1

An Introduction to Density Functional Theory (DFT) and Derivatives

1.1 The Problem of a N-electron System

The density functional theory (DFT) is first resulted from the work by Hohenberg and Kohn [1], wherein the complicated individual electron orbitals are substituted by the electron density. Namely, the DFT is entirely expressed in terms of the functional of electron density, rather than the many-electron wave functions. In this case, DFT significantly reduces the calculations of the ground state properties of materials. That is why DFT is useful for calculating electronic structures, especially with many electrons. As the foundation of DFT, two theorems are proposed by Hohenberg and Kohn [1]. The first theorem presents that the ground state energy is a functional of electron density. The second theorem shows that the ground state energy can be achieved by minimizing system energy on the basis of electron density.

It should be noted that, although Hohenberg and Kohn point out there are relations between properties and electron density functional, they do not present the exact relationship. But fortunately, soon after the work of Hohenberg and Kohn, Kohn and Sham simplified the many-electron problems into a model of individual electrons in an effective potential [2]. Such a potential contains the external potential and exchange-correlation interactions. For exchange-correlation potential, it is a challenge to describe it rigorously.

The simplest approximation for treating the exchange-correlation interaction is the local density approximation (LDA) [3], wherein the exchange and correlation energies are obtained by the uniform electron gas model and fitting to the uniform electron gas, respectively. LDA can provide a realistic description of the atomic structure, elastic, and vibrational properties for a wide range of systems. Yet, because LDA treats the energy of the true density using the energy of a local constant density, it cannot describe the situations where the density features rapid changes such as in molecules [4, 5]. To address this problem, the generalized gradient approximation (GGA) is proposed [6–8], which depends on both the local density and the spatial variation of the density. And in principle, GGA is as simple to use as LDA. Currently, in the vast majority of DFT calculations for solids, these two approximations are adopted.

Calculations and Simulations of Low-Dimensional Materials: Tailoring Properties for Applications,
First Edition. Ying Dai, Wei Wei, Yandong Ma, and Chengwang Niu.

By considering the Born–Oppenheimer and non-relativistic approximations, the effective Hamiltonian of a N-electron system in the position representation can be given by,

$$H(r_1, r_2, \dots r_N) = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i \hat{\nu}_{\text{ne}}(r_i) + \frac{1}{2}\sum_i \sum_{j \neq i} \frac{1}{|r_i - r_j|} \tag{1.1}$$

The first term is kinetic energy operator. The second term is an external potential operator. In systems of interest to us, the external potential is simply the Coulomb interaction of electrons with atomic nuclei:

$$\hat{\nu}_{\text{ne}}(r_i) = -\sum_\alpha \frac{Z_\alpha}{|r_i - R_\alpha|} \tag{1.2}$$

where the r_i is the coordinate of electron i and the charge on the nucleus at R_α is Z_α. The third term of Eq. (1.1) is the electron-electron operator. The electronic state can be obtained by the Schrödinger equation:

$$H(r_1, r_2, r_N)\Psi(r_1, r_2, r_N) = E\Psi(r_1, r_2, r_N) \tag{1.3}$$

Here, $\Psi(r_1, r_2, r_N)$ is a wave function in terms of space-spin coordinates. Apparently, the wave function is antisymmetric under exchanging the coordinates. Under Dirac notation, the Eq. (1.1) can be expressed in representation-independent formalism:

$$H|\Psi\rangle = E|\Psi\rangle \tag{1.4}$$

In principle, the ground state energy E_0 of the N-electron system can be found based on the variational theorem, which is obtained by the minimization:

$$E_0 = \min_\Psi \langle\Psi|\hat{H}|\Psi\rangle \tag{1.5}$$

Here, the search is over all antisymmetric wave functions Ψ. In this regard, better approximations for Ψ can readily result in the ground state energy E_0 of the N-electron system, but the computational cost would be very high. Therefore, the direct solution is not feasible. To address this issue, DFT is developed, which is based on a reformulation of the variational theorem in terms of electron density.

We know that $|\Psi|^2 = \Psi^*\Psi$ represents the probability density of measuring the first electron at r_1, the second electron at r_2, … and the Nth electron at r_N. By integrating $|\Psi|^2$ over the first $N - 1$ electrons, the probability density of the Nth electron at r_N is determined. Then the probability electron density that defines any of the N electrons at the position r is given by:

$$\rho(r) = N \int \dots \int \Psi^*(r_1, r_2, r_N)\Psi(r_1, r_2, \dots r_N) d_{r_2} \dots d_{r_N} \tag{1.6}$$

And the electron density is normalized to the electron number:

$$\int \rho(r)dr = N \tag{1.7}$$

The energy of the system is expressed as:

$$E = \langle\Psi|\hat{H}|\Psi\rangle = \langle\Psi|\hat{T}|\Psi\rangle + \langle\Psi|\hat{V}_{\text{ext}}|\Psi\rangle + \langle\Psi|\hat{V}_{\text{ee}}|\Psi\rangle = T + V_{\text{ext}} + V_{\text{ee}} \tag{1.8}$$

Here,

$$T = \langle \Psi | \hat{T} | \Psi \rangle = \sum_i \int \Psi^*(r_1, r_2, r_N) \left(-\frac{1}{2} \nabla_i^2 \right) \Psi(r_1, r_2, r_N) d_{r_1} d_{r_2} \dots d_{r_N} \tag{1.9}$$

$$\begin{aligned} V_{\text{ext}} &= \langle \Psi | \hat{V}_{\text{ext}} | \Psi \rangle = \sum_i \int \Psi^*(r_1, r_2, r_N) \hat{v}_{\text{ne}}(r_i) \Psi(r_1, r_2, \dots r_N) d_{r_1} d_{r_2} \dots d_{r_N} \\ &= \int v_{\text{ne}}(r) \rho(r) dr = V_{\text{ext}}[\rho] \end{aligned} \tag{1.10}$$

