

Part I
Theoretical Background

Chapter 2

Density-functional Theory

The Schrödinger equation provides a theoretical way to investigate the quantum-mechanical properties of systems like atoms, molecules *etc.* Unfortunately, for most systems of practical interest the Schrödinger equation cannot be solved exactly anymore and approximations must be applied. In this respect, Density-functional theory (DFT) has become a most popular and powerful tool in studying physics and chemistry topics, and Walter Kohn was awarded by Nobel Prize in 1998 [1] for his marvelous contribution on DFT. As Isaac Newton said:” *If I have seen further it is by standing on ye shoulders of giants*”. Before heading on to the modern DFT, we are going to retrace the history of this theory.

2.1 The Many-electron Problem

In order to understand or predict properties of materials including N electrons and M nuclei from *ab initio* calculations, the Schrödinger equation for time-independent nonrelativistic cases can be written as,

$$\hat{H}\Psi(\mathbf{r}_i, \mathbf{R}_j) = E\Psi(\mathbf{r}_i, \mathbf{R}_j) \quad , \quad (2.1)$$

where \hat{H} is the Hamiltonian operator, and E is the eigenvalue of the operator. Ψ is the corresponding wave function, and \mathbf{r}_i and \mathbf{R}_j are the coordinates of electron i and nucleus j , respectively. The Hamiltonian contains several contributions: Kinetic energies of electrons (T_e) and nuclei (T_n); interaction potentials between particles, in form of repulsive potentials between nuclei (V_{nn}), attractive potentials between nuclei and electrons (V_{ne}) and repulsive potentials between electrons (V_{ee}). Then, the Hamiltonian operator \hat{H} is rewritten as,

$$\hat{H} = T_n + T_e + V_{nn} + V_{ne} + V_{ee} \quad . \quad (2.2)$$

Applying atomic units, *i.e.* $m_e = \hbar = e = 1$, these components of the Hamiltonian

operator are written as,

$$T_n = -\frac{1}{2} \sum_{j=1}^M \frac{1}{m_j} \nabla_j^2 \quad , \quad (2.3)$$

$$T_e = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 \quad , \quad (2.4)$$

$$V_{nn} = \sum_{i=1}^M \sum_{j>i}^M \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \quad , \quad (2.5)$$

$$V_{ne} = \sum_{i=1}^N \sum_{j=1}^M \frac{-Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \quad , \quad (2.6)$$

$$V_{ee} = \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad , \quad (2.7)$$

where m_j , Z_j and ∇_j are the mass of nucleus j , its atomic number and Laplacian operator, respectively.

Although Eq. 2.1 is the most accurate way to obtain the properties of materials, it is too complex to solve due to the large number of variables that the wave function depends on. For a system with M nuclei and N electrons, we have $3M + 3N$ variables, *i.e.* 3 coordinates and 3 momenta for each particle, respectively. In order to solve Eq. 2.1 in practice for many-electron systems, approximations have to be applied. The first important approximation is the Born-Oppenheimer (*BO*) nonrelativistic approximation: Loosely stated, it assumes that the movement of the electrons is so fast that they can catch up the movement of the nuclei instantaneously and relax to the corresponding ground state instantaneously, because the mass of an electron is so much smaller than that of a nucleus. (In the worst case of a H atom, 1 electron $\approx 1/1800$ of the mass of nucleus) Then, nuclei can be viewed as static, enabling a splitting of the full Hamiltonian (\hat{H}) in Eq. 2.1 into two parts: Nuclear part (\hat{H}_N) and electronic part (\hat{H}_e). Correspondingly, the total wave function (Ψ) is split into the following form:

$$\Psi(\mathbf{r}_i, \mathbf{R}_j) = \Psi_e^{\mathbf{R}}(\mathbf{r}_i) \Psi(\mathbf{R}_j) \quad , \quad (2.8)$$

where $\Psi_e^{\mathbf{R}}(\mathbf{r}_i)$ is the wave function of electrons with current nuclei positions, and $\Psi(\mathbf{R}_j)$ is the wave function of nuclei. The total energy of a system then equals the sum of the nuclear energy (E_N) and electronic energy (E_e), $E = E_N + E_e$. For fixed atomic coordinates E_N is a constant and we can focus on the electron Schrödinger equation,

$$\hat{H}_e \Psi_e(\mathbf{r}_i) = E_e \Psi_e(\mathbf{r}_i) \quad , \quad (2.9)$$

where

$$\hat{H}_e = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N \sum_{j=1}^M \frac{-Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad . \quad (2.10)$$

