Density Functional Theory

Foundations, Formalism, Functionals, and Frontiers

Pulkit Joshi

2026

Contents

Preface

Density Functional Theory (DFT) stands today as the most widely used quantum mechanical method for the study of the electronic structure of matter. Its domain of application spans condensed matter physics, quantum chemistry, materials science, geophysics, and increasingly biochemistry and pharmacology. The 1998 Nobel Prize in Chemistry, awarded jointly to Walter Kohn and John Pople, recognized both the theoretical foundations of DFT and the development of computational chemistry methods, underscoring the field's profound impact.

The central idea of DFT is striking in its elegance: rather than working with the enormously complex many-body wavefunction $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$, one can, in principle, obtain all ground-state properties from the much simpler electron density $n(\mathbf{r})$, a function of only three spatial coordinates. This realization, made rigorous by Hohenberg and Kohn in 1964, and made practically useful by Kohn and Sham in 1965, transformed quantum mechanics from a theory restricted to small systems into a predictive tool for systems containing hundreds or thousands of atoms.

These lecture notes are designed to be *fully self-contained*. A motivated student who works through them should need no other reference for a thorough understanding of DFT. The treatment is mathematically rigorous: every major result is either proved in full or the proof is developed in sufficient detail that the gaps are minor exercises. Physical intuition accompanies every formal development, because mathematics divorced from meaning is sterile, and intuition without mathematical grounding is unreliable.

Prerequisites. The reader is expected to have a solid foundation in:

- Quantum mechanics (Dirac notation, perturbation theory, variational principle).
- Linear algebra (eigenvalue problems, Hilbert spaces, functional analysis at a physics level).
- Classical electrostatics (Poisson equation, multipole expansions).
- Some exposure to many-body physics or quantum chemistry is helpful but not strictly required.

Structure. The notes proceed from the exact many-body problem, through the historical development of Thomas–Fermi theory, to the rigorous Hohenberg–Kohn framework, the practical Kohn–Sham scheme, the hierarchy of exchange-correlation approximations, formal topics (XC hole, derivative discontinuity, ensemble extensions), and finally to advanced methods (TD-DFT, DFT+U, GW, and machine-learning functionals). Each chapter begins with a conceptual overview and ends with a summary of key results.

Equations are numbered continuously within each chapter. Boxes highlight key definitions, theorems, and physical insights.

Chapter 1

The Many-Body Problem in Electronic Structure Theory

1.1 Overview and Goals

The goal of electronic structure theory is to solve, exactly or approximately, the quantum mechanical equations governing a collection of electrons and nuclei. From the solution, one can extract total energies, geometries, phonon spectra, optical excitations, magnetic properties, chemical reaction pathways, and a vast array of other observables. The fundamental difficulty is that electrons interact with each other through a long-ranged Coulomb potential, making the problem genuinely *many-body*: the motion of each electron depends on the instantaneous positions of all others.

In this chapter we set up the exact problem, introduce the Born–Oppenheimer approximation that separates nuclear and electronic motion, and discuss different representations of the quantum state—the wavefunction, the reduced density matrices, and the electron density. Understanding what information is contained in each representation is essential preparation for appreciating why the density, despite being a much simpler object than the wavefunction, can (in principle) serve as the complete descriptor of the ground state.

1.2 The Full Molecular Hamiltonian

Consider a system of N electrons (mass m_e, charge $-e$) and M nuclei (charges Z_I, masses M_I). In atomic units, where $\hbar = m_e = e = 4\pi\epsilon_0 = 1$, the Hamiltonian is:

$$\boxed{\hat{\mathcal{H}} = \hat{T}_e + \hat{T}_n + \hat{V}_{en} + \hat{V}_{ee} + \hat{V}_{nn}} \tag{1.1}$$

where each term is:

$$\hat{T}_e = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 \qquad \text{(electron kinetic energy)} \tag{1.2}$$

$$\hat{T}_n = -\sum_{I=1}^{M} \frac{1}{2M_I}\nabla_I^2 \qquad \text{(nuclear kinetic energy)} \tag{1.3}$$

$$\hat{V}_{en} = -\sum_{i=1}^{N}\sum_{I=1}^{M} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \qquad \text{(electron-nucleus attraction)} \tag{1.4}$$

$$\hat{V}_{ee} = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \qquad \text{(electron-electron repulsion)} \tag{1.5}$$

$$\hat{V}_{nn} = \sum_{I=1}^{M}\sum_{J>I}^{M} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \qquad \text{(nucleus-nucleus repulsion)} \tag{1.6}$$

The stationary states satisfy:

$$\hat{\mathcal{H}}\,\Psi(\mathbf{r}_1, \sigma_1, \ldots, \mathbf{r}_N, \sigma_N, \mathbf{R}_1, \ldots, \mathbf{R}_M) = E\,\Psi \tag{1.7}$$

Here $\mathbf{r}_i$ are electron positions, $\sigma_i \in \{\uparrow, \downarrow\}$ are spin labels, and $\mathbf{R}_I$ are nuclear positions. We use the compact notation $\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$ for the combined space-spin coordinate.

Physical Insight

Scale of the problem. For a single carbon atom ($N = 6$), the wavefunction depends on $6 \times 3 = 18$ spatial coordinates. For a typical biomolecule or solid with $N \sim 100$–1000 electrons, direct wavefunction methods become computationally intractable: the Hilbert space dimension grows exponentially with N. This is the fundamental motivation for density-based methods.