1.2 The Thomas–Fermi Theory for Electron Density

Before discussing the Hohenberg–Kohn theorems, we first introduce the Thomas–Fermi theory. The Thomas–Fermi theory is important as it gives the relation between external potential and the density distribution for interacting electrons moving in an external potential:

$$\rho(r) = \gamma(\mu - v_{\text{eff}}(r))^{3/2} \tag{1.11}$$

$$v_{\text{eff}}(r) \equiv v_{\text{ne}}(r) + \int \frac{\rho(r')}{|r - r'|} dr' \tag{1.12}$$

Here,

$$\gamma = \frac{1}{3\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \tag{1.13}$$

and μ is the r independent chemical potential. The second term is Eq. (1.12) is the classical electrostatic potential raised by the density $\rho(r)$. Based on the Thomas–Fermi theory, Hohenberg and Kohn build up the connection between electron density and the Schrödinger equation. And in the following, we will introduce the two Hohenberg–Kohn theorems, which lie at the heart of DFT.

1.3 The First Hohenberg–Kohn Theorem

By replacing the external potential $v_{\text{ne}}(r)$ with an arbitrary external local potential $v(r)$, the corresponding ground state wave function Ψ can be found by solving the Schrödinger equation. Based on the obtained wave function, the ground state density $\rho(r)$ can be computed. And obviously, two different local potentials would give two different wave functions and thus two different electron densities. This gives the map:

$$v(r) \rightarrow \rho(r) \tag{1.14}$$

Based on the Thomas–Fermi theory, Hohenberg and Kohn demonstrated that the preceding mapping can be inverted, namely, *the ground state electron density $\rho(r)$ of a bound system of interacting electrons in some external potential $v(r)$ determines the potential uniquely*:

$$\rho(r) \rightarrow v(r) \tag{1.15}$$

This is known as the first Hohenberg–Kohn theorem.

To demonstrate this theorem, we consider two different local potentials $v_1(r)$ and $v_2(r)$, which differ by more than the constant. These two potentials yield two different ground state wave functions Ψ and Ψ', respectively. And apparently, these two ground state wave functions are different. Assume $v_1(r)$ and $v_2(r)$ correspond to the same ground state wave function, then

$$\hat{H}|\Psi\rangle = E_0|\Psi\rangle \tag{1.16}$$

$$\hat{H}'|\Psi\rangle = E_0'|\Psi\rangle \tag{1.17}$$

By subtracting Eq. (1.17) from Eq. (1.16), we can obtain:

$$(\hat{V}_1 - \hat{V}_2)|\Psi\rangle = \left(E_0 - E_0'\right)|\Psi\rangle \tag{1.18}$$

which can be expressed in position representation,

$$\sum_i [\hat{v}_1(r_1) - \hat{v}_2(r_i)]\Psi(r_1, r_2, r_N) = \left(E_0 - E_0'\right)\Psi(r_1, r_2, r_N) \tag{1.19}$$

This suggests that

$$v_1(r) - v_2(r) = \text{const} \tag{1.20}$$

thus in contradiction with the assumption that $v_1(r)$ and $v_2(r)$ differ by more than a constant. Accordingly, two different local potentials that differ by more than the constant cannot share the same ground state wave function, which demonstrate the map:

$$v(r) \to \Psi \tag{1.21}$$

Then, we demonstrate the map:

$$\Psi \to \rho(r) \tag{1.22}$$

Let Ψ and Ψ' be the ground state wave functions corresponding to $v_1(r)$ and $v_2(r)$, respectively. Assuming that Ψ and Ψ' exhibit the same ground state electron density $\rho(r)$, then the variational theorem gives the ground state energy as:

$$\begin{aligned} E_0 &= \langle\Psi|\hat{H}|\Psi\rangle < \langle\Psi'|\hat{H}|\Psi'\rangle = \langle\Psi'|\hat{H}' + \hat{V}_1 - \hat{V}_2|\Psi'\rangle \\ &= E_0' + \int [v_1(r) - v_2(r)]\rho(r)dr \end{aligned} \tag{1.23}$$

$$\begin{aligned} E_0' &= \langle\Psi'|\hat{H}'|\Psi'\rangle < \langle\Psi|\hat{H}'|\Psi\rangle = \langle\Psi|\hat{H} - \hat{V}_1 + \hat{V}_2|\Psi'\rangle \\ &= E_0 + \int [v_2(r) - v_1(r)]\rho(r)dr \end{aligned} \tag{1.24}$$

By subtracting Eq. (1.23) from Eq. (1.24), we can obtain:

$$E_0' - E_0 < E_0 - E_0' \tag{1.25}$$

This makes no sense. This finally leads to the conclusion that there cannot exist two local potentials differing by more than an additive constant that has the same ground state density.