It is clear that the *BO* approximation decreases the number of variables in Eq. 2.1 from $3N + 3M$ to $3N$, but this computational burden still requires further approximations for most practical systems. One fundamental approach to solve the electronic Schrödinger equation (Eq. 2.9) numerically is the Hartree-Fock approximation, which transfers the many-body problem into a single particle problem through approximating the electronic wave function $\Psi_e(\mathbf{r}_i)$ by a Slater-determinant of single particle wave functions. This ensures the antisymmetry of the wave function, required to fulfill to Pauli principle, and thereby accounts for a quantum mechanical contribution to the potential, termed exchange potential (V_x). On the other hand, this approximation does not account for the quantum mechanical interaction between electrons of like spin. Compared to the full potential (V_{ee}), the Hartree-Fock potential thus includes the classical Coulomb potential (V_{ee}^c) and the exchange potential (V_x), but misses a part termed correlation potential (V_c).

Although this term is smaller than the other two, it is significant to obtain accurate results. In order to improve the original Hartree-Fock approximation, there are thus further advanced approaches to account for the correlation energy. The most popular are second or fourth order perturbation theory by Møller and Plesset (MP2 or MP4) [3], configuration interaction (CI) [4], multiconfiguration self-consistent field (MCSCF) [5], and coupled cluster approaches (CC) [6]. These methods are quite accurate, but are computationally very intense.

Alternatively, density-functional theory (DFT) is a remarkable theory that replaces the complicated N -electron wave function and the associated Schrödinger equation by a formulation based on the simpler electron density (ρ).

2.2 Density-functional Theory

2.2.1 Original Idea: Thomas-Fermi Model

In 1927 Thomas and Fermi realized that statistical considerations can be used to approximate the distribution of electrons in an atom. *Electrons are distributed uniformly in the six-dimensional phase space for the motion of an electron at the rate of two for each h^3 of volume*, and there is an effective potential field that "is itself determined by the nuclear charge and this distribution of electrons." Based on this realization the Thomas-Fermi formula for the electron density can be derived [7], and finally, the famous Thomas-Fermi kinetic energy functional is,

$$T_{\text{TF}}(\rho) = C_F \int \rho^{5/3}(\mathbf{r}) d\mathbf{r}, \quad C_F = \frac{3}{10}(3\pi^2)^{2/3} = 2.871 \quad . \quad (2.11)$$

Additionally, the *electron density* [$\rho(\mathbf{r})$] is defined as the number of electrons per unit volume in a given state. The electron density at point \mathbf{r}_1 means that the probability to find an electron at this point, is (without spin),

$$\rho(\mathbf{r}_1) = N \int \dots \int |\Psi(\mathbf{r}_1, \mathbf{r}_2 \dots \mathbf{r}_N)|^2 d\mathbf{r}_2 \dots d\mathbf{r}_N \quad . \quad (2.12)$$

If we integrate Eq. 2.12 over space, we will get the total number of electrons,

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \quad . \quad (2.13)$$

Applying this definition on Eq. 2.10, the electron-nucleus attractive energy and classical electron-electron repulsive energy can be written as a function of $\rho(\mathbf{r})$,

$$V_{\text{ne}} = Z \int \frac{\rho(\mathbf{r})}{r} d\mathbf{r} \quad , \quad (2.14)$$

$$V_{\text{ee}}^{\text{c}} = \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad . \quad (2.15)$$

