quantum numbers ($m_1 = m_2$ if without SOC), and I represents image atoms if the periodic boundary condition is imposed, i.e. $\chi_{\alpha I}^m(\mathbf{r}) = \chi_{\alpha}^m(\mathbf{r} - \mathbf{R}_I)$. In a periodic system the Bloch wavefunction $\psi_{\mathbf{k}}(\mathbf{r})$ extends in the entire real-space; therefore, it is not so obvious how to calculate $\sum_I |\chi_{\alpha I}^{m_1}\rangle \langle \chi_{\alpha I}^{m_2}| \psi_{\mathbf{k}}\rangle$. To overcome this difficulty we can show that if we define

$$\eta_{\alpha}^m(\mathbf{r}) = \sum_I e^{i\mathbf{k} \cdot \mathbf{R}_I} \chi_{\alpha}^m(\mathbf{r} - \mathbf{R}_I)$$

then $\sum_I |\chi_{\alpha I}^{m_1}\rangle \langle \chi_{\alpha I}^{m_2}| \psi_{\mathbf{k}}\rangle$ can be computed as

$$\sum_I |\chi_{\alpha I}^{m_1}\rangle \langle \chi_{\alpha I}^{m_2}| \psi_{\mathbf{k}}\rangle = \int_{0 < x, y, z < L} d^3r \eta_{\alpha}^{m_1}(\mathbf{r}) [e^{-i\mathbf{k} \cdot \mathbf{r}} \eta_{\alpha}^{m_2}(\mathbf{r})]^* u_{\mathbf{k}}(\mathbf{r}), \quad (3.4)$$

where $u_{\mathbf{k}}(\mathbf{r})$ is the periodic part of the Bloch wavefunction [$u_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}(\mathbf{r})e^{-i\mathbf{k} \cdot \mathbf{r}}$]. Here

notice that the integral domain in Eq.(3.4) is confined to one single supercell only.

In practice when setting up the Hamiltonian we calculate each $\eta_{\alpha}^m(\mathbf{r})$ and store them in a matrix[2]

$$\mathbf{\Gamma} = [\eta_{\alpha_1}^{m_{1,1}}, \eta_{\alpha_1}^{m_{1,2}}, \eta_{\alpha_1}^{m_{1,3}}, \dots, \eta_{\alpha_2}^{m_{2,1}}, \eta_{\alpha_2}^{m_{2,2}}, \dots, \eta_{\alpha_3}^{m_{3,1}}, \eta_{\alpha_3}^{m_{3,2}}, \dots],$$

which is usually quite sparse. Then the operation $\hat{V}_{\text{ps}} |\psi\rangle$ is accomplished by the matrix product

$$\hat{V}_{\text{ps}} |\psi\rangle = \mathbf{\Gamma}^a \mathbf{v}^{a,b} \mathbf{\Gamma}^{b\dagger} \Psi,$$

where $\mathbf{v}$ is a diagonal matrix storing all KB-energies (Eq.2.48) and the superscript a or b indicates that $\mathbf{\Gamma}^{a,b}$ ($\mathbf{v}^{a,b}$) may be just part of $\mathbf{\Gamma}$ ($\mathbf{v}$), c.f. Eq.(2.45); when there is no SOC we simply have $\mathbf{\Gamma}^{a,b} = \mathbf{\Gamma}$ and $\mathbf{v}^{a,b} = \mathbf{v}$.

3.1.3 Hartree Potential

In the previous chapter we have defined the Hartree potential as

$$V_H(\mathbf{r}) = \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (3.5)$$

But since we have screened the ionic potential by an artificial neutral atomic density (Eq.2.36), it is then

$$\delta V_H(\mathbf{r}) = \int d^3r' \frac{\delta \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

that needs to be calculated instead of $V_H(\mathbf{r})$. δV_H can be obtained by solving the Poisson equation

$$\nabla^2 \delta V_H(\mathbf{r}) = -4\pi \delta \rho(\mathbf{r}). \quad (3.6)$$

If the periodic boundary condition is imposed, the Poisson equation can be solved by the Fourier transform method. One first transforms $\delta \rho(\mathbf{r})$ into the reciprocal space, then solves for $\delta V_H(\mathbf{K})$, and finally revert δV_H back to real-space. The Fourier

transforms can be performed efficiently by the Fast Fourier Transform (FFT) method. The Fourier transform of the Poisson equation reads

$$-K^2 \delta V_H(\mathbf{K}) = -4\pi \delta \rho(\mathbf{K}) \quad (3.7)$$

with the transform convention

$$f(\mathbf{K}) = v_0^{-1} \int d^3r e^{-i\mathbf{K}\cdot\mathbf{r}} f(\mathbf{r})$$

$$f(\mathbf{r}) = \sum_{\mathbf{K}} f(\mathbf{K}) e^{i\mathbf{K}\cdot\mathbf{r}}$$

where $\mathbf{K}$ is the reciprocal lattice vector and v_0 is the supercell volume. For $\mathbf{K} \neq \mathbf{0}$ Eq.(3.7) can be readily solved, but $\delta V_H(\mathbf{K})$ is not well defined for $\mathbf{K} = \mathbf{0}$; this reflects the fact that the electrostatic potential is defined always up to an arbitrary constant shift. For convenience we can set $\delta V_H(\mathbf{K} = \mathbf{0}) = 0$.

When a nonperiodic (isolated) boundary condition is imposed, we need to solve Eq.(3.6) directly. In the finite difference representation, Eq.(3.6) is transformed into an equivalent algebraic system $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is the electron density, $\mathbf{A}$ is the matrix representing the Laplacian operator, and $\mathbf{x}$ is the potential to be sought. This linear system can be solved using a number of methods, direct (e.g. performing an LU decomposition) or iterative (e.g. conjugate gradient)[67]. Under the nonperiodic boundary condition, one thing nontrivial is that the potential at the boundary should be calculated accurately before solving $\mathbf{A}\mathbf{x} = \mathbf{b}$. The most straightforward approach to obtaining these boundary values is to perform numerical integrations using Eq.(3.5), but the computational cost is proportioned to the 5/3 power of the model size[68]. To solve this problem Chelikowsky *et al.*[8, 62, 69] proposed a procedure where a multipole expansion is used to express the potential at the boundaries, and it is then calculated term by term until convergence is achieved.

3.2 Eigensolvers

The bottleneck of the self-consistency loop consists in finding the eigenpairs of the Hamiltonian [c.f. Eq.(3.1)]. As mentioned before it is unrealistic to write down the entire Hamiltonian in real-space representation and diagonalize it exactly during each self-consistency iteration. Actually the algorithm can be much simplified based on two observations. The first observation is that we only need those eigenpairs below the Fermi energy because the unoccupied states do not contribute to the density, which is the only quantity relevant to our convergence criterion. The second observation is that before the final self-consistency is achieved, it makes no sense to compute very accurately every eigenpair of an interim Hamiltonian which differs from the converged one for sure. Any efficient DFT algorithm should be aware of these two observations.

In our work the Chebyshev-filtered subspace iteration (CFSI) method is used as the main eigensolver[2], which was devised for real-space DFT by Zhou *et al.*[3] for the first time. Remarkably, in CFSI the explicit matrix form of the Hamiltonian never shows up, only the operation $\hat{H}|\psi\rangle$ being needed, and during each self-consistency iteration we use the filtering method to refine the occupied subspace only, which is much smaller than the entire eigen-space. In the filtering process the accuracy can be controlled by the Chebyshev-filter degree, that is, the eigenpairs are calculated only approximately. Therefore, starting with an initial guess for the occupied subspace and then iterating the self-consistency loop, the subspace gradually approaches the true occupied subspace of the converged Hamiltonian.