1.4 The Second Hohenberg–Kohn Theorem

According to the first Hohenberg–Kohn theorem, the ground state density $\rho(r)$ determines the local potential $v(r)$, and in turn determines the Hamiltonian. Therefore, for a given ground state density $\rho_0(r)$ that is generated by a local potential, it is possible to compute the corresponding ground state wave function Ψ_0. That is to say, Ψ_0 is also a unique functional of $\rho_0(r)$:

$$\Psi_0 = \Psi_0[\rho_0] \tag{1.26}$$

According to Eq. (1.26), the ground state energy E_0 is also a functional of $\rho_0(r)$:

$$E_0 = E_0[\rho_0] \tag{1.27}$$

Hohenberg and Kohn define the universal density functional:

$$F[\rho] = \langle \Psi[\rho] | \hat{T} + \hat{V}_{\text{ee}} | \Psi[\rho] \rangle \tag{1.28}$$

Here, $\Psi[\rho]$ is any ground state wave function corresponding to the ground state density $\rho(r)$. By combining Eq. (1.10), the total energy functional can be defined as:

$$E[\rho] = F[\rho] + \int v_{\text{ne}}(r)\rho(r)dr \tag{1.29}$$

From the Ritz principle, we have:

$$E_0 = \min_{\rho} E[\rho] = \min_{\rho} \left\{ F[\rho] + \int v_{\text{ne}}(r)\rho(r)dr \right\} \tag{1.30}$$

This is known as the second Hohenberg–Kohn theorem.

1.5 The Kohn–Sham Equations

For a system with noninteracting electrons, the effective Hamiltonian n can be given by:

$$\hat{H}_{\text{s}} = \hat{T} + \hat{V}_{\text{s}} \tag{1.31}$$

The corresponding Schrödinger equation is:

$$\left(-\frac{1}{2}\nabla^2 + \hat{v}_s(r)\right)\psi_i(r) = \varepsilon_i \psi_i \tag{1.32}$$

Then the density is given by:

$$\rho_{\text{s}}(r) = \sum_i |\psi_i(r)|^2 \tag{1.33}$$

Here, the single particle orbital $\psi_i(r)$ is constructed based on the effective potential $v_{\text{s}}(r)$.

The total energy can be expressed as:

$$E_{\text{s}}[\rho] = T_{\text{s}}[\rho] + V_{\text{s}}[\rho] \tag{1.34}$$

The first term is the kinetic energy of the noninteracting electrons, which is given by:

$$T_s[\rho] = \langle \Psi[\rho] | \hat{T} | \Psi[\rho] \rangle = \sum_i \left\langle \psi_i(r) \mid -\frac{1}{2}\nabla^2 \mid \psi_i(r) \right\rangle \tag{1.35}$$

The second term is the effective potential, which is given by:

$$V_s[\rho] = \langle \Psi[\rho] | \hat{V}_s | \Psi[\rho] \rangle = \int \hat{v}_s(r)\rho(r)dr \tag{1.36}$$

Accordingly, the total energy can be given by:

$$E_s[\rho] = \sum_i \varepsilon_i \tag{1.37}$$

By combining Eqs. (1.36) and (1.37), the kinetic energy can be expressed as:

$$T_s[\rho] = \sum_i \left\langle \psi_i(r) \middle| -\frac{1}{2}\nabla^2 \middle| \psi_i(r) \right\rangle = \sum_i \varepsilon_i - \int \hat{v}_s(r)\rho(r)dr \tag{1.38}$$

Using the method of Lagrange multipliers, we can obtain the following equation:

$$\mu = \frac{\delta E_s[\rho]}{\delta \rho(r)} = \frac{\delta T_s[\rho]}{\delta \rho(r)} + \frac{\delta V_s[\rho]}{\delta \rho(r)} = \frac{\delta T_s[\rho]}{\delta \rho(r)} + v_s(r) \tag{1.39}$$

Solution of Eq. (1.39) yields the density $\rho_s(r)$.

The classical electrostatic interaction energy is given by:

$$E_H[\rho] = \frac{1}{2} \int V_H(r)\rho(r)dr \tag{1.40}$$

And

$$V_H(r) = \int \frac{\rho(r')}{|r - r'|} dr' \tag{1.41}$$

Then the following equation is obtained:

$$\begin{aligned} E[\rho] = T[\rho] + V_{ee}[\rho] + V[\rho] &= T_s[\rho] + E_H[\rho] \\ &+ (T[\rho] - T_s[\rho] + V_{ee}[\rho] - E_H[\rho]) + V[\rho] \end{aligned} \tag{1.42}$$

The third term is exchange and correlation energy functional:

$$E_{xc} = T[\rho] - T_s[\rho] + V_{ee}[\rho] - E_H[\rho] \tag{1.43}$$

The exchange-correlation potential is defined as:

$$V_{xc} = \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \tag{1.44}$$

Using the method of Lagrange multipliers, we can obtain the following equation:

$$\mu = \frac{\delta E[\rho]}{\delta \rho(r)} = \frac{\delta T_s[\rho]}{\delta \rho(r)} + V_H(r) + V_{xc}(r) + v_{ne}(r) \tag{1.45}$$

And solution to Eq. (1.45) is $\rho(r)$.

Therefore, given the relation:

$$v_s(r) \equiv V_H(r) + V_{xc}(r) + v_{ne}(r) \tag{1.46}$$

We have:

$$\rho_s(r) \equiv \rho(r) \tag{1.47}$$

We then arrive at the Kohn–Sham equation:

$$\left(-\frac{1}{2}\nabla^2 + \hat{V}_H(r) + \hat{V}_{xc}(r) + \hat{v}_{ne}(r)\right)\psi_i(r) = \varepsilon_i\psi_i \tag{1.48}$$

Solving this equation gives the orbital and then the density of the original interacting system.