1.3 The Born–Oppenheimer Approximation

The proton-to-electron mass ratio is $M_p/m_e \approx 1836$. Nuclei are therefore far heavier than electrons and move much more slowly on electronic timescales. This suggests separating the problem: for each fixed nuclear configuration $\{\mathbf{R}_I\}$, solve the electronic problem, then treat nuclear motion on the resulting potential energy surface.

1.3.1 Formal derivation

Write the total wavefunction as a product (or, more generally, a sum of products—the Born–Huang expansion):

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{R}_1, \ldots, \mathbf{R}_M) = \Psi_e(\mathbf{x}_1, \ldots, \mathbf{x}_N; \{\mathbf{R}_I\})\,\chi(\mathbf{R}_1, \ldots, \mathbf{R}_M) \tag{1.8}$$

where Ψ_e depends parametrically on nuclear positions. Substituting into Eq. (1.7) and invoking the approximation that *nuclear kinetic energy acting on* Ψ_e *is negligible* (i.e., the non-adiabatic coupling terms $\langle \Psi_e | \nabla_I | \Psi_e \rangle$ and $\langle \Psi_e | \nabla_I^2 | \Psi_e \rangle$ are small), one obtains two separate equations.

The electronic Schrödinger equation:

$$\hat{H}_e\,\Psi_e = E_e(\{\mathbf{R}_I\})\,\Psi_e \tag{1.9}$$

with the electronic Hamiltonian:

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{en}(\{\mathbf{R}_I\}) \tag{1.10}$$

The nuclear repulsion $\hat{V}_{nn}$ is a constant for fixed $\{\mathbf{R}_I\}$ and merely shifts the energy.

The nuclear Schrödinger equation:

$$\left[\hat{T}_n + E_e(\{\mathbf{R}_I\}) + \hat{V}_{nn}\right] \chi = E\,\chi \tag{1.11}$$

The electronic energy $E_e(\{\mathbf{R}_I\})$ acts as the potential energy surface for nuclear motion.

1.3.2 Validity

The BO approximation breaks down when two adiabatic potential energy surfaces come close together (avoided crossings), as happens in conical intersections relevant to photochemistry. For ground-state DFT calculations, the BO approximation is almost universally employed and its validity is well established. Henceforth we drop the nuclear kinetic energy and work entirely with the electronic Hamiltonian Eq. (1.10).

1.4 The Electronic Problem: A Many-Body Challenge

The electronic Hamiltonian, written explicitly, is:

$$\hat{H}_e = -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i v_{\text{ext}}(\mathbf{r}_i) \tag{1.12}$$

where the external potential:

$$v_{\text{ext}}(\mathbf{r}) = -\sum_I \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \tag{1.13}$$

represents the electron-nucleus attraction. The first term $\hat{T}_e$ and last term $\hat{V}_{en} = \sum_i v_{\text{ext}}(\mathbf{r}_i)$ are one-body operators (they act on one electron at a time). The electron-electron repulsion $\hat{V}_{ee}$ is a two-body operator—it irreducibly couples all pairs of electrons. This two-body coupling is the source of all the difficulty.

1.4.1 Why exact solutions are impossible for large N

The ground-state wavefunction $\Psi_0(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is antisymmetric under exchange of any two electron coordinates (Pauli principle). If we discretize each spatial coordinate on a grid of K points and two spin values, the wavefunction has $(2K)^N$ coefficients. For $K = 100$ and $N = 10$, this is $200^{10} \approx 10^{23}$ numbers—well beyond any imaginable storage. The computational cost of exact diagonalization (Full Configuration Interaction) scales as $(2K)^N$, rendering it impossible for all but the smallest systems. This exponential wall is the fundamental motivation for approximate methods, including DFT.

1.5 Reduced Density Matrices

Since $\hat{H}_e$ contains at most two-body interactions, all observable properties can be extracted from objects far simpler than the full N-body wavefunction. These are the *reduced density matrices.*

1.5.1 The p-body reduced density matrix

Definition

The p-**body reduced density matrix** (p-**RDM**) is defined by:

$$\gamma^{(p)}(\mathbf{x}_1,\dots,\mathbf{x}_p;\mathbf{x}'_1,\dots,\mathbf{x}'_p) = \binom{N}{p}\int \Psi(\mathbf{x}_1,\dots,\mathbf{x}_N)\Psi^*(\mathbf{x}'_1,\dots,\mathbf{x}'_p,\mathbf{x}_{p+1},\dots,\mathbf{x}_N)\prod_{i=p+1}^{N} d\mathbf{x}_i \tag{1.14}$$

The one-body RDM (1-RDM)

For $p = 1$:

$$\gamma(\mathbf{x};\mathbf{x}') = N\int \Psi(\mathbf{x},\mathbf{x}_2,\dots,\mathbf{x}_N)\Psi^*(\mathbf{x}',\mathbf{x}_2,\dots,\mathbf{x}_N)\, d\mathbf{x}_2\cdots d\mathbf{x}_N \tag{1.15}$$