To illustrate the basic working principle of CFSI, in what follows we first present the power method and the inverse iteration method, and then we present the complete algorithm of CFSI.

3.2.1 Power Method and Inverse Iteration

Since the eigenvectors of a Hermitian matrix A form a complete orthonormal basis $[v_1, \dots, v_n]$, we can thus express a given vector as

$$y = \sum_i a_i v_i.$$

If we apply the operator A^k to y we obtain

$$y = \sum_i \lambda_i^k a_i v_i = \lambda_n^k \sum_i \left(\frac{\lambda_i}{\lambda_n} \right)^k a_i v_i$$

where $|\lambda_i| < |\lambda_{i+1}|$ are the eigenvalues. From this expression, it is obvious that the eigenvector with the largest absolute eigenvalue will eventually dominate and y will become parallel to v_n . The algorithm goes as follows:[70]

Algorithm 1 Power Method

generate an initial random guess for the eigenvector $y = y_0$

while 1 do

$v = y / \|y\|$

$y = Av$

$\theta = v'y$

if $\|y - \theta v\| \leq \varepsilon$

break

end if

end while, set $\lambda = \theta$ and $x = v$

Using the numerical power method we obtain the largest eigenpair $\{\lambda_n, x_n\}$ in the end.

The rate of convergence of this algorithm is derived from the following observation. The largest eigenvector v_n eventually wins, but it has to dwarf the second largest one

v_{n-1} and hence the convergence rate is λ_{n-1}/λ_n .

Though the power method is efficient, it only gives the eigenpair at the top of the spectrum. If we are interested in the eigenpair around a certain number σ , we can perform the following inverse iteration algorithm[70], which is based on the observation that

$$Ax = \lambda x \implies (A - \sigma I)^{-1}x = (\lambda - \sigma)^{-1}x$$

in the case where $(A - \sigma I)$ is invertible.

Algorithm 2 Inverse Iteration

generate an initial random guess for the eigenvector $y = y_0$

while 1 **do**

$$v = y/\|y\|$$

$$y = (A - \sigma I)^{-1}v$$

$$\theta = v'y$$

$$\mathbf{if} \ \|y - \theta v\| \leq \varepsilon$$

break

end if

end while, set $\lambda = \sigma + 1/\theta$ and $x = v$

Following the analysis of the power method, the convergence rate is found to be $(\lambda_k - \sigma)/(\lambda_l - \sigma)$ where λ_k and λ_l are the closest and second closest eigenvalues to σ respectively.

What we have learnt from the inverse iteration and the power method is that suppose we are interested in an eigenpair of A close to some number σ then we simply build a function $p(A)$ which has a particularly large magnitude at σ while suppressing other part of the eigenspectrum of A . By applying $p(A)$ to any vector x iteratively, only the wanted eigenvector will finally survive and its corresponding eigenvalue is then simply computed from $\lambda = v^\dagger Av$.

The above two algorithms focus on computing only one eigenpair of interest, but in the electronic structure problem all the eigenpairs below Fermi energy are wanted. Naively one may try computing these eigenpairs one by one using the inverse iteration method to sweep the spectrum below Fermi energy. This is apparently inefficient. Moreover, in real electronic structure problems there are always some degenerate eigenpairs, which are unlikely to be distinguished in this way. A better way to do is to use the block method together with orthogonality constraints. We illustrate this method along with the power method as follows (same thing applies to the inverse iteration):[70]

Algorithm 3 Orthogonal Iteration

generate an initial random guess for the orthonormal block vector V

while 1 do

$Y = AV$

$\Theta = V'Y$

if $\|y - \theta v\| \leq \varepsilon$

break

end if

orthonormalize Y

end while, set $\Lambda = \Theta$ and $X = V$

Note that the orthonormalization step is essential in this algorithm. If there were no such constraint, every vector in the block would eventually fall into the one with the largest eigenvalue (absolute value). With this constraint imposed, the vectors are constantly “pushed away” from each other so that no two vectors would fall into the same “pit”; this ensures that these vectors will finally span out the desired eigensubspace. The orthonormality can be maintained using any algorithm. For example, dividing Y by the Cholesky factor of $Y^\dagger Y$ is one efficient way.

3.2.2 Chebyshev-Filtered Subspace Iteration

Now we are in a good position to see how CFSI works. The Chebyshev-filter is nothing but a carefully designed matrix function $p(A)$ suitable for electronic structure calculations, plus a Rayleigh-Ritz step which is only a little bit more advanced than the orthogonalization procedure.

The well-known Chebyshev polynomial of the first kind is defined as [3, 71, 72]

$$C_m(t) = \begin{cases} \cos(m \cos^{-1}(t)), & -1 \leq t \leq 1 \\ \cosh(m \cosh^{-1}(t)), & |t| > 1 \end{cases}$$

Note that $C_0(t) = 1, C_1(t) = t$. The following recurrence relation is easy to derive from properties of $\cos(t)$ and $\cosh(t)$,

$$C_{m+1}(t) = 2tC_m(t) - C_{m-1}(t). \quad (3.8)$$

The general property is depicted in Fig. . The Chebyshev polynomial grows rapidly outside the region $[-1,1]$; the higher the polynomial degree m , the more rapidly it grows. Therefore, we can utilize this feature of Chebyshev polynomial to calculate the wanted eigenpairs if we can map the “useless” part of the Hamiltonian spectrum (the unoccupied states) into $[-1,1]$ and the lower end of the spectrum (the occupied states) into $[-\infty, -1)$. This can be accomplished by a simple linear mapping:

$$l(t) := \frac{t - c}{e} \quad c = \frac{a + b}{2}, e = \frac{b - a}{2},$$

where a is an energy value slightly above the Fermi energy and b is an upper bound of the entire Hamiltonian spectrum. Note that it is important to make sure that b is greater than the largest eigenvalue of Hamiltonian and that a is greater than the highest occupied energy level. The estimation of b can be done using the Lanczos algorithm (see details in Appendix B) and a can be set equal to the largest Ritz value (see Algorithm 5 below) obtained from the last self-consistency iteration. a and b are

updated in each self-consistency iteration.

With this mapping $l(t) < -1$ for $t < a$ and $-1 < l(t) < 1$ for $a < t < b$. Then, just like what we did in Algorithms 1-3, we simply compute

$$Y = p_m(H)X, \text{ where } p_m(t) = C_m \left(\frac{t - c}{e} \right).$$

In practice, this is actually accomplished by the following algorithm (notice the recurrence relation Eq.3.8):[2, 3]

Algorithm 4 $Y = \text{Chebyshev_filter}(H, X, m, a, b)$

$e = (b - a)/2$, $c = (b + a)/2$

$\psi_1 = (HX - cX)/e$

for $i = 2, \dots, m$ **do**

$\psi_2 = 2(H\psi_1 - c\psi_1)/e - X$

$X = \psi_1$, $\psi_1 = \psi_2$

end for

$Y = \psi_2$

To avoid all the eigenpairs falling into the bottom of the spectrum, an orthogonalization step is needed following the Chebyshev-filter. Instead of direct orthogonalization, in fact the Rayleigh-Ritz method is more preferred, because it not only orthonormalizes the vectors but also serves to refine the subspace[3, 5]. The complete CFSI algorithm in one self-consistency iteration is presented in Algorithm 5.