The exchange and correlation functional can be written as:

$$E_{xc}[\rho] = \int \varepsilon_{xc}(r'; \rho)\rho(r')dr' \tag{1.49}$$

Here, $\varepsilon_{xc}(r'; \rho)$ is the exchange-correlation energy density. And the exchange-potential is defined as:

$$V_{xc}(r; \rho) = \frac{\delta E_{xc}[\rho]}{\delta\rho(r)} = \varepsilon_{xc}(r; \rho) + \int \frac{\delta\varepsilon_{xc}(r'; \rho)}{\delta\rho(r)}\rho(r')dr \tag{1.50}$$

Then the total energy can be given by:

$$E[\rho] = \sum_i \varepsilon_i - E_H[\rho] + E_{xc}[\rho] - \int V_{xc}(r; \rho)\rho(r)dr \tag{1.51}$$

From preceding equation, we can see that, except for the exchange-correlation functional, all the aforementioned expressions are exact. In practice, we have to use approximations for exchange-correlation potential, as the exact form is unknown.

1.6 The Local Density Approximation (LDA)

LDA is one of the most widely used and simplest approximations for E_{xc}. In LDA, the exchange-correlation functional is approximated as:

$$E_{xc}^{LD}[\rho] = \int \rho(r)\varepsilon_{xc}^{unif}[\rho(r)dr] \tag{1.52}$$

Here, $\varepsilon_{xc}^{unif}[\rho(r)]$ is the exchange-correlation energy per electron in homogeneous electron gas at density ρ. LDA works well for homogeneous electron gas, and thus is valid for systems where electron density does not change rapidly.

The exchange-correlation energy density can be broken down into two parts:

$$\varepsilon_{xc}^{unif}[\rho(r)] = \varepsilon_x^{unif}[\rho(r)] + \varepsilon_c^{unif}[\rho(r)] \tag{1.53}$$

The first term is the exchange term, which is given by:

$$\varepsilon_x^{unif}[\rho(r)] = \text{Const.} \times \rho^{1/3}(r) \tag{1.54}$$

While for the second term of Eq. (1.53), it is the correlation density, which does not have an analytic formula. However, the correlation energies can be obtained numerically from quantum Monte Carlo (QMC) calculations by Ceperley and Alder [9].

And for spin-polarized systems, the spin-up and spin-down densities are taken as two independent densities in the exchange-correlation energy. And in this case, the Eq. (1.52) can be expressed as:

$$E_{\mathrm{xc}}^{\mathrm{LSD}}[\rho] = \int \rho(r)\varepsilon_{\mathrm{xc}}^{\mathrm{unif}}[\rho_{\downarrow}(r), \rho_{\downarrow}(r)]dr \tag{1.55}$$

For calculating the electronic structure, LDA approach is estimated to be successful. However, for some systems, it does not work. As a result, many efforts are devoted to improve it. One of them is to include the gradient of the density in the exchange correlation functional, as we will show next.

1.7 The Generalized Gradient Approximation (GGA)

To introduce the gradient of the density in the exchange correlation functional, the gradient expansion approximation (GEA) is first proposed. Starting from the uniform electron gas, a slowly varying external potential $v(r)$ is introduced. And then the exchange-correlation energy is expanded in terms of the gradients of the density:

$$E_{\mathrm{xc}}^{\mathrm{GEA}}[\rho] = E_{\mathrm{xc}}^{\mathrm{LD}}[\rho] + \int C_{\mathrm{xc}}(\rho(r))\rho(r)^{4/3}\left(\frac{\nabla\rho(r)}{\rho(r)^{4/3}}\right)^{2} dr \tag{1.56}$$

Here, $C_{\mathrm{xc}}(\rho)$ is the sum of the exchange and correlation coefficients of the gradient expansion. As the reduced density gradient is small, GEA approach should be superior to LDA approach. However, because the reduced density gradient can be large in some region of space for real systems, GEA is shown to be worse than LDA.

To overcome the shortcomings of GEA, the GGA is developed. In GGA, the exchange-correlation functional is approximated as:

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[\rho] = \int f(\rho(r), \nabla\rho(r))dr \tag{1.57}$$

Here, f is some function. Many GGA functionals have been proposed, including B88 [10], Lee-Yang-Parr (LYP) [11], PW91 [12], and Perdew–Burke–Ernzerhof (PBE) [13] exchange-correlation functionals.

1.8 The LDA+U Method

While LDA and GGA are estimated to be able to deal with the many systems and phenomena, they do not work well for the systems with rare-earth and late-transition metal elements. This is because the effective single-particle methods are applicable for highly delocalized band states but not for strongly localized states. For the d and f electrons of rare-earth and late-transition metal elements, they essentially retain their atomic character in solids. As a consequence, standard DFT functionals such as LDA, local spin density approximation (LSDA), and GGA itinerant d states and metallic ground state for many transition metal oxides, for which semiconducting behavior is demonstrated experimentally. For improving these issues, LDA+U method is developed [14–16]. Here, if not specified elsewhere, +U indicates a Hubbard, and LDA indicates the standard DFT functionals, i.e. LDA, LSDA, and

GGA. The idea of LDA+U method is on the basis that the strongly correlated electronic states (i.e. d and f sates) are treated by the Hubbard model, and the rest of the valence electrons are described by the standard DFT functionals. Therefore, the total energy within LDA+U method can be given by:

$$E_{\mathrm{LDA+U}}[\rho] = E_{\mathrm{LDA}}[\rho] + E_{\mathrm{Hub}}\left[\left\{\rho^{I\sigma}_{mm'}\right\}\right] - E_{\mathrm{dc}}[\{\rho^{I\sigma}\}] \tag{1.58}$$

Here, the first term is the standard DFT total energy functional being corrected, the second term represents the Hubbard Hamiltonian to model correlated states, and the third is the double-counting term. The LDA+U method can well describe the electronic properties of the Mott insulators and increase the band gaps in the Kohn–Sham spectrum.