For simplicity, we here neglect the exchange-correlation energy in the expectation value of Eq. 2.10 ($\langle \Psi_e | \hat{H}_e | \Psi_e \rangle$) for the moment. Together with the kinetic energy functional, Eq. 2.11 we then reach the energy functional of the Thomas-Fermi theory of atoms in terms of the electron density,

$$E_{\text{TF}}[\rho(\mathbf{r})] = C_F \int \rho^{5/3}(\mathbf{r}) d\mathbf{r} + Z \int \frac{\rho(\mathbf{r})}{r} d\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad . \quad (2.16)$$

Under the constraint of Eq. 2.13, the electron density is used to minimize the energy functional $E_{\text{TF}}[\rho(\mathbf{r})]$, and to obtain the ground state energy. Unfortunately, while the Thomas-Fermi formula is quite simple and decrease the variables from $3N$ to 3, the approximations behind it (neglected V_{xc} and kinetic energy functional based on a non-interacting, homogenous electron gas) are too bold to make it of any practical use for actual calculations.

2.2.2 Hohenberg-Kohn Theorems

While the Thomas-Fermi model already casts the electronic energy into a functional of electron density, it did not become popular because of its low accuracy. The situation changed with the landmark paper published by Hohenberg and Kohn [8] in 1964, which demonstrates that the Thomas-Fermi formula is only an approximation to an exact theory, the *density-functional theory*. For nondegenerate ground states, the Hohenberg-Kohn theory is based on two theorems, which legitimize the use of the electron density $\rho(\mathbf{r})$ as basic variable to obtain the ground state energy. The first theorem states: *The external potential $v(\mathbf{r})$ is determined, within a trivial additive constant, by the electron density $\rho(\mathbf{r})$.* Here $v(\mathbf{r})$ does not restrict to the Coulomb potential. The proof of this theorem is quite simple by employing the minimum-energy principle for the ground state. Consider two external potentials v and v' that differ by more than a constant, but both give the same ρ for their ground state. The different Hamiltonians H and H' whose ground state densities are the same then determine two different normalized wave functions Ψ and Ψ' . Taking Ψ' as a trial function for

H , we would have,

$$\begin{aligned} E_0 < \langle \Psi' | \hat{H} | \Psi' \rangle &= \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle \\ &= E'_0 + \int \rho(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r} \quad , \end{aligned} \quad (2.17)$$

where E_0 and E'_0 are the ground state energies for \hat{H} and \hat{H}' , respectively. In the same way, taking Ψ as a trial function for H' we would get,

$$E'_0 < \langle \Psi | \hat{H}' | \Psi \rangle = E_0 - \int \rho(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r} \quad . \quad (2.18)$$

Adding Eq. 2.17 and 2.18, we get $E_0 + E'_0 < E'_0 + E_0$. This contradiction indicates that there cannot be two different v that give the same ρ for their ground states. The first Hohenberg-Kohn theorem thus tell us that the ground state density ρ determines all properties of a system. Then the total energy of a configuration at a certain potential v can be written as,

$$\begin{aligned} E_v[\rho] &= T[\rho] + V_{ne}[\rho] + V_{ee}[\rho] \\ &= F_{\text{HK}}[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \quad , \end{aligned} \quad (2.19)$$

where

$$F_{\text{HK}}[\rho] = T[\rho] + V_{ee}[\rho] \quad . \quad (2.20)$$

F_{HK} is called *universal functional* of ρ due to its independence on the external potential.

The second Hohenberg-Kohn theorem states: *For a trial density $\tilde{\rho}(\mathbf{r})$, such that $\tilde{\rho}(\mathbf{r}) \geq 0$ and $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$, then,*

$$E_0 \leq E_v[\tilde{\rho}(\mathbf{r})] \quad , \quad (2.21)$$

where $E_v[\tilde{\rho}(\mathbf{r})]$ is the energy depending on v . This theorem is a justification to use the variational principle to determine the ground state energy E_0 . Based on the first theorem, $\tilde{\rho}$ determines its own potential \tilde{v} , Hamiltonian $\hat{\tilde{H}}$ and wave function $\tilde{\Psi}$. Taking $\tilde{\Psi}$ as a trial function for a Hamiltonian (\hat{H}) with the true external potential v ,

$$\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle = \int \tilde{\rho}(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + F_{\text{HK}}[\tilde{\rho}] = E_v[\tilde{\rho}] \geq E_v[\rho] \quad , \quad (2.22)$$

which is the desired result.