The 1-RDM gives the expectation value of any one-body operator $\hat{O} = \sum_i o(\mathbf{x}_i)$:

$$\langle\hat{O}\rangle = \int o(\mathbf{x})\,\gamma(\mathbf{x};\mathbf{x}')\Big|_{\mathbf{x}'=\mathbf{x}} d\mathbf{x} \tag{1.16}$$

In particular, the *electron density* is the diagonal of the 1-RDM (summed over spin):

$$\boxed{n(\mathbf{r}) = \sum_\sigma \gamma(\mathbf{r}\sigma;\mathbf{r}\sigma) = N\sum_\sigma \int |\Psi(\mathbf{r}\sigma,\mathbf{x}_2,\dots,\mathbf{x}_N)|^2\, d\mathbf{x}_2\cdots d\mathbf{x}_N} \tag{1.17}$$

and the kinetic energy:

$$\langle\hat{T}_e\rangle = -\frac{1}{2}\int \nabla^2_{\mathbf{r}}\,\gamma(\mathbf{x};\mathbf{x}')\Big|_{\mathbf{x}'=\mathbf{x}} d\mathbf{x} \tag{1.18}$$

The two-body RDM (2-RDM)

For $p = 2$:

$$\Gamma(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}'_1,\mathbf{x}'_2) = \binom{N}{2}\int \Psi(\mathbf{x}_1,\mathbf{x}_2,\dots,\mathbf{x}_N)\Psi^*(\mathbf{x}'_1,\mathbf{x}'_2,\mathbf{x}_3,\dots,\mathbf{x}_N)\, d\mathbf{x}_3\cdots d\mathbf{x}_N \tag{1.19}$$

The electron-electron repulsion energy is obtained from the diagonal of the 2-RDM:

$$\langle\hat{V}_{ee}\rangle = \int \frac{\Gamma(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1,\mathbf{x}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}\, d\mathbf{x}_1\, d\mathbf{x}_2 \tag{1.20}$$

Physical Insight

Key point. Since $\hat{H}_e$ contains at most two-body operators, the ground-state energy and all ground-state properties are *completely determined* by the 2-RDM. Knowledge of the full N-body wavefunction is not required. The 2-RDM is a function of 8 variables (4 space-spin coordinates), vastly simpler than the wavefunction. However, the problem of finding the 2-RDM directly (without computing the full wavefunction) is extremely hard because characterizing which 2-RDMs arise from antisymmetric N-body wavefunctions (the N-representability problem) is computationally intractable in general.

1.5.2 The pair density

The diagonal of the 2-RDM (summed over spins) is the *pair density*:

$$n^{(2)}(\mathbf{r},\mathbf{r}') = \sum_{\sigma\sigma'} \Gamma(\mathbf{r}\sigma,\mathbf{r}'\sigma';\mathbf{r}\sigma,\mathbf{r}'\sigma') \tag{1.21}$$

It measures the probability of simultaneously finding one electron at $\mathbf{r}$ and another at $\mathbf{r}'$. The pair density satisfies:

$$\int n^{(2)}(\mathbf{r},\mathbf{r}')\,d\mathbf{r}' = (N-1)\,n(\mathbf{r}) \tag{1.22}$$

1.6 Properties of the Electron Density

1.6.1 Normalization and positivity

The density is non-negative and normalizes to the electron number:

$$n(\mathbf{r}) \geq 0 \quad \forall\,\mathbf{r}, \qquad \int n(\mathbf{r})\,d\mathbf{r} = N \tag{1.23}$$

1.6.2 Nuclear cusps

The electron density has cusps (discontinuities in gradient) at nuclear positions. For a nucleus at $\mathbf{R}_I$ with charge Z_I:

$$\left.\frac{\partial \bar{n}(r)}{\partial r}\right|_{r=0} = -2Z_I\,\bar{n}(0) \tag{1.24}$$

where $\bar{n}(r)$ is the spherically averaged density around nucleus I. This is the *Kato cusp condition*, and it has important consequences for the design of approximate functionals and basis sets.

1.6.3 Asymptotic decay

In the exact ground state, the density decays exponentially at large distances:

$$n(\mathbf{r}) \xrightarrow{r\to\infty} A\,r^{\alpha}\,e^{-2\sqrt{2I_P}\,r} \tag{1.25}$$

where $I_P = -\epsilon_{HOMO}$ is the ionization potential and α is a system-dependent parameter. This exact asymptotic behavior is not reproduced by most approximate functionals, a fact with significant consequences for charge transfer and ionization calculations.

1.6.4 The density as a basic variable

A fundamental question arises: can the density $n(\mathbf{r})$, a function of only 3 coordinates, contain the same information as the $3N$-coordinate wavefunction? The answer—yes, for ground states—is the content of the Hohenberg–Kohn theorem, which we prove rigorously in Chapter 3. Before reaching that theorem, we examine the historical precursor: Thomas–Fermi theory.

Chapter 2

Thomas–Fermi Theory and Its Extensions

2.1 Historical Context

Long before Hohenberg and Kohn, Thomas (1927) and Fermi (1928) independently proposed an approach to the electronic structure of atoms based entirely on the electron density. Their model was motivated by treating the electrons as a uniform gas whose local properties could be described by the exact results for the homogeneous electron gas (HEG). While Thomas–Fermi theory is quantitatively poor—it cannot describe chemical bonding and gives incorrect shell structure—it contains the germs of modern DFT and provides an essential warm-up for the Hohenberg–Kohn formalism.