1.9 The Heyd–Scuseria–Ernzerhof Density Functional

In standard DFT, the Fock exchange energy is computed based on a local energy density and its derivatives. However, the exact form for the Fock exchange energy is known as nonlocal from the Hartree–Fock theory. To improve the accuracy, the PBE exchange energy should be mixed with a fraction of the exact nonlocal Fock exchange energy, giving rise to the hybrid functionals, such as the Heyd–Scuseria–Ernzerhof (HSE) functional [17]. The HSE exchange-correlation energy is given by:

$$E_{\mathrm{xc}}^{\mathrm{HSE}} = \alpha E_{\mathrm{x}}^{\mathrm{HF,short}}[\mu] + (1-\alpha)E_{\mathrm{x}}^{\mathrm{PBE,short}}[\mu] + E_{\mathrm{x}}^{\mathrm{PBE,long}}[\mu] + E_{\mathrm{c}}^{\mathrm{PBE}}[\mu] \tag{1.59}$$

Here, $E_{\mathrm{x}}^{\mathrm{HF}}$ and $E_{\mathrm{x}}^{\mathrm{PBE}}$, respectively, represent the exact Fock exchange energy and the PBE exchange energy. $E_{\mathrm{c}}^{\mathrm{PBE}}$ is the PBE correlation energy. μ is the range-separation parameter. α is the mixing parameter. From Eq. (1.59), it can be seen that the HSE functional is split into short- and long-range terms. In this case, it can improve the accuracy, while avoiding the computational cost.

1.9.1 Introduction to Tight-Binding Approximation

Consider in a single atom there are multiple atomic orbitals $\varphi_{\mathrm{m}}(\mathbf{r})$ with m being the orbital indices. Here, $\varphi_{\mathrm{m}}(\mathbf{r})$ must be eigenfunctions of the Hamiltonian of that single atom H_{atom}. When we place it in a crystal with plenty of atoms, the wave function of different atoms overlap each other to form a different wave function. Due to that $\varphi_{\mathrm{m}}(\mathbf{r})$ is not a real eigenfunction for a Hamiltonian of crystal and we need to find out what the true eigenfunctions are. If the overlap of one atom on another is small enough, we can still assume that electrons are tightly bound to the corresponding atoms, which is exactly the reason why we call it as tight-binding approximation. The approximate Hamiltonian is $H(\mathbf{r}) = H_{\mathrm{atom}}(\mathbf{r}) + \Delta U(\mathbf{r})$, whose Bloch wave function can be taken as a combination of all the isolated orbitals: [18, 19]

$$\psi_{\mathrm{m}}(\mathbf{r}) = \sum_{\mathbf{R_n}} b_{\mathrm{m}}(\mathbf{R_n})\varphi_{\mathrm{m}}(\mathbf{r} - \mathbf{R_n}),$$

where $\mathbf{R_n}$ denotes all lattice points and $b_{\mathrm{m}}(\mathbf{R_n})$ is just a coefficient number for orbital m. In the presence of translation symmetry, coefficient numbers can be replaced by a Bloch form, which gives:

$$\psi_{\mathrm{m}}(\mathbf{r}) \approx \frac{1}{\sqrt{N}} \sum_{\mathbf{R_n}} e^{i\mathbf{k}\cdot\mathbf{R_n}} \varphi_{\mathrm{m}}(\mathbf{r}-\mathbf{R_n}).$$

1.9.2 Matrix Elements of Tight-Binding Hamiltonian

To get Hamiltonian in the momentum space, we shall do a basis transformation such as: [20]

$$\psi_{\mathbf{k}}(\mathbf{r})|H|\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{N}\sum_{i,j} e^{-\mathrm{i}\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)} \langle \varphi(\mathbf{r}-\mathbf{R}_i)|H|\varphi(\mathbf{r}-\mathbf{R}_j)\rangle.$$

When $\mathbf{R}_i$ is equal to $\mathbf{R}_j$, we will find the onsite energy represented for the atomic energy shift due to the overlap of other atoms, which can be given as:

$$\beta_i = \langle \varphi(\mathbf{r}-\mathbf{R}_i)|H|\varphi(\mathbf{r}-\mathbf{R}_i)\rangle.$$

If $\mathbf{R}_i$ is not equal to $\mathbf{R}_j$, hopping energy between different lattice sites can be defined as:

$$t_{ij} = -\langle \varphi(\mathbf{r}-\mathbf{R}_i)|H|\varphi(\mathbf{r}-\mathbf{R}_j)\rangle.$$

1.9.3 Matrix Elements with the Help of Wannier Function

Usually, Bloch wave functions are not orthogonal, which may result in some problems. To resolve that we should define the orthogonal Wannier function as:

$$w(\mathbf{r}-\mathbf{R}_i) = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i}\psi_{\mathbf{k}}(\mathbf{r}).$$