The Rayleigh-Ritz method has been widely used in a number of DFT codes based on different subspace iteration methods and it has been shown to improve both the stability and the convergence rate[2, 5, 73, 74]. The basic mechanism in the Rayleigh-Ritz method is that if the block vector Φ spans an eigen-subspace of the Hamiltonian, then the Ritz values D are equal to the corresponding true eigenvalues and after

subspace rotation the block vector ΦQ will be the true eigenvectors.

Algorithm 5 Chebyshev – filtered subspace iteration (CFSI) method

1. Get the lower bound a from previous Ritz values (largest one)
 2. Use Lanczos to get an upper bound b of the spectrum of the current H
 3. Chebyshev – filter the subspace block vector Φ which contains the current trial wavefunctions $\Psi_i(\mathbf{r})$, $i = 1, \dots, s$: $\Phi = \text{Chebyshev_filter}(H, \Phi, m, a, b)$
 - (4. Orthonormalize Φ)
 5. Perform Rayleigh – Ritz step :
 - (a) Compute $\tilde{H} = \Phi^T H \Phi$, $\tilde{S} = \Phi^T \Phi$
 - (b) Solve $\tilde{H}Q = \tilde{S}QD$, where D contains the Ritz values of H
 - (c) Subspace rotation $\Phi := \Phi Q$
-

As can be seen the central task of CFSI is to refine the subspace Φ iteratively. But how do we get an initial guess for Φ at the beginning? It has been shown that the quality of the initial guess strongly affects the overall efficiency of the algorithm[2, 3, 75]. In practice, to construct the initial Φ we first fill the blockvector with single- or double-zeta atomic orbitals originally designed for LCAO calculations, and then we perform a Rayleigh-Ritz refinement with respect to the initial Hamiltonian to get the initial subspace[2]. Alternatively, one can start with a full LCAO calculation and use the LCAO solution as the initial subspace for the following real-space calculation. For spin polarized calculations this latter procedure is recommended.

Tests and Applications

To test our code as well as to verify the real-space method for spin-polarized DFT calculations, the numerical scheme demonstrated above is applied to perform simulations on several systems. As the GGA functional is known to tend to stabilize a magnetic state over the nonmagnetic state[76], in what follows we use GGA for all of our simulations.

4.1 *Collinear-spin Calculations: Ni and Fe*

As the first part of our application, we have used our code to investigate the magnetization of bulk Ni and Fe, which are the most typical magnetic metals with a collinear-spin configuration. Ni has an FCC lattice with $a = 3.52\text{\AA}$ and Fe has a BCC lattice with $a = 5.42\text{\AA}$.

In the self-consistency calculations we use an orthogonal real-space mesh with the resolution of $\sim 6\text{points}/\text{\AA}$, for both systems. A convergence criterion of 10^{-5}Hartree for the total energy difference has been used in self-consistency calculations. As Ni and Fe are metallic bulk systems, we have found that the use of the metric Eq.(2.30) in our charge mixers is crucial for suppressing the charge-sloshing effect and for improving the convergence performance. The calculated magnetic moment per Ni/Fe atom is shown in Table 4.1. We notice that our result for Ni agrees very well with the experimental value, but our result for Fe shows a noticeable overestimation. The

Table 4.1: Magnetic moment per atom $M_{at}(\mu_B)$ from experiment and calculated by different DFT methods.

Method	Ni	Fe
LAPW-NCPP[77]	0.64	2.32
Real-space-NCPP (ours)	0.61	2.55
Exp.[78]	0.61	2.22

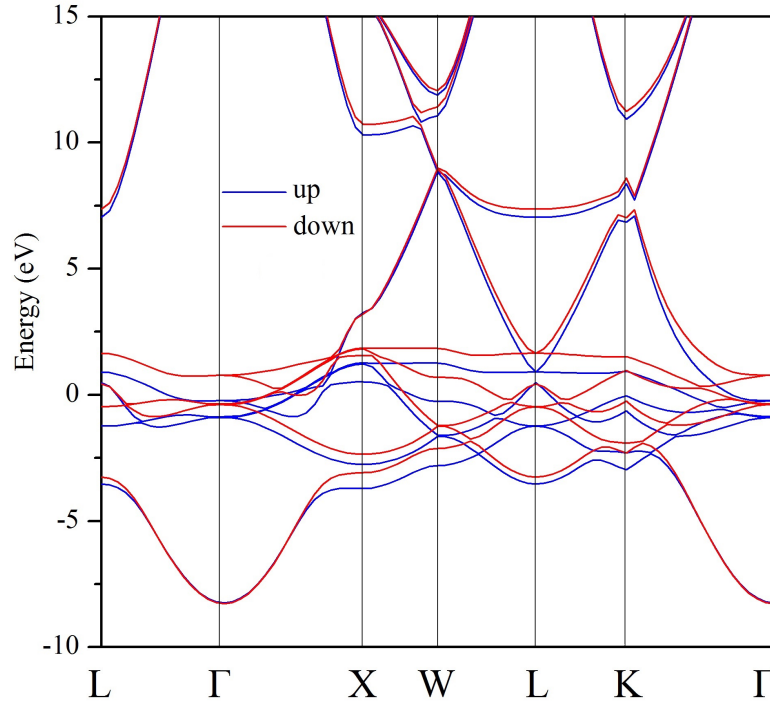


Figure 4.1: Bandstructure of Ni (FCC) calculated with our real-space method.

reason for this overestimation can be multiple and is not clear yet; this will be the next stage of our work.

We have also calculated the spin-resolved bandstructures of Ni and Fe, see Fig.4.1 and Fig.4.2 respectively. The first Brillouin zones (FBZ) are shown in Fig.4.3.

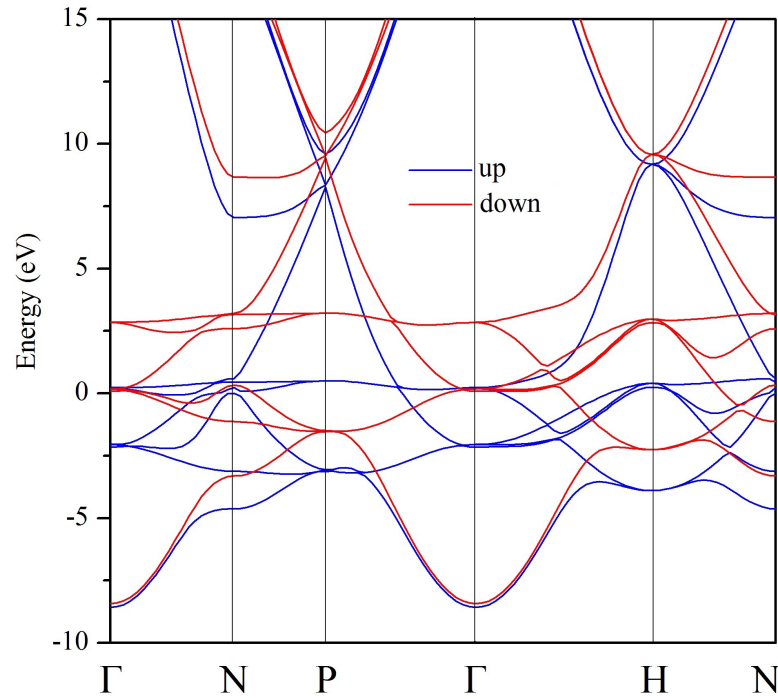


Figure 4.2: Bandstructure of Fe (BCC) calculated with our real-space method.