2.2 The Homogeneous Electron Gas

The *homogeneous electron gas* (HEG), also called the uniform or jellium model, consists of N electrons in a volume Ω with a uniform positive background charge maintaining overall neutrality. As $N, \Omega \to \infty$ with $n = N/\Omega$ held fixed, this defines the HEG at density n.

2.2.1 Kinetic energy of the HEG

The non-interacting kinetic energy of the HEG is obtained by filling single-particle plane-wave states $\phi_{\mathbf{k}}(\mathbf{r}) = \Omega^{-1/2} e^{i\mathbf{k}\cdot\mathbf{r}}$ up to the Fermi wavevector k_F:

$$n = \frac{k_F^3}{3\pi^2} \quad \Rightarrow \quad k_F = (3\pi^2 n)^{1/3} \tag{2.1}$$

The kinetic energy per unit volume (kinetic energy density) is:

$$\tau_s^{\mathrm{HEG}}(n) = 2 \cdot \frac{1}{(2\pi)^3} \int_{|\mathbf{k}| \le k_F} \frac{k^2}{2} d^3k = \frac{3}{10}(3\pi^2)^{2/3} n^{5/3} \equiv C_F\, n^{5/3} \tag{2.2}$$

where the factor of 2 accounts for spin degeneracy, and:

$$C_F = \frac{3}{10}(3\pi^2)^{2/3} \approx 2.871 \text{ (in atomic units)} \tag{2.3}$$

2.2.2 Exchange energy of the HEG

Fock (1930) showed that the exchange energy per electron of the HEG is:

$$\epsilon_x^{\mathrm{HEG}}(n) = -\frac{3}{4}\left(\frac{3n}{\pi}\right)^{1/3} \equiv -C_x\, n^{1/3}, \quad C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \tag{2.4}$$

The exchange energy per unit volume is thus $\epsilon_x^{\text{HEG}} \cdot n = -C_x n^{4/3}$.

Derivation sketch. The exchange energy arises from the Fock integral in Hartree–Fock theory. For a single Slater determinant of plane waves:

$$E_x = -\frac{1}{2} \int \int \frac{|\gamma(\mathbf{r}, \mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}\, d\mathbf{r}' \tag{2.5}$$

where $\gamma(\mathbf{r}, \mathbf{r}') = 2 \sum_{|\mathbf{k}| \leq k_F} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}/(2\pi)^3$ is the 1-RDM of the HEG. Evaluating this integral yields Eq. (2.4).

2.2.3 Correlation energy of the HEG

The correlation energy of the HEG is defined as the difference between the exact energy and the Hartree–Fock energy:

$$E_c^{\text{HEG}} = E_{\text{exact}}^{\text{HEG}} - E_{HF}^{\text{HEG}} \tag{2.6}$$

The exact correlation energy of the HEG is known only from quantum Monte Carlo (QMC) simulations. The standard reference data are from Ceperley and Alder (1980), later parameterized by Perdew and Zunger (1981) and by Vosko, Wilk, and Nusair (1980). In the high-density (weakly correlated) limit $r_s \to 0$ (where $r_s = (3/4\pi n)^{1/3}$ is the Wigner-Seitz radius):

$$\epsilon_c^{\text{HEG}}(r_s) = A \ln r_s + B + C\, r_s \ln r_s + D\, r_s + \mathcal{O}(r_s^2 \ln r_s) \tag{2.7}$$

and in the low-density (strongly correlated, Wigner crystal) limit $r_s \to \infty$:

$$\epsilon_c^{\text{HEG}}(r_s) = -\frac{d_0}{r_s} + \frac{d_1}{r_s^{3/2}} + \cdots \tag{2.8}$$

2.3 The Thomas–Fermi Energy Functional

Thomas and Fermi proposed to approximate the kinetic energy of the inhomogeneous electron gas by treating each small volume element as a piece of the HEG at the local density $n(\mathbf{r})$:

The Thomas–Fermi Functional

$$E_{TF}[n] = T_{TF}[n] + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r})\, d\mathbf{r} + J[n] \tag{2.9}$$

where:

$$T_{TF}[n] = C_F \int n(\mathbf{r})^{5/3}\, d\mathbf{r} \tag{2.10}$$

$$J[n] = \frac{1}{2} \iint \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' \tag{2.11}$$

$J[n]$ is the classical electrostatic (Hartree) energy.

2.3.1 Minimization and the TF equation

Minimizing $E_{TF}[n]$ subject to $\int n\, d\mathbf{r} = N$ (using a Lagrange multiplier μ):

$$\frac{\delta E_{TF}}{\delta n(\mathbf{r})} = \mu \tag{2.12}$$

gives the *Thomas–Fermi equation*:

$$\frac{5}{3} C_F n(\mathbf{r})^{2/3} + v_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' = \mu \tag{2.13}$$

Identifying the local Fermi kinetic energy with the local chemical potential, this can be rewritten as $\frac{1}{2}(3\pi^2)^{2/3}n(\mathbf{r})^{2/3} = \mu - v(\mathbf{r})$, where $v(\mathbf{r})$ is the total electrostatic potential.