By using a bra-ket notation, the Hamiltonian in the real space takes the form of: [21]

$$H = \beta_i \sum_{\mathbf{R}_i} |\mathbf{R}_i\rangle\langle\mathbf{R}_i| - t_{ij}\sum_{\mathrm{i}\neq\mathrm{j}} |\mathbf{R}_i\rangle\langle\mathbf{R}_j|,$$

where β_i and t_{ij} denote the onsite energy and hopping energy, respectively. To get the energy of Hamiltonian, we shall do a basis transformation into the momentum space, which is implemented by a Fourier transform:

$$H = \sum_i \beta_i \sum_{\mathbf{k}} |\mathbf{k}\rangle\langle\mathbf{k}| - t_{ij}\sum_{i\neq j}\frac{1}{N}\sum_{\mathbf{k}} e^{-\mathrm{i}\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)}|\mathbf{k}\rangle\langle\mathbf{k}|.$$

1.9.4 Example for a Graphene Model

By using graphene lattice [22] as an example, we give a detailed description about how to use tight-binding method. Here, we consider two atoms with s orbital in a single unit cell of unit lattice constant, whose lattice vector can be written as:

$$a_1 = \frac{1}{2}\left(3, \sqrt{3}\right), a_2 = \frac{1}{2}\left(3, -\sqrt{3}\right).$$

Each atom is connected with three nearest-neighbor atoms with a distance of:

$$\delta_1 = \frac{1}{2}\left(1, \sqrt{3}\right), \delta_2 = \frac{1}{2}\left(1, -\sqrt{3}\right), \delta_3 = (-1, 0).$$

In the momentum space, we consider the onsite energy for two s orbitals is ϵ_1. The nearest-neighbor hopping strength is t, which leaves a Hamiltonian matrix as follows:

$$H = \begin{vmatrix} \epsilon_1 & -t\sum_j e^{ik\delta_j} \\ -t\sum_j e^{-ik\delta_j} & \epsilon_1 \end{vmatrix}.$$

We can get the energy dispersion with an implementation of diagonalization, which results in:

$$E(k_x, k_y) = \epsilon_1 \pm t\sqrt{1 + 4\cos\left(\frac{\sqrt{3}k_y}{2}\right)\cos\left(\frac{3k_x}{2}\right) + 4\cos^2\left(\frac{\sqrt{3}k_y}{2}\right)}.$$

1.10 Introduction to $k \cdot p$ Perturbation Theory

A single electron in a periodic potential $V(\mathbf{r})$ obeys a Schrödinger equation, and such a form can be written as [23, 24]:

$$\left[\frac{\mathbf{p}^2}{2m} + V(\mathbf{r})\right]\psi_{(\mathbf{k},\mathbf{r})} = E\psi_{(\mathbf{k},\mathbf{r})},$$

where the eigenvalues and eigenfunctions can be written as $E_n(\mathbf{k})$ and $\psi_n(\mathbf{k}, \mathbf{r}) = \mathrm{e}^{\mathbf{ik}\cdot\mathbf{r}}u_n(\mathbf{k}, \mathbf{r})$. The periodic part that is called as a Bloch function satisfies such an equation as:

$$H_0(\mathbf{k})u_n(\mathbf{k}, \mathbf{r}) = E_n(\mathbf{k})u_n(\mathbf{k}, \mathbf{r}),$$

where the Hamiltonian located at momentum point of $\mathbf{k}_0$ can be given as: $H_0(\mathbf{k}_0) = \frac{\mathbf{p}^2}{2m} + \frac{\hbar}{m}\mathbf{k}_0 \cdot \mathbf{p} + \frac{\hbar^2\mathbf{k}_0^2}{2m} + V(\mathbf{r})$. If the eigenvalues and eigenfunctions are assumed to be solved for point as $\mathbf{k}_0$, we can get the solutions of nearby points such as $\mathbf{k} = \mathbf{k}_0 + \delta\mathbf{k}$ through the following equation:

$$\left[H_0(\mathbf{k}_0) + \frac{\hbar}{m}\delta\mathbf{k} \cdot \mathbf{p}\right]u_n(\mathbf{k}, \mathbf{r}) = \left[E_n(\mathbf{k}_0) + \frac{\hbar^2}{2m}\left(\mathbf{k}_0^2 - \mathbf{k}^2\right)\right]u_n(\mathbf{k}, \mathbf{r}).$$

Then, the perturbation Hamiltonian gains a form of $H'(\mathbf{k}) = \frac{\hbar}{m}\delta\mathbf{k} \cdot \mathbf{p}$. With the help of perturbation theory, we can solve the energy of nearby points under two different situations, which is illustrated in the following chapter.

1.10.1 Solution for Non-degenerate Bands

If the bands are not degenerate for $\mathbf{k}_0$, the perturbated eigenvalues of the points $\mathbf{k} = \mathbf{k}_0 + \delta\mathbf{k}$ can be given by [25, 26]:

$$\begin{aligned} E_n(k) = {} & E_n(k_0) + \frac{2}{2m}\left(k_0^2 - k^2\right) + \frac{\hbar}{m}\delta k\langle u_{nk0}(r) \mid p \mid u_{nk0}(r)\rangle \\ & + \frac{\hbar 2}{m^2}\sum_{n'\neq n}\frac{\langle u_{nk_0}(r) \mid p \mid u_{n'k_0}(r)\rangle\langle u_{n'k_0}(r) \mid p \mid u_{nk_0}(r)\rangle\delta k^2}{E_n(k_0) - E_{n'}(k_0)}, \end{aligned}$$

where the first term and second term of right-hand side are zero order approximation, and the third term and fourth term serve as the first and second order approximation. Here, we give an explicit example to describe how to get a Hamiltonian based on the following equation. We assume the eigenvalues and eigenfunctions of $\mathbf{k}_0 = 0$ are known by first-principles calculations or experiment. Consider a cubic lattice with a point group of O_h and two bands $|u_{1\mathbf{k}}(\mathbf{r})\rangle$ and $|u_{2\mathbf{k}}(\mathbf{r})\rangle$ that transform as Γ_1^+ and Γ_1^- representation, and the symmetry representation of H' is the vector representation as Γ_{15}^-.