Additionally, applying the variational principle with constrained $\int \rho(\mathbf{r}) d\mathbf{r} = N$, the differentiability of $E_v[\rho]$ for the ground state density is satisfying the stationary principle,

$$\delta \{ E_v[\rho] - \mu [\int \rho(\mathbf{r}) d\mathbf{r} - N] \} = 0 \quad , \quad (2.23)$$

which gives the Euler-Lagrange equation, and

$$\mu = \frac{\delta E_v[\rho]}{\delta \rho(\mathbf{r})} = v(\mathbf{r}) + \frac{\delta F_{\text{HK}}[\rho]}{\delta \rho(\mathbf{r})} \quad , \quad (2.24)$$

where μ is the chemical potential.

2.2.3 Kohn-Sham Equation

In principle, if we knew the exact *universal functional* (F_{HK}), Eq. 2.23 is an exact equation to obtain the ground-state properties. Unfortunately, this exact F_{HK} is elusive, leaving the Hohenberg-Kohn theorems without much practical relevance. The break-through came only one year after the HK theorems. In 1965 Kohn and Sham [9] approximated the F_{HK} functional, and made the theorems practically useful by introducing the concept of a non-interacting reference system. Within this approach, the kinetic energy of the non-interacting reference system still exhibiting the real density $\rho(\mathbf{r})$ can be written as,

$$T_s = -\frac{1}{2} \sum_i^N \langle \varphi_i | \nabla^2 | \varphi_i \rangle \quad , \quad (2.25)$$

where φ_i are single-particle Kohn-Sham orbitals. Here the non-interacting kinetic energy is not equal to the true kinetic energy (T) of the interacting system, $T_s \neq T$, and actually $T = T_s + T_c$. Therefore, Kohn and Sham rewrote the universal functional as,

$$F[\rho] = T_s[\rho] + J[\rho] + E_{\text{xc}}[\rho] \quad , \quad (2.26)$$

and the exchange-correlation energy is defined as,

$$E_{\text{xc}}[\rho] = (T[\rho] - T_s[\rho]) + (E_{\text{ee}}[\rho] - J[\rho]) \quad . \quad (2.27)$$

The exchange-correlation energy thus contains everything that is unknown: The non-classical effects of exchange and correlation, which are contributions to the potential energy of the system, and a portion belonging to the kinetic energy. With this formulation, the many-body problem is mapped onto an effective single particle problem. Kohn and Sham thus established that for any real (interacting) system with ground-state density $\rho(\mathbf{r})$ there always exists a noninteracting system with the same ground-state density $\rho(\mathbf{r})$. This leads to the famous Kohn-Sham equation,

$$\left[-\frac{1}{2} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \varphi_i = \epsilon_i \varphi_i \quad , \quad (2.28)$$

where φ_i are the single particle wave functions, or Kohn-Sham orbitals, and ϵ_i are the Kohn-Sham orbital energies. $V_{\text{eff}}(\mathbf{r})$ is the effective potential, which contains the classical Coulomb potential, exchange-correlation potential and the external potential $V_{\text{ext}}(\mathbf{r})$,

$$V_{\text{eff}}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + V_{\text{xc}}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r}) \quad . \quad (2.29)$$

The density of the real system, $\rho(\mathbf{r})$, can be expressed in terms of the Kohn-Sham orbitals, $\rho(\mathbf{r}) = \sum_i^N |\varphi_i(\mathbf{r})|^2$. The exchange-correlation potential, $V_{\text{xc}}(\mathbf{r})$, can be written as,

$$V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \quad , \quad (2.30)$$

i.e. as the functional derivative of $E_{xc}[\rho]$ with respect to the density. If the exact forms of E_{xc} and V_{xc} were known, the Kohn-Sham strategy would lead to the exact energy, *i.e.* the correct eigenvalue of the Hamilton operator \hat{H} of the Schrödinger equation. The Kohn-Sham approach is thus in principle exact! The approximation only enters when we have to decide on an explicit form for the unknown functional for the exchange-correlation energy E_{xc} and its corresponding potential V_{xc} . The central goal of modern density-functional theory is therefore to find better and better approximations to these two quantities.