2.3.2 The Thomas–Fermi–Dirac model

Dirac (1930) added the exchange energy:

$$E_{TFD}[n] = E_{TF}[n] - C_x \int n(\mathbf{r})^{4/3}\, d\mathbf{r} \tag{2.14}$$

2.4 Failures of Thomas–Fermi Theory

Thomas–Fermi theory, despite being the first density functional theory, has severe quantitative and qualitative failures:

(i) **Teller's theorem.** Molecules do not bind in TF theory: the total energy of a molecule is always greater than the sum of the atomic energies. This was proven rigorously by Teller (1962) and Balazs (1967). The physical reason is that TF theory misses the quantum mechanical lowering of kinetic energy associated with delocalization across multiple nuclear centers.

(ii) **Incorrect shell structure.** TF theory gives a smooth, monotonically decreasing density that does not reflect atomic shell structure. The oscillations in the density associated with filled shells—which determine chemical periodicity—are absent.

(iii) **Wrong asymptotic decay.** The TF density decays as r^{-6} at large distances, in conflict with the exact $e^{-2\sqrt{2I_P}\,r}$ decay (Eq. 1.25).

(iv) **Poor kinetic energy.** The TF approximation to T_s is only accurate to leading order in an expansion parameter proportional to $(\nabla n/n^{4/3})^2$. For real systems, density gradients are far from small.

Despite these failures, TF theory contains two enduring insights: (1) the density alone can serve as the basic variable, and (2) the key challenge is to find accurate kinetic and exchange-correlation energy functionals of the density.

2.5 The von Weizsäcker Correction

von Weizsäcker (1935) proposed a gradient correction to the TF kinetic energy:

$$T_W[n] = \frac{1}{8} \int \frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})}\, d\mathbf{r} \tag{2.15}$$

This can be derived by writing $n = |\psi|^2$ and computing $\langle \hat{T} \rangle$ for a single orbital ψ, giving an exact expression for one-electron and two-electron systems. The full kinetic energy functional for the interacting system interpolates between T_{TF} (accurate at high density, slow variation) and T_W (accurate for one or two electrons, rapid variation). The combination $T_{TF} + \lambda T_W$ (with λ fitted empirically) was later generalized into the family of Thomas–Fermi–Weizsäcker (TFW) models, which remain active research topics in orbital-free DFT.

Chapter 3

The Hohenberg–Kohn Theorems

3.1 Overview and Significance

The Hohenberg–Kohn (HK) theorems, published in 1964, placed the use of the electron density as the basic variable on rigorous mathematical footing. They established that:

(I) The ground-state electron density $n(\mathbf{r})$ *uniquely* determines the external potential $v_{\text{ext}}(\mathbf{r})$ (and hence the entire Hamiltonian, and hence all properties of the ground state).

(II) There exists a universal energy functional $E_v[n]$ whose minimum over all v-representable densities gives the exact ground-state energy.

These theorems justify replacing the wavefunction with the density as the fundamental descriptor. They do not, however, tell us the form of the functional, nor do they provide a practical scheme for computation. That comes in Chapter 4.

3.2 v-Representability and N-Representability

Before stating the HK theorems precisely, we need two definitions.

v-Representable Densities

A density $n(\mathbf{r})$ is **v-representable** if there exists some external potential $v_{\text{ext}}(\mathbf{r})$ such that $n(\mathbf{r})$ is the ground-state density of the Hamiltonian $\hat{H} = \hat{T} + \hat{V}_{ee} + \int v_{\text{ext}} n \, d\mathbf{r}$.

N-Representable Densities

A density $n(\mathbf{r})$ is **N-representable** if there exists some antisymmetric N-electron wavefunction Ψ (not necessarily a ground state of any Hamiltonian) such that:

$$n(\mathbf{r}) = N \int |\Psi(\mathbf{r}, \mathbf{x}_2, \ldots, \mathbf{x}_N)|^2 \, d\mathbf{x}_2 \cdots d\mathbf{x}_N$$

N-representability is weaker than v-representability: every v-representable density is N-representable, but not vice versa. Harriman (1981) showed that the set of N-representable densities is simply all non-negative, normalized functions with finite kinetic energy—a large and tractable set. The set of v-representable densities is more restrictive and not fully characterized; this is a subtle problem in functional analysis. The Levy constrained-search formulation (Section 3.6) circumvents the v-representability issue by extending the domain of the HK functional to all N-representable densities.

3.3 The First Hohenberg–Kohn Theorem: Uniqueness

Hohenberg–Kohn Theorem I

(1964) The ground-state electron density $n_0(\mathbf{r})$ uniquely determines the external potential $v_{\text{ext}}(\mathbf{r})$, up to an additive constant. Equivalently, two different external potentials (differing by more than a constant) cannot have the same ground-state density.

Proof. We proceed by reductio ad absurdum. Assume two external potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ with $v - v' \neq \text{const}$, both yielding the *same* ground-state density $n_0(\mathbf{r})$.

Let $\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}$, $\hat{H}' = \hat{T} + \hat{V}_{ee} + \hat{V}'$, with ground-state wavefunctions Ψ_0 and Ψ_0' and ground-state energies E_0 and E_0' respectively.