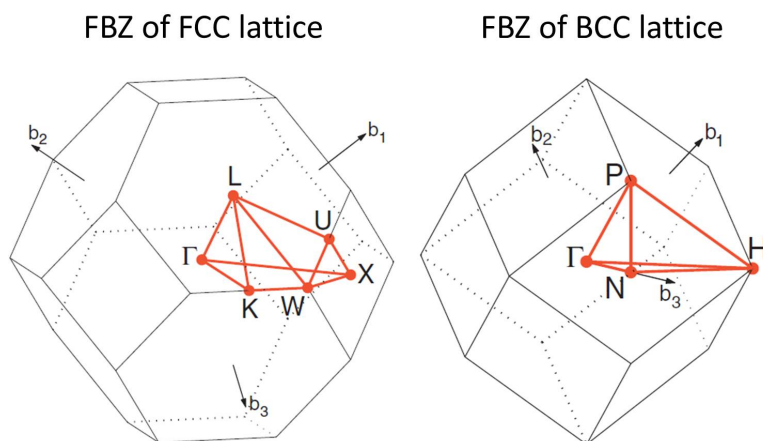


Figure 4.3: FBZs of FCC and BCC lattices.

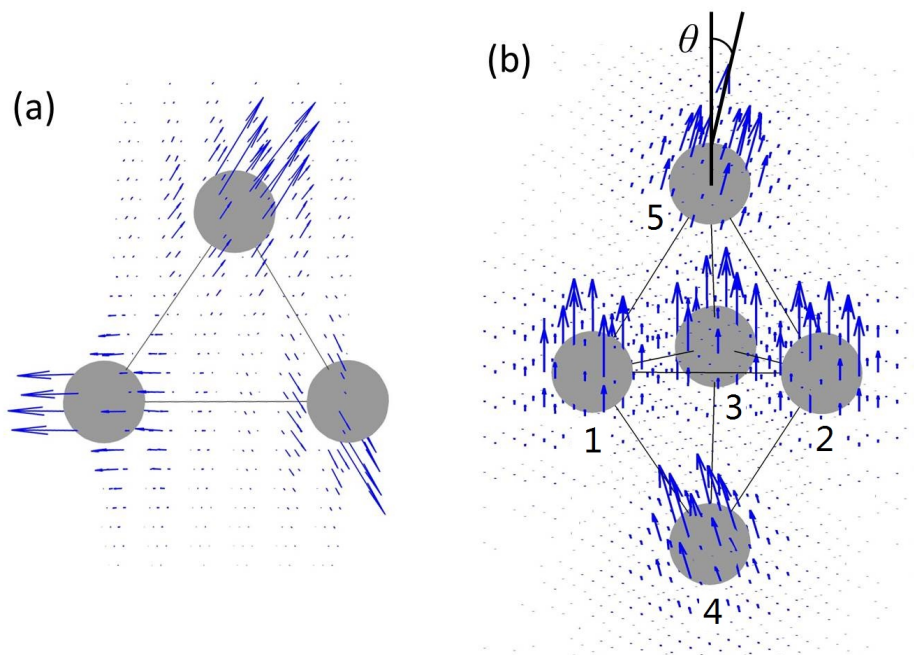


Figure 4.4: Non-collinear spin texture of Cr_3 (a) and Fe_5 (b) clusters.

4.2 Noncollinear-spin Calculations: Fe_5 and Cr_3 clusters

The Cr_3 trimer (Fig.4.4a) and the Fe_5 bipyramidal pentamer (Fig.4.4b) are magnetic clusters that are known to have a noncollinear-spin texture in one of their stable states[76, 79]. We thus use these two systems as the benchmark tests for our noncollinear-spin implementation. A dense real-space mesh with the resolution of $\sim 10\text{points}/\text{\AA}$ has been used for both systems. Vacuum buffer layers are added around the clusters in order to confine the wavefunctions within the simulation box, the size of which is set to be $11\text{\AA}$ in each dimension.

In our calculation of the Cr_3 trimer we adopt the geometric constant from Ref.[76]: the bond-lengths are equally taken to be $2.51\text{\AA}$, at which the trimer is known to have a spiral noncollinear-spin ground-state[76]. We set the initial magnetic moment direction of each Cr atom to be $\pm 15^\circ$ off the correct value. After convergence is reached in our DFT calculation, we see the magnetization be relaxed to the correct spiral distribution with a 3-fold symmetry (Fig.4.4a) with each atomic magnetic moment equal to $3.89\mu_B$ (to be compared with the previously reported value $4.11\mu_B$ obtained

Table 4.2: The calculated and previously reported atomic magnetic moment (M_{at}, θ) of each atom in the Fe_5 pentamer. The numbers are ordered according to atom 1-5 (Fig.4.4b)

	$M_{at}(\mu_B)$	θ
PAW[76]	3.04,3.04,3.04,3.06,3.06	$0^\circ, 0^\circ, 0^\circ, 31.3^\circ, 31.3^\circ$
ours	3.13,3.13,3.13,3.18,3.18	$0^\circ, 0^\circ, 0^\circ, 27.8^\circ, 27.8^\circ$

with the projector augmented wave method (PAW)[76]).

The Fe_5 pentamer is known to have a noncollinear-spin ground-state with the bond-lengths $l_{12} = l_{32} = l_{13} = 2.38\text{\AA}$ and the rest equal to $2.33\text{\AA}$ [76]. Thus this set of geometric constants is used in our calculation. The initial guess for the magnetic moments of the two apical atoms is again set to be $\pm 15^\circ$ off the final solution. After full relaxation these two magnetic moments eventually rotate to a stable angle $\theta_4 = \theta_5 = 27.8^\circ$ in opposite directions but still stay coplanar (Fig.4.4b). The magnetic moments of the other three atoms in the middle are aligned parallel, in a way similar to ferromagnetism. These results mean that our real-space noncollinear-spin method is capable of capturing both the antiferromagnetic-like exchange between the apical atoms and the ferromagnetic exchange among the central atoms. To make a comparison we list our result for the Fe_5 pentamer in table 4.2, together with the PAW result reported in Ref.[76].

4.3 Spin-orbit Coupling Calculations: *SnTe* and *TaAs*

The band splitting effect due to spin-orbit coupling (SOC) is most significant in materials containing heavy atoms. In this section we present our calculation results of the bandstructures of tin telluride (SnTe) and titanium arsenide (TaAs), both of which are well-known topological materials: SnTe is known to be a topological crystalline insulator[80] and TaAs is a newly discovered Weyl semimetal[1].

As a fully self-consistent noncollinear-spin calculation is often hard to get con-

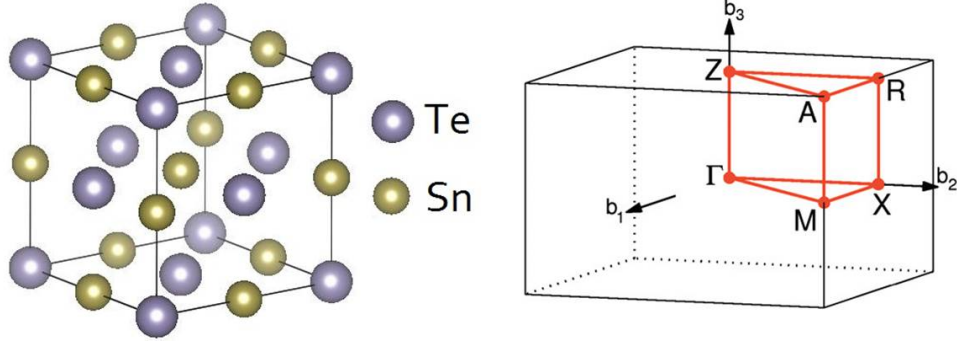


Figure 4.5: The conventional cell and the corresponding first Brillouin zone of SnTe.