2.2.4 Present-day Exchange-correlation Functionals

The most basic approximation to V_{xc} was introduced by Kohn and Sham in their seminal paper [9]. The so-called local density approximation (LDA) is based on the homogeneous electron gas to approximate the exchange-correlation energy. For an inhomogeneous system, the idea is to view the electron density as locally homogeneous at each point in space. The LDA xc energy then has the form,

$$E_{xc}^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc}[\rho(\mathbf{r})] d\mathbf{r} \quad , \quad (2.31)$$

where $\epsilon_{xc}(\rho)$ indicates the exchange-correlation energy per particle of a uniform electron gas of density ρ . The corresponding exchange-correlation potential becomes,

$$v_{xc}^{\text{LDA}}(\mathbf{r}) = \frac{\delta E_{xc}^{\text{LDA}}}{\delta \rho(\mathbf{r})} = \epsilon_{xc}[\rho(\mathbf{r})] + \rho(\mathbf{r}) \frac{\partial \epsilon_{xc}(\rho)}{\partial \rho} \quad . \quad (2.32)$$

Moreover, $\epsilon_{xc}(\rho)$ contains two parts: exchange $\epsilon_x(\rho)$ part and correlation $\epsilon_c(\rho)$ part,

$$\epsilon_{xc}(\rho) = \epsilon_x(\rho) + \epsilon_c(\rho) \quad , \quad (2.33)$$

where the exchange part $\epsilon_x(\rho)$ is expressed as [7, 10],

$$\epsilon_x(\rho) = -\frac{3}{4} \left(\frac{3\rho(\mathbf{r})}{\pi} \right)^{\frac{1}{3}} \quad . \quad (2.34)$$

The correlation part $\epsilon_c(\rho)$ can not be expressed by such an explicit functional form. Expressions are only known for the high-density [12, 13] and low-density [14, 15] limit, whereas for intermediate densities only numerical values are known from highly accurate quantum Monte Carlo calculations [11]. This insight was then used to parameterize suitable expressions for DFT-LDA calculations [16, 17].

The LDA gives already surprisingly good results for a wide range of realistic systems, but fails for systems that are far from the uniform electron gas limit like atoms or molecules. A straightforward correction to the LDA is a formal expansion of E_{xc} in gradients of the density. The idea also came from Hohenberg and Kohn [8]. In

general, such a density-gradient expansion (DGE) of the exchange-correlation energy has the form,

$$E_{xc}^{\text{DGE}}[\rho] = \int [\epsilon_{xc}^{(0)}(\rho) + \epsilon_{xc}^{(1)}(\rho)\nabla\rho + \epsilon_{xc}^{(2)}(\rho)|\nabla\rho|^2 + \dots]d\mathbf{r} \quad . \quad (2.35)$$

However, a first order truncated DGE does not work very well, and sometimes give even worse results than LDA. In order to solve this problem, one can introduce a more generalized expansion and consider *generalized-gradient approximation* (GGA) functionals of the form.

$$E_{xc}^{\text{GGA}}[\rho] = \int f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}))d\mathbf{r} = \int \rho\epsilon_{xc}^{\text{GGA}}(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}))d\mathbf{r} \quad . \quad (2.36)$$

As in the LDA, the GGA exchange-correlation energy also divides into two parts: E_x^{GGA} and E_c^{GGA} . There are lots of GGA versions based on different constructions mainly using as many known constraints on the V_{xc} . In the present work the functional from Perdew, Burke and Ernzerhof in 1996 (PBE) [18] is almost exclusively used.