Since Ψ_0' is not the ground state of $\hat{H}$ (assuming non-degeneracy), the variational principle gives a strict inequality:

$$\begin{aligned} E_0 &< \langle \Psi_0' | \hat{H} | \Psi_0' \rangle \\ &= \langle \Psi_0' | \hat{H}' | \Psi_0' \rangle + \langle \Psi_0' | \hat{H} - \hat{H}' | \Psi_0' \rangle \\ &= E_0' + \int [v(\mathbf{r}) - v'(\mathbf{r})]\, n_0(\mathbf{r})\, d\mathbf{r} \end{aligned} \tag{3.1}$$

By the same argument with primed and unprimed exchanged:

$$E_0' < E_0 + \int [v'(\mathbf{r}) - v(\mathbf{r})]\, n_0(\mathbf{r})\, d\mathbf{r} \tag{3.2}$$

Adding Eqs. (3.1) and (3.2):

$$E_0 + E_0' < E_0' + E_0 \tag{3.3}$$

This is $0 < 0$, which is a contradiction. Therefore, the assumption that $v \neq v' + \text{const}$ yet both yield the same n_0 must be false. ☐ ☐

Remark 3.1. The proof assumes a non-degenerate ground state. For the degenerate case, if multiple wavefunctions share the same ground-state energy, the original Hohenberg–Kohn argument needs modification. The result still holds but the proof requires additional care: one can show that any member of the degenerate ground-state subspace determines the potential, and different potentials (up to a constant) cannot share any common ground-state density.

3.3.1 Consequences of HK-I

Since $n_0(\mathbf{r})$ uniquely determines $v_{\text{ext}}(\mathbf{r})$ (up to a constant), and $v_{\text{ext}}(\mathbf{r})$ fixes the Hamiltonian $\hat{H}$, it follows that $n_0(\mathbf{r})$ determines:

- The complete ground-state wavefunction $\Psi_0[v_{\text{ext}}] = \Psi_0[n_0]$ (as a functional of n_0).
- All excited states $\Psi_k[v_{\text{ext}}] = \Psi_k[n_0]$ (since $\hat{H}$ is determined).
- All ground-state observables: $\langle \Psi_0 | \hat{O} | \Psi_0 \rangle = O[n_0]$ for any operator $\hat{O}$.

In principle, the density determines everything. In practice, the functional dependence of observables on n_0 is generally unknown and the challenge of DFT is to approximate these functionals.

3.4 The Universal Hohenberg–Kohn Functional

The HK Functional

Define the **universal functional** $F_{HK}[n]$ for v-representable densities n as:

$$F_{HK}[n] = \langle \Psi_0[n] | \hat{T} + \hat{V}_{ee} | \Psi_0[n] \rangle \tag{3.4}$$

This functional is universal: it does not depend on the external potential v_{ext}, only on the density n (through the many-body wavefunction $\Psi_0[n]$).

The total energy for a given external potential v_{ext} is then:

$$E_v[n] = F_{HK}[n] + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r})\, d\mathbf{r} \tag{3.5}$$

3.5 The Second Hohenberg–Kohn Theorem: Variational Principle

Hohenberg–Kohn Theorem II

For any trial density $\tilde{n}(\mathbf{r})$ that is v-representable (i.e., is the ground-state density of some Hamiltonian with the same $\hat{T} + \hat{V}_{ee}$ but a different external potential):

$$E_v[\tilde{n}] = F_{HK}[\tilde{n}] + \int v_{\text{ext}}(\mathbf{r}) \tilde{n}(\mathbf{r})\, d\mathbf{r} \geq E_0 \tag{3.6}$$

with equality if and only if $\tilde{n} = n_0$.

Proof. Let $\tilde{n} \neq n_0$ be v-representable, with associated ground state $\tilde{\Psi}$. Then:

$$\begin{aligned} E_v[\tilde{n}] &= F_{HK}[\tilde{n}] + \int v\, \tilde{n}\, d\mathbf{r} \\ &= \langle \tilde{\Psi} | \hat{T} + \hat{V}_{ee} | \tilde{\Psi} \rangle + \int v\, \tilde{n}\, d\mathbf{r} \\ &= \langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle \geq E_0 \end{aligned} \tag{3.7}$$

The last step follows from the variational principle: $\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle \geq E_0$ for any normalized antisymmetric $\tilde{\Psi}$, with equality iff $\tilde{\Psi} = \Psi_0$, which (by HK-I) holds iff $\tilde{n} = n_0$. □ □

Physical Insight

Physical significance. HK-II tells us that the ground-state energy and density can be found by minimizing $E_v[n]$ over all (well-behaved) trial densities. This is the density analog of the wavefunction variational principle. The practical obstacle is that the exact form of $F_{HK}[n]$ is unknown.

3.6 Levy–Lieb Constrained Search

The HK proof requires v-representability of the trial density, which limits the domain of the functional and is difficult to verify in practice. Levy (1979) and independently Lieb (1983) proposed a powerful generalization that lifts this restriction.

Levy–Lieb Constrained Search

Define the **Levy–Lieb functional**:

$$F_{LL}[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \tag{3.8}$$

where the minimum is over all antisymmetric N-body wavefunctions Ψ that yield the density $n(\mathbf{r})$ [written $\Psi \to n$].

The constrained search proceeds in two stages:

1. **Inner minimization:** For fixed n, minimize $\langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle$ over all Ψ yielding that n. This defines $F_{LL}[n]$.
2. **Outer minimization:** Minimize $F_{LL}[n] + \int v_{\text{ext}} n \, d\mathbf{r}$ over all N-representable densities n.