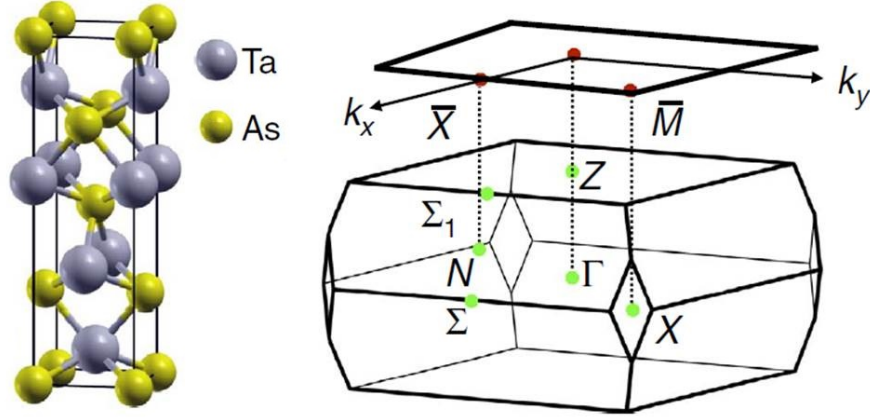


Figure 4.6: Body-centered tetragonal structure of TaAs and its first Brillouin zone. Figure from Ref.[1].

verged, here we treat the SOC effect as a perturbation when calculating the bandstructures. Namely, we first perform a self-consistent calculation to get the converged effective potential without SOC taken into account, and then j -dependent pseudopotential terms are added when calculating the bandstructures. This two-step procedure is justified because SOC has a very small effect on the effective potential considering its order of magnitude. The unit cell and the corresponding first Brillouin zone of SnTe (TaAs) used in the simulation are shown in Fig.4.5 (Fig.4.6). SnTe has a simple rocksalt crystal structure with the lattice constant $6.403\text{\AA}$. In order to avoid real-space interpolations when calculating the bandstructure, the simple cubic conventional cell is used as the unit cell in our simulation. TaAs has a body-centered tetragonal crystal structure of the space group $I4_1md$ with lattice constants $a = 3.437\text{\AA}$ and $c = 11.656\text{\AA}$.

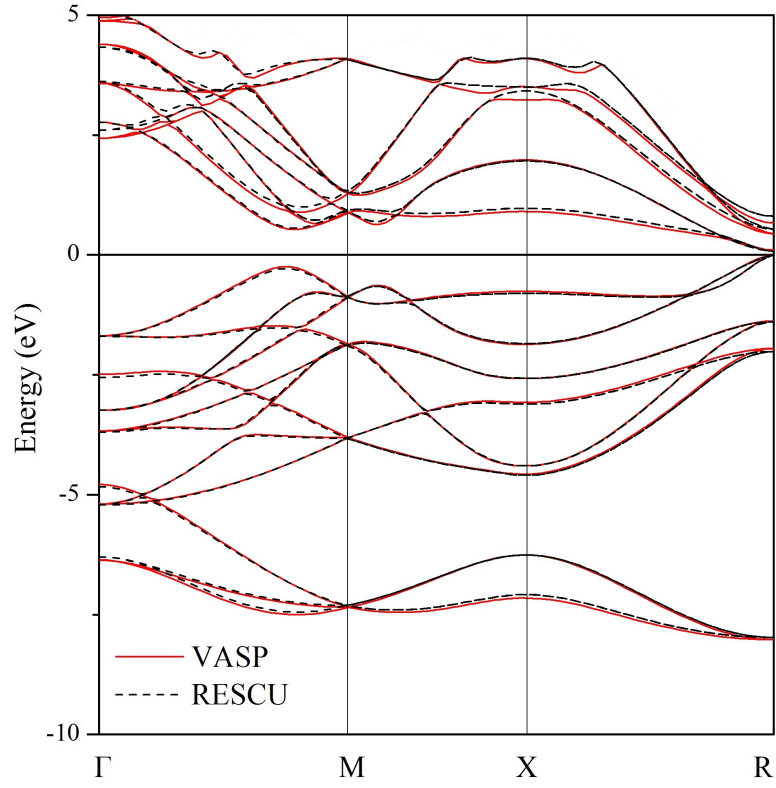


Figure 4.7: Bandstructure of SnTe calculated with our real-space method and with PAW.

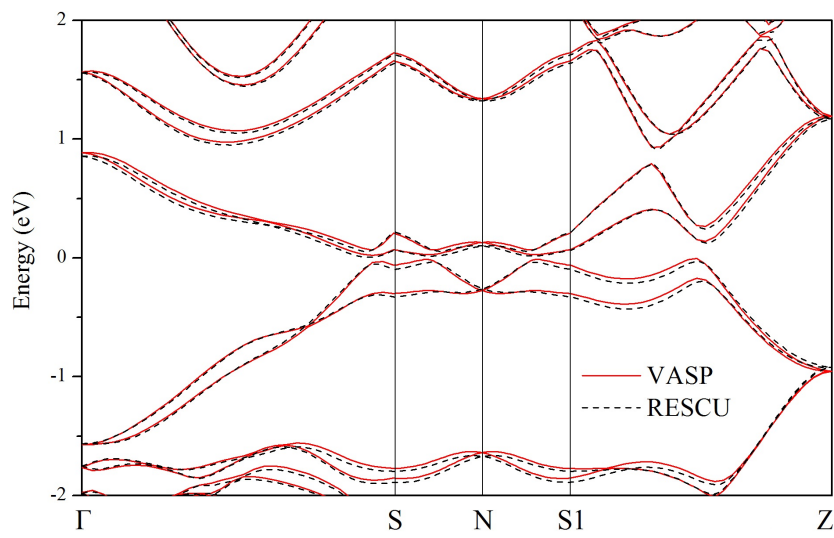


Figure 4.8: Bandstructure of TaAs calculated with our real-space method and with PAW.

The calculated bandstructure is shown in Fig.4.7 (Fig.4.8), where the comparison to the standard bandstructure obtained with PAW (implemented in the VASP package[45]) is also provided. As for TaAs, from the comparison we see that there is a bit deviation from the standard PAW result. There could be two possible origins for this deviation. One is the error induced when we interpolate the effective potential expressed on the orthogonal real-space grid onto the nonorthogonal grid of the TaAs primitive cell. The other possible origin is that our norm-conserving pseudopotentials is unable to handle the d-electrons in Ta and As as well as the PAW pseudopotentials[43].

Conclusions

This thesis is devoted to presenting an implementation of the various spin-involved first-principles calculations, including collinear, noncollinear, and SOC calculations, within the real-space DFT framework.

We started by reviewing the basic quantum many-body formulation of a general many-electron problem and we lay our emphasis on the fundamental difficulty in solving the problem. We then reviewed the Hartree-Fock method, where the concepts of exchange and correlation were clarified. Moving to the Density Functional Theory, we demonstrated how a many-electron problem can be formally cast down to a single-body problem with an opaquely defined XC-potential as the price. To proceed we resort to the L(S)DA or the GGA scheme for approximating the XC-functional. In particular we demonstrated the extension of collinear-spin XC-functionals to the non-collinear case. Besides XC-functionals the accuracy of a DFT method also depends on how the atomic core electrons are treated therein. Here a real-space pseudopotential method has been employed to describe the ions. As relativistic effects mainly reside in the core region of an atom, the SOC effect can be incorporated by generating the Kleinman-Bylander orbitals from solving the atomic Dirac equation.