For the first order approximation, the direct product $\Gamma_1^+ \otimes \Gamma_{15}^- \otimes \Gamma_1^+ = \Gamma_{15}^-$ changes the representation Γ_1^+ into Γ_{15}^-, resulting in a vanishing matrix. While for the second order approximation, the direct product $\Gamma_{15}^- \otimes \Gamma_1^+ = \Gamma_{15}^-$ limits the $|u_{n'\mathbf{k}_0}(\mathbf{r})\rangle$ to the Γ_{15}^- representation, which is just the band representation of antibonding p bands. There are three basis x, y, and z for Γ_{15}^- representation; only three terms can exist as a cross term of $\langle \Gamma_1^+ | p_x | x \rangle$, $\langle \Gamma_1^+ | p_y | y \rangle$, and $\langle \Gamma_1^+ | p_z | z \rangle$ according to the selection rules. Finally, we can get the total eigenvalues as:

$$E_1(\mathbf{k}) = E_1(0) + \frac{\hbar^2\mathbf{k}^2}{2m} + \frac{\hbar^2\mathbf{k}^2}{m^2} \sum_{n' \neq 1} \frac{|\langle u_{1\mathbf{k}_0}(\mathbf{r}) | p_x | x \rangle|^2}{E_1(0) - E_{n'}(0)}.$$

1.10.2 Solution for Degenerate Bands

If the bands are degenerate for $\mathbf{k}_0$, the eigenfunction must be a linear combination of degenerated bands. Here, we assume band i is degenerate with band j, and the first order perturbation equation can be written as: [27]

$$\begin{vmatrix} \langle i | H_0 + H' | i \rangle - \varepsilon & \langle i | H_0 + H' | j \rangle \\ \langle j | H_0 + H' | i \rangle & \langle j | H_0 + H' | j \rangle - \varepsilon \end{vmatrix} = 0.$$

The four terms on left-hand side can change the forms as $\langle i | H_0 + H' | i \rangle = E_i^0 - \varepsilon$, $\langle j | H_0 + H' | j \rangle = E_j^0 - \varepsilon$, $\langle i | H_0 + H' | j \rangle = (\hbar/m)\mathbf{k} \cdot \langle i | \mathbf{p} | j \rangle$, and $\langle j | H_0 + H' | i \rangle = (\hbar/m)\mathbf{k} \cdot \langle j | \mathbf{p} | i \rangle$. The solution yields:

$$\varepsilon(\vec{\kappa}) = \frac{E_i^0 + E_j^0}{2} \pm \frac{1}{2}\sqrt{\left(E_i^0 - E_j^0\right)^2 + \frac{4\hbar^2}{m^2}\mathbf{k} \cdot \mathbf{p}_{ij}^{\leftrightarrow} \cdot \mathbf{k}},$$

where $\mathbf{p}_{ij}^{\leftrightarrow}$ is a third order tensor. For a cubic lattice, such a tensor has a form of:

$$\mathbf{p}_{ij}^{\leftrightarrow} = \begin{pmatrix} p_{ij}^2 & 0 & 0 \\ 0 & p_{ij}^2 & 0 \\ 0 & 0 & p_{ij}^2 \end{pmatrix}.$$

1.10.3 Explicit Hamiltonian of $k \cdot p$ Perturbation Theory

Generally, any 4×4 Hamiltonian can be expanded with 16 Dirac matrices as:

$$H = \epsilon(\mathbf{k})I + \sum_i d_i(\mathbf{k})\Gamma_i + \sum_{ij} d_{ij}(\mathbf{k})\Gamma_{ij},$$

where I is the identify matrices, and the five Dirac Γ_i matrices can be defined as $\Gamma_1 = \sigma_1 \otimes \tau_1, \Gamma_2 = \sigma_2 \otimes \tau_1, \Gamma_3 = \sigma_3 \otimes \tau_1, \Gamma_4 = 1 \otimes \tau_2, \Gamma_5 = 1 \otimes \tau_4$, and $\Gamma_{ij} = [\Gamma_i, \Gamma_j]/2i$ with σ_i and τ_i are two sets of Pauli matrices.

First, we should write the symmetry matrix in the basis we choose. Taken Bi_2Se_3 as an example, [28, 29] we choose four states as: $\left|P1^+_-, \frac{1}{2}\right\rangle, \left|P2^-_+, \frac{1}{2}\right\rangle, \left|P1^+_-, -\frac{1}{2}\right\rangle$, and $\left|P2^-_+, -\frac{1}{2}\right\rangle$. According to the transformation formula, four symmetry matrices D(R) can be constructed as:

1. Time-reversal symmetry: $T = i\sigma_y K \otimes 1$.
2. Threefold rotation symmetry along z axis: $R_3 = e^{i(\sigma_3 \otimes 1)\pi/3}$.
3. Twofold rotation symmetry along x axis: $R_2 = i\sigma_1 \otimes \tau_3$.
4. Inversion symmetry: $P = 1 \otimes \tau_3$.