Theorem 3.2. *The Levy–Lieb functional agrees with the Hohenberg–Kohn functional on v-representable densities: $F_{LL}[n] = F_{HK}[n]$ for all v-representable n. Moreover, F_{LL} is defined on the larger set of all N-representable densities.*

Proof. Let n be v-representable with ground state $\Psi_0[n]$. Then:

$$F_{LL}[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \leq \langle \Psi_0[n] | \hat{T} + \hat{V}_{ee} | \Psi_0[n] \rangle = F_{HK}[n] \tag{3.9}$$

Now let $\tilde{\Psi}$ be any $\Psi \to n$ and write $\tilde{n}$ for any v'-representable density with $\tilde{n} \to \tilde{\Psi}$... The full proof uses the fact that for the minimizing $\tilde{\Psi}$, there exists a v' such that $\tilde{\Psi}$ is a ground state of $\hat{T} + \hat{V}_{ee} + V'$. A detailed argument shows $\langle \tilde{\Psi} | \hat{T} + \hat{V}_{ee} | \tilde{\Psi} \rangle \geq \langle \Psi_0[n] | \hat{T} + \hat{V}_{ee} | \Psi_0[n] \rangle$ via the HK variational argument, giving $F_{LL}[n] = F_{HK}[n]$. □ □

3.6.1 Lieb's convex conjugate formulation

Lieb (1983) provided a more mathematically rigorous formulation using convex analysis. He showed that the appropriate generalization is:

$$F_{Lieb}[n] = \sup_{v} \left\{ E_v - \int v \, n \, d\mathbf{r} \right\} \tag{3.10}$$

where the supremum is over all $v \in L^{3/2}(\mathbb{R}^3) + L^{\infty}(\mathbb{R}^3)$ and E_v is the ground-state energy. This is the convex conjugate (Legendre–Fenchel transform) of $-E_v$ as a function of v. The Lieb functional is defined on all N-representable densities and is convex—a key property lacking in the original HK formulation.

3.7 Ensemble Extensions and the Zero-Temperature Limit

For systems at finite temperature, or for degenerate ground states, one works with density matrices rather than pure states. The *finite-temperature DFT* was developed by Mermin (1965), who proved that the HK theorem extends to finite temperature T by replacing the wavefunction with the grand canonical density matrix $\hat{\rho} = e^{-\beta(\hat{H} - \mu \hat{N})}/\mathcal{Z}$. The free energy $A = E - TS$ plays the role of the energy functional, and the minimizing density is the equilibrium density at temperature T and chemical potential μ.

3.8 Limitations of the Hohenberg–Kohn Theorems

The HK theorems, while profound, have important limitations:

(i) **Existence, not constructiveness.** The theorems prove that a universal functional $F[n]$ exists, but give no prescription for computing it. Finding accurate approximations to $F[n]$ is the central challenge of DFT.

(ii) **Ground states only.** Standard HK theory applies only to the ground state. Extension to excited states requires additional work (ensemble DFT, time-dependent DFT).

(iii) **Non-degenerate ground states.** The original proof assumes a non-degenerate ground state. Extensions to the degenerate case exist but require more care.

(iv) **v-representability.** Not every non-negative normalized function is v-representable. This subtlety is handled by the Levy–Lieb constrained search.

Chapter 4

The Kohn–Sham Formalism

4.1 The Need for the Kohn–Sham Approach

The Hohenberg–Kohn theorems guarantee the existence of an exact density functional, but the functional $F[n] = T[n] + V_{ee}[n]$ cannot be evaluated without the full many-body wavefunction—and computing the wavefunction is exactly what we are trying to avoid. The main obstacle is the kinetic energy functional $T[n]$: while the exchange-correlation energy E_{xc} is on the order of 10–20% of the total energy, the kinetic energy is typically the largest contribution, and the Thomas–Fermi approximation to $T[n]$ is simply not accurate enough for chemistry or materials science.

Kohn and Sham (1965) made a brilliant observation: instead of trying to approximate the kinetic energy directly as a functional of the density, introduce an auxiliary system of non-interacting electrons that has the same ground-state density as the interacting system. The kinetic energy of the non-interacting system can be computed exactly, and the remaining (small) correction is absorbed into the exchange-correlation functional.

4.2 The Non-Interacting Reference System

The Kohn–Sham System

The **Kohn–Sham (KS) system** is an auxiliary system of N non-interacting electrons moving in an effective one-body potential $v_{\text{eff}}(\mathbf{r})$ chosen such that the ground-state density of the KS system equals the ground-state density of the interacting system:

$$n_{KS}(\mathbf{r}) = n_0(\mathbf{r}) \tag{4.1}$$

The existence of such a non-interacting system with the correct density is guaranteed by the *non-interacting v-representability* assumption: we assume that the interacting ground-state density $n_0(\mathbf{r})$ is also the ground-state density of some non-interacting Hamiltonian. This is not proven in general but is believed to hold for all physically relevant densities.