Next, we demonstrated the implementation of each of the Hamiltonian operators on a uniformly meshed real-space grid. Our eigensolver is based on the polynomial-filtering method. From a theoretical point of view, the advantages of using the polynomial-filtering method are that the full real-space representation of the Hamil-

tonian is not necessary and that it allows one to only focus on a relatively small eigen-subspace. The machinery of the polynomial-filtering method was exposed and the complete eigensolver algorithm was then presented.

To verify our method we have applied it onto several systems most representative within the topic. We compared our results with previously reported results or with those calculated from other DFT softwares and good agreements have been found. We have shown that our scheme for the noncollinear-spin XC-functionals is capable of capturing the exchange effects in the noncollinear-spin configuration and that our norm-conserving j -dependent pseudopotential method leads to a correct description of the SOC induced band-splitting phenomenon.

Even though all the numerical tests of this work were performed on simple systems only, we have already observed the drastically increased convergence difficulties after the spin degree of freedom is included. As the overall convergence performance largely depends on the mixer and the mixer is highly associated with the specific basis set that has been chosen, needless to say a better real-space mixer is urgently demanded at this point.

A

SPIN-ORBIT COUPLING WITH LCAO BASIS

In this appendix we sketch the scheme how SOC computation is implemented within the LCAO formalism.

We start with Eq.(2.43). By utilizing the general identity

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} |j, M\rangle = \begin{cases} \frac{l}{2} |l + \frac{1}{2}, M\rangle & \text{for } j = l + 1/2 \\ -\frac{l+1}{2} |l - \frac{1}{2}, M\rangle & \text{for } j = l - 1/2 \end{cases}$$

Eq.(2.43) can be transformed into

$$\begin{aligned} \hat{V}_{\text{ion}}^{\text{PS}} &= \sum_{lm} |lm\rangle \bar{V}_l^{\text{SC}} \langle lm| + \sum_l \bar{V}_l^{\text{SO}}(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \\ &\quad \left(\sum_M \left| l + \frac{1}{2}, M \right\rangle \left\langle l + \frac{1}{2}, M \right| + \sum_{M'} \left| l - \frac{1}{2}, M' \right\rangle \left\langle l - \frac{1}{2}, M' \right| \right) \\ &= \sum_{lm} \left(\bar{V}_l^{\text{SC}} + \bar{V}_l^{\text{SO}}(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \right) |lm\rangle \langle lm| \end{aligned} \quad (\text{A.1})$$

In getting the last equality we have used the completeness relation

$$\sum_M \left| l + \frac{1}{2}, M \right\rangle \left\langle l + \frac{1}{2}, M \right| + \sum_{M'} \left| l - \frac{1}{2}, M' \right\rangle \left\langle l - \frac{1}{2}, M' \right| = \sum_m |lm\rangle \langle lm| \otimes I_{\text{spin}}.$$

Now we identify the SOC part of the total pseudopotential as

$$\hat{V}^{\text{SO}} = \sum_{\mathbf{R}} \hat{V}^{\text{SO}}(\mathbf{R}) = \sum_{\mathbf{R}lm} \bar{V}_l^{\text{SO}}(\mathbf{r} - \mathbf{R}) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} |lm(\mathbf{R})\rangle \langle lm(\mathbf{R})|.$$

In the LCAO method the matrix elements of $\hat{V}^{\text{SO}}$ under the atomic orbital basis are wanted, that is we need to calculate $\langle \phi_{i\sigma} | \hat{V}^{\text{SO}} | \phi_{j\sigma'} \rangle$, where i represents a combination of indices $i \equiv (\mathbf{R}_i, n_i, l_i, m_i)$ with $\mathbf{R}_i$ being the center of $\phi_{i\sigma}(\mathbf{r})$. Note that atomic orbitals on the same site with different n values are not necessarily orthogonal.

The calculation of $\langle \phi_{i\sigma} | \hat{V}^{\text{SO}} | \phi_{j\sigma'} \rangle$ can be largely simplified based on the following observation[61]: as $\bar{V}_l^{\text{SO}}(r - \mathbf{R})$ is very short-ranged and localized around its center $\mathbf{R}$, $\langle \phi_{i\sigma} | \hat{V}^{\text{SO}}(\mathbf{R}) | \phi_{j\sigma'} \rangle$ becomes nonzero only if $\mathbf{R} = \mathbf{R}_i = \mathbf{R}_j$. Thus the SOC term only adds an on-site (block diagonal) contribution to the total Hamiltonian. For the block of $\mathbf{R}_i$ ($\mathbf{R}_j$) site we have

$$\begin{aligned} \langle \phi_{i\sigma} | \hat{V}^{\text{SO}}(\mathbf{R}_i) | \phi_{j\sigma'} \rangle &\approx \delta_{\mathbf{R}_i, \mathbf{R}_j} \sum_{lm} \langle \phi_i | \bar{V}_l^{\text{SO}}(r - \mathbf{R}_i) \hat{L}_{\sigma\sigma'} | lm(\mathbf{R}_i) \rangle \langle lm(\mathbf{R}_i) | \phi_j \rangle \\ &= \delta_{\mathbf{R}_i, \mathbf{R}_j} \delta_{l_i, l_j} \int dr r^2 R_{n_i, l_i}(r) \bar{V}_{l_i}^{\text{SO}}(r) R_{n_j, l_j}(r) \\ &\quad \cdot \langle l_i m_i(\mathbf{R}_i) | \hat{L}_{\sigma\sigma'} | l_j m_j(\mathbf{R}_j) \rangle, \end{aligned} \quad (\text{A.2})$$

where we have used the notation

$$\hat{L}_{\sigma\sigma'} = \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2} \begin{pmatrix} \hat{L}_z & \hat{L}_- \\ \hat{L}_+ & -\hat{L}_z \end{pmatrix}$$

$$\hat{L}_+ |l, m\rangle = \sqrt{(l-m)(l+m+1)} |l, m+1\rangle$$

$$\hat{L}_- |l, m\rangle = \sqrt{(l+m)(l-m+1)} |l, m-1\rangle.$$

In LCAO method usually the real spherical harmonics are used instead of their complex counterparts. The transformation convention that we adopt is

$$\bar{Y}_{l,m} = \begin{cases} \frac{1}{\sqrt{2}} [Y_{l,-m} + (-1)^m Y_{l,m}] & \text{if } m > 0 \\ \frac{i}{\sqrt{2}} [Y_{l,-|m|} - (-1)^m Y_{l,|m|}] & \text{if } m < 0 \\ Y_{l,0} & \text{if } m = 0 \end{cases}$$

where $\bar{Y}_{l,m}$ is the real spherical harmonic. This transformation should be noticed when calculating $\hat{L}_{\sigma\sigma'} |lm\rangle$.