Second, we apply the symmetry matrices into 16 Dirac matrices to get the representation of matrices, which is implemented by $D(R)\Gamma_i D(R)^{-1}$. In the presence of time-reversal symmetry, we only need to take into consideration one identity matrix and five Dirac Γ_i matrices.

Third, we combined the Dirac matrices with corresponding polynomials of the momentum k who share get the Hamiltonian. It can be written as:

$$H = \epsilon_{\mathbf{k}} + M(\mathbf{k})\Gamma_5 + B(k_z)\Gamma_4 k_z + A(k_\parallel)(\Gamma_1 k_y - \Gamma_2 k_x),$$

where $\epsilon_{\mathrm{k}} = C_0 + C_1 k_z^2 + C_2 k_\parallel^2$, $M(\mathbf{k}) = M_0 + M_1 k_z^2 + M_2 k_\parallel^2, B(k_z) = B_0 + B_2 k_z^2$, and $k_\parallel^2 = k_x^2 + k_y^2$.

References

1 Hohenberg, P. and Kohn, W. (1964). Inhomogeneous electron gas. *Phys. Rev.* 136: B864–B871.

2 Kohn, W. and Sham, L.J. (1965). Self-consistent equations including exchange and correlation effects. *Phys. Rev.* 140: A1133–A1138.

3 Jones, R.O. and Gunnarsson, O. (1989). The density functional formalism, its applications and prospects. *Rev. Mod. Phys.* 61: 689–746.

4 Grossman, J.C., Mitas, L., and Raghavachari, K. (1995). Structure and stability of molecular carbon: importance of electron correlation. *Phys. Rev. Lett.* 75: 3870–3873.

5 Zupan, A., Blaha, P., Schwarz, K., and Perdew, J.P. (1998). Pressure-induced phase transitions in solid Si, SiO2, and Fe: Performance of local-spin-density and generalized-gradient-approximation density functionals. *Phys. Rev. B* 58: 11266–11272.

6 Perdew, J.P., Burke, K., and Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Phys. Rev. Lett.* 77: 3865–3868.

7 Perdew, J.P., Burke, K., and Wang, Y. (1996). Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B* 54: 16533–16539.

8 Becke, A.D. (1997). Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* 102: 8554–8560.

9 Ceperley, D.M. and Alder, B.J. (1980). Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* 45: 566–569.

10 Becke, A.D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38: 3098–3100.

11 Lee, C., Yang, W., and Parr, R.G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* 37: 785–789.

12 Colle, R. and Salvetti, O. (1975). Approximate calculation of the correlation energy for the closed shells. *Theor. Chim. Acta* 37: 329–334.

13 Miehlich, B., Savin, A., Stoll, H., and Preuss, H. (1989). Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. *Chem. Phys. Lett.* 157: 200–206.

14 Anisimov, V.I. and Gunnarsson, O. (1991). Density-functional calculation of effective Coulomb interactions in metals. *Phys. Rev. B* 43: 7570–7574.

15 Anisimov, V.I., Zaanen, J., and Andersen, O.K. (1991). Band theory and Mott insulators: Hubbard U instead of Stoner I. *Phys. Rev. B* 44: 943–954.

16 Anisimov, V.I., Solovyev, I.V., Korotin, M.A. et al. (1993). Density-functional theory and NiO photoemission spectra. *Phys. Rev. B* 48: 16929–16934.

17 Heyd, J., Scuseria, G.E., and Ernzerhof, M. (2003). Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* 118: 8207–8215.

18 Papaconstantopoulos, D.-A. (1986). *Handbook of the Band Structure of Elemental Solids*. Springer.

19 Madelung, O. (2012). *Introduction to Solid-State Theory*, vol. 2. Springer Science & Business Media.

20 Wieder, B.-J., Bradlyn, B., Wang, Z.-J. et al. (2018). Wallpaper fermions and the nonsymmorphic Dirac insulator. *Science* 361: 246.

21 Liu, G.-B., Shan, W.-Y., Yao, Y. et al. (2013). Three-band tight-binding model for monolayers of group-VIB transition metal dichalcogenides. *Phys. Rev. B* 88: 085433.

22 Kane, C.-L. and Mele, E.-J. (2005). Quantum spin hall effect in graphene. *Phys. Rev. Lett.* 95: 226801.

23 Luttinger, J.-M. (1955). Quantum theory of cyclotron resonance in semiconductors: general theory. *Phys. Rev.* 102: 1030.

24 Pikus, G.-E. (1962). A new method of calculating the energy spectrum of carriers in semiconductors. I. Neglecting spin-orbit interaction. *Sov. Phys. JETP* 14: 898.

25 Bir, G.-L. and Pikus, G.-E. (1975). *Symmetry and Strain-Induced Effects in Semiconductors*. New York: Wiley.

26 Ivchenko, E.-L. and Pikus, G.-E. (1995). *Superlattices and Other Heterostructures*. Berlin: Springer.

27 Voon, L.-C.-L.-Y. and Willatzen, M. (2009). *The kp Method: Electronic Properties of Semiconductors*. (Springer Science & Business Media.

28 Liu, C.-X., Qi, X.-L., Zhang, H.-J. et al. (2010). Model Hamiltonian for topological insulators. *Phys. Rev. B* 82: 045122.

29 Zhang, W., Yu, R., Zhang, H.-J. et al. (2010). First-principles studies of the three-dimensional strong topological insulators Bi_2Te_3, Bi_2Se_3 and Sb_2Te_3. *New J. Phys.* 12: 065013.