B

LANCZOS ALGORITHM

The Lanczos algorithm automatically generates an orthogonal basis in which the given Hamiltonian H is tridiagonal. It is especially powerful for generating the highest eigenpairs of large matrices, which can be utilized in our real-space method to estimate an upper bound of the electronic spectrum[2, 3, 65]. The algorithm proceeds as follows[22]: starting with a normalized trial vector v_1 , form a second vector $v_2 = C_2[Hv_1 - T_{11}v_1]$, where $T_{11} = v_1^\dagger H v_1$ and C_2 is chosen so that v_2 is normalized. Note that v_2 is orthogonal to v_1 . Subsequent vectors of the basis are generated recursively by

$$v_{n+1} = C_{n+1}[Hv_n - T_{nn}v_n - T_{n,n-1}v_{n-1}], \quad (\text{B.1})$$

where $T_{ij} = v_i^\dagger H v_j$. The matrix T is tridiagonal since Eq.(B.1) shows that H operating on v_n yields only terms proportional to v_n , v_{n-1} , and v_{n+1} . Furthermore, it can be shown that each vector v_n is orthogonal to all other vectors[22]. Going to step M yields a tridiagonal matrix

$$T = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & & & \\ & & \ddots & \ddots & \\ & & \ddots & \alpha_{M-1} & \beta_{M-1} \\ & & & \beta_{M-1} & \alpha_M \end{bmatrix}$$

where α and β are determined via Eq.(B.1). The wanted upper bound u for our problem is then estimated by diagonalizing T and finding its highest eigenvalue.

A few iterations M is usually enough for a good estimation[3, 65]. The complete algorithm used in the program is sketched as follows[65]:

Algorithm 6 Lanczos Algorithm

generate an initial guess for smallest eigenvector $v = r$

set $\beta_0 = \|r\|$

for $j = 1 : \text{niter}$ **do**

$v_j = r / \beta_{j-1}$

$r = H v_j$

$r = r - v_{j-1} \beta_{j-1}$

$\alpha_j = v_j^\dagger r$

$r = r - v_j \alpha_j$

$\beta_j = \|r\|$

diagonalize $T = S \Theta S^\dagger$

test convergence

end for

set $u = \max(\Theta)$

BIBLIOGRAPHY

- [1] Shin-Ming Huang, Su-Yang Xu, Ilya Belopolski, Chi-Cheng Lee, Guoqing Chang, BaoKai Wang, Nasser Alidoust, Guang Bian, Madhab Neupane, Chenglong Zhang, et al. A Weyl fermion semimetal with surface fermi arcs in the transition metal mononictide taas class. *Nature communications*, 6, 2015.
- [2] Vincent Michaud-Rioux, Lei Zhang, and Hong Guo. Rescu: A real space electronic structure method. *Journal of Computational Physics*, 307:593–613, 2016.
- [3] Yunkai Zhou, Yousef Saad, Murilo L Tiago, and James R Chelikowsky. Self-consistent-field calculations using Chebyshev-filtered subspace iteration. *Journal of Computational Physics*, 219(1):172–184, 2006.
- [4] F Zirkelbach, P-Y Prodhomme, Peng Han, R Cherian, and G Bester. Large-scale atomic effective pseudopotential program including an efficient spin-orbit coupling treatment in real space. *Physical Review B*, 91(7):075119, 2015.
- [5] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.
- [6] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.
- [7] J Callaway and NH March. Density functional methods: theory and applications. *Solid State Physics*, 38:135–221, 1984.

- [8] James R Chelikowsky, N Troullier, and Y Saad. Finite-difference-pseudopotential method: electronic structure calculations without a basis. *Physical review letters*, 72(8):1240, 1994.
- [9] Ajit Srivastava, Meinrad Sidler, Adrien V Allain, Dominik S Lembke, Andras Kis, and A Imamoğlu. Valley zeeman effect in elementary optical excitations of monolayer WSe₂. *Nature Physics*, 2015.
- [10] SJ Youn and AJ Freeman. First-principles electronic structure and its relation to thermoelectric properties of Bi₂Te₃. *Physical Review B*, 63(8):085112, 2001.
- [11] M Zahid Hasan and Charles L Kane. Colloquium: topological insulators. *Reviews of Modern Physics*, 82(4):3045, 2010.
- [12] Shin-Ming Huang, Su-Yang Xu, Ilya Belopolski, Chi-Cheng Lee, Guoqing Chang, BaoKai Wang, Nasser Alidoust, Guang Bian, Madhab Neupane, Chenglong Zhang, et al. A Weyl fermion semimetal with surface fermi arcs in the transition metal mononictide taas class. *Nature communications*, 6, 2015.
- [13] Leonard Kleinman. Relativistic norm-conserving pseudopotential. *Physical Review B*, 21(6):2630, 1980.
- [14] Friedhelm Bechstedt. *Many-Body Approach to Electronic Excitations*, volume 181. Springer, 2014.
- [15] Henrik Bruus and Karsten Flensberg. *Many-body quantum theory in condensed matter physics: an introduction*. OUP Oxford, 2004.
- [16] Gerald D Mahan. *Many-particle physics*. Springer Science & Business Media, 2013.
- [17] Ferdi Aryasetiawan and Olle Gunnarsson. The gw method. *Reports on Progress in Physics*, 61(3):237, 1998.

-
- [18] Prof. Dr. Eberhard K. U. Gross (auth.) Prof. Dr. Reiner M. Dreizler. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer Berlin Heidelberg, 1990.
- [19] Gabriele Giuliani and Giovanni Vignale. *Quantum theory of the electron liquid*. Cambridge university press, 2005.
- [20] Otfried Madelung. *Introduction to solid-state theory*, volume 2. Springer Science & Business Media, 2012.
- [21] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.
- [22] Richard M Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2004.
- [23] Mel Levy, John P Perdew, and Virah Sahni. Exact differential equation for the density and ionization energy of a many-particle system. *Physical Review A*, 30(5):2745, 1984.
- [24] JF Janak. Proof that $\partial E / \partial n_i = \varepsilon_i$ in density-functional theory. *Physical Review B*, 18(12):7165, 1978.
- [25] David M Ceperley and BJ Alder. Ground state of the electron gas by a stochastic method. *Physical Review Letters*, 45(7):566, 1980.
- [26] Murray Gell-Mann and Keith A Brueckner. Correlation energy of an electron gas at high density. *Physical Review*, 106(2):364, 1957.
- [27] John P Perdew and Yue Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244, 1992.
- [28] John P Perdew and Alex Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10):5048, 1981.

- [29] SH Vosko, LI Wilk, and M Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of physics*, 58(8):1200–1211, 1980.
- [30] John P Perdew and Kieron Burke. Comparison shopping for a gradient-corrected density functional. *International journal of quantum chemistry*, 57(3):309–319, 1996.
- [31] JA White and DM Bird. Implementation of gradient-corrected exchange-correlation potentials in car-parrinello total-energy calculations. *Physical Review B*, 50(7):4954, 1994.
- [32] Miguel AL Marques, Micael JT Oliveira, and Tobias Burnus. Libxc: A library of exchange and correlation functionals for density functional theory. *Computer physics communications*, 183(10):2272–2281, 2012.
- [33] D Hobbs, G Kresse, and J Hafner. Fully unconstrained noncollinear magnetism within the projector augmented-wave method. *Physical Review B*, 62(17):11556, 2000.
- [34] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics*. Addison-Wesley, 2011.
- [35] <http://www.openmx-square.org>
- [36] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169, 1996.
- [37] GP Kerker. Efficient iteration scheme for self-consistent pseudopotential calculations. *Physical Review B*, 23(6):3082, 1981.
- [38] Peter Pulay. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters*, 73(2):393–398, 1980.

-
- [39] GB Bachelet, DR Hamann, and M Schlüter. Pseudopotentials that work: From H to Pu. *Physical Review B*, 26(8):4199, 1982.
- [40] Supriyo Datta. *Quantum transport: atom to transistor*. Cambridge University Press, 2005.
- [41] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 43:1494–1497, Nov 1979.
- [42] David Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B*, 41(11):7892, 1990.
- [43] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953–17979, Dec 1994.
- [44] J Enkovaara, C Rostgaard, J J Mortensen, and J Chen et.al. Electronic structure calculations with GPAW: a real-space implementation of the projector augmented-wave method. *Journal of Physics: Condensed Matter*, 22(25):253202, 2010.
- [45] Georg Kresse and D Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B*, 59(3):1758, 1999.
- [46] Norman Troullier and José Luís Martins. Efficient pseudopotentials for plane-wave calculations. *Physical review B*, 43(3):1993, 1991.
- [47] GP Kerker. Non-singular atomic pseudopotentials for solid state applications. *Journal of Physics C: Solid State Physics*, 13(9):L189, 1980.
- [48] <http://www.nanoacademic.ca/>
- [49] E Engel, A Höck, RN Schmid, RM Dreizler, and N Chetty. Role of the core-valence interaction for pseudopotential calculations with exact exchange. *Physical Review B*, 64(12):125111, 2001.

- [50] Steven G Louie, Sverre Froyen, and Marvin L Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Physical Review B*, 26(4):1738, 1982.
- [51] David Vanderbilt. Optimally smooth norm-conserving pseudopotentials. *Physical Review B*, 32(12):8412, 1985.
- [52] Mike C Payne, Michael P Teter, Douglas C Allan, TA Arias, and JD Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Reviews of Modern Physics*, 64(4):1045, 1992.
- [53] Leonard Kleinman and DM Bylander. Efficacious form for model pseudopotentials. *Physical Review Letters*, 48(20):1425, 1982.
- [54] Xavier Gonze, Roland Stumpf, and Matthias Scheffler. Analysis of separable potentials. *Physical Review B*, 44(16):8503, 1991.
- [55] Xavier Gonze, Peter Käckell, and Matthias Scheffler. Ghost states for separable, norm-conserving, ab initio pseudopotentials. *Physical Review B*, 41(17):12264, 1990.
- [56] C. Kittel. *Quantum Theory of Solids*. Wiley, 1987.
- [57] DD Koelling and BN Harmon. A technique for relativistic spin-polarised calculations. *Journal of Physics C: Solid State Physics*, 10(16):3107, 1977.
- [58] LA Hemstreet, CY Fong, and JS Nelson. First-principles calculations of spin-orbit splittings in solids using nonlocal separable pseudopotentials. *Physical Review B*, 47(8):4238, 1993.
- [59] Doron Naveh, Leeor Kronik, Murilo L Tiago, and James R Chelikowsky. Real-space pseudopotential method for spin-orbit coupling within density functional theory. *Physical Review B*, 76(15):153407, 2007.

- [60] Gerhard Theurich and Nicola A. Hill. Self-consistent treatment of spin-orbit coupling in solids using relativistic fully separable *ab initio* pseudopotentials. *Phys. Rev. B*, 64:073106, Jul 2001.
- [61] Lucas Fernández-Seivane, Miguel A Oliveira, Stefano Sanvito, and Jaime Ferrer. On-site approximation for spin-orbit coupling in linear combination of atomic orbitals density functional methods. *Journal of Physics: Condensed Matter*, 18(34):7999, 2006.
- [62] James R Chelikowsky, N Troullier, K Wu, and Y Saad. Higher-order finite-difference pseudopotential method: An application to diatomic molecules. *Physical Review B*, 50(16):11355, 1994.
- [63] Xiaodun Jing, N Troullier, David Dean, N Binggeli, James R Chelikowsky, K Wu, and Y Saad. Ab initio molecular-dynamics simulations of Si clusters using the higher-order finite-difference-pseudopotential method. *Physical Review B*, 50(16):12234, 1994.
- [64] Leeor Kronik, Adi Makmal, Murilo L Tiago, MMG Alemany, Manish Jain, Xiangyang Huang, Yousef Saad, and James R Chelikowsky. Parsec—the pseudopotential algorithm for real-space electronic structure calculations: recent advances and novel applications to nano-structures. *physica status solidi (b)*, 243(5):1063–1079, 2006.
- [65] Vincent Michaud-Rioux. Real space DFT by locally optimal block preconditioned conjugate gradient method. Master’s thesis, McGill University, Montreal, Canada, 2012.
- [66] Amir Natan, Ayelet Benjamini, Doron Naveh, Leeor Kronik, Murilo L Tiago, Scott P Beckman, and James R Chelikowsky. Real-space pseudopotential method for first principles calculations of general periodic and partially periodic systems. *Physical Review B*, 78(7):075109, 2008.

-
- [67] Magnus Rudolph Hestenes and Eduard Stiefel. *Methods of conjugate gradients for solving linear systems*, volume 49. NBS, 1952.
- [68] Kikuji Hirose, Tomoya Ono, Yoshitaka Fujimoto, and Shigeru Tsukamoto. *First-principles calculations in real-space formalism: electronic configurations and transport properties of nanostructures*. World Scientific, 2005.
- [69] W Russell Burdick, Yousef Saad, Leeor Kronik, Igor Vasiliev, Manish Jain, and James R Chelikowsky. Parallel implementation of time-dependent density functional theory. *Computer Physics Communications*, 156(1):22–42, 2003.
- [70] Zhaojun Bai, James Demmel, Jack Dongarra, Axel Ruhe, and Henk van der Vorst. *Templates for the solution of algebraic eigenvalue problems: a practical guide*, volume 11. Siam, 2000.
- [71] Beresford N Parlett. *The symmetric eigenvalue problem*, volume 7. SIAM, 1980.
- [72] Youcef Saad. *Numerical methods for large eigenvalue problems*, volume 158. SIAM, 1992.
- [73] Sorin Costiner and Shlomo Taafasan. Adaptive multigrid techniques for large-scale eigenvalue problems: solutions of the schrödinger problem in two and three dimensions. *Physical Review E*, 51(4):3704, 1995.
- [74] Paul Ferdinand Baumeister. *Real-Space Finite-Difference Paw Method for Large-Scale Applications on Massively Parallel Computers*, volume 53. Forschungszentrum Jülich, 2013.
- [75] Yunkai Zhou, James R Chelikowsky, and Yousef Saad. Chebyshev-filtered subspace iteration method free of sparse diagonalization for solving the kohn–sham equation. *Journal of Computational Physics*, 274:770–782, 2014.
- [76] D Hobbs, G Kresse, and J Hafner. Fully unconstrained noncollinear magnetism within the projector augmented-wave method. *Physical Review B*, 62(17):11556, 2000.

-
- [77] Jun-Hyung Cho and Matthias Scheffler. Ab initio pseudopotential study of Fe, Co, and Ni employing the spin-polarized lapw approach. *Physical Review B*, 53(16):10685, 1996.
- [78] Charles Kittel. *Introduction to solid state physics*. Wiley, 1986.
- [79] Tatsuki Oda, Alfredo Pasquarello, and Roberto Car. Fully unconstrained approach to noncollinear magnetism: application to small fe clusters. *Physical review letters*, 80(16):3622, 1998.
- [80] Timothy H Hsieh, Hsin Lin, Junwei Liu, Wenhui Duan, Arun Bansil, and Liang Fu. Topological crystalline insulators in the snte material class. *Nature communications*, 3:982, 2012.