4.2.1 The non-interacting kinetic energy

For the non-interacting system, the many-body wavefunction is a single Slater determinant Φ of orthonormal one-electron orbitals $\{\phi_i\}$:

$$\Phi(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det[\phi_1(\mathbf{x}_1) \cdots \phi_N(\mathbf{x}_N)] \tag{4.2}$$

The non-interacting kinetic energy is:

$$T_s[n] = \langle \Phi | \hat{T} | \Phi \rangle = -\frac{1}{2} \sum_{i=1}^{N} \langle \phi_i | \nabla^2 | \phi_i \rangle = -\frac{1}{2} \sum_{i=1}^{N} \int \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) \, d\mathbf{r} \tag{4.3}$$

This is the *exact* kinetic energy of the KS system. It differs from the true kinetic energy of the interacting system $T[n_0]$ by a small amount absorbed into E_{xc}.

4.3 The Kohn–Sham Energy Partitioning

The exact ground-state energy is partitioned as:

$$\boxed{E[n] = T_s[n] + J[n] + E_{xc}[n] + \int v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \, d\mathbf{r}} \tag{4.4}$$

where:

$$T_s[n] = -\frac{1}{2} \sum_i \int \phi_i^* \nabla^2 \phi_i \, d\mathbf{r} \qquad \text{(non-interacting KE, computed exactly)} \tag{4.5}$$

$$J[n] = \frac{1}{2} \iint \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} \, d\mathbf{r}' \qquad \text{(classical Hartree energy)} \tag{4.6}$$

$$E_{xc}[n] = F[n] - T_s[n] - J[n] \qquad \text{(exchange-correlation energy)} \tag{4.7}$$

$$\int v_{\text{ext}} n \, d\mathbf{r} \qquad \text{(electron-nuclear attraction)} \tag{4.8}$$

The exchange-correlation energy $E_{xc}[n]$ is defined as:

$$E_{xc}[n] \equiv (T[n] - T_s[n]) + (V_{ee}[n] - J[n]) \tag{4.9}$$

It contains two physically distinct contributions:

- **Kinetic correlation:** $T_c[n] = T[n] - T_s[n]$, the difference between the true kinetic energy and the non-interacting kinetic energy. This is always positive: correlation increases the kinetic energy.
- **Non-classical electron-electron interaction:** $E_{ncl}[n] = V_{ee}[n] - J[n]$, which includes the exchange energy (from Pauli antisymmetry) and the Coulomb correlation energy.

Physical Insight

The genius of the KS decomposition. The exact kinetic energy $T[n]$ is a large, difficult-to-approximate functional. The KS scheme sidesteps this by introducing orbitals: T_s is computed exactly and only the small correction $T_c = T - T_s$ (typically $\sim 1\%$ of T) is left to approximation. This is why KS-DFT is dramatically more accurate than orbital-free Thomas–Fermi theory, which approximates the full T.

4.4 Derivation of the Kohn–Sham Equations

We minimize $E[n]$ (Eq. 4.4) with respect to the orbitals $\{\phi_i\}$, subject to the orthonormality constraint $\langle \phi_i | \phi_j \rangle = \delta_{ij}$.

4.4.1 Functional derivatives

The key functional derivative relations are:

$$\frac{\delta T_s}{\delta \phi_i^*(\mathbf{r})} = -\frac{1}{2}\nabla^2 \phi_i(\mathbf{r}) \tag{4.10}$$

$$\frac{\delta J}{\delta n(\mathbf{r})} = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' = v_H(\mathbf{r}) \tag{4.11}$$

$$\frac{\delta}{\delta n(\mathbf{r})} \int v_{\text{ext}}\, n\, d\mathbf{r} = v_{\text{ext}}(\mathbf{r}) \tag{4.12}$$

$$\frac{\delta E_{xc}}{\delta n(\mathbf{r})} = v_{\text{xc}}(\mathbf{r}) \tag{4.13}$$

4.4.2 The Euler–Lagrange equations

Using the chain rule $\delta T_s/\delta n = (\delta T_s/\delta \phi_i^*)(\delta \phi_i^*/\delta n)$, the Euler–Lagrange equation for the orbitals is:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{4.14}$$

where the effective KS potential is:

$$\boxed{v_{\text{eff}}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + v_H(\mathbf{r}) + v_{\text{xc}}(\mathbf{r})} \tag{4.15}$$

with:

$$v_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' \quad \text{(Hartree/Coulomb potential)} \tag{4.16}$$

$$v_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \quad \text{(exchange-correlation potential)} \tag{4.17}$$

The density is reconstructed from the occupied orbitals:

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 = \sum_{i=1}^{N} f_i |\phi_i(\mathbf{r})|^2 \tag{4.18}$$

where $f_i \in \{0, 1\}$ are occupation numbers (or fractional in metallic systems with partial occupancies).

The Kohn–Sham Equations

The ground-state density of an N-electron system in an external potential v_{ext} can be found by solving the self-consistent set of single-particle equations:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{\text{xc}}[n](\mathbf{r})\right] \phi_i = \epsilon_i \phi_i \tag{4.19}$$

where $n(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2$. These equations are exact given the exact v_{xc}.

4.5 Self-Consistent Field Procedure

The KS equations are nonlinear: v_{eff} depends on the density n, which depends on the orbitals $\{\phi_i\}$, which are determined by v_{eff}. The solution is found iteratively via the self-consistent field (SCF) procedure:

1. **Initial guess:** Start with an initial density $n^{(0)}(\mathbf{r})$ (e.g., superposition of atomic densities).
2. **Construct v_{eff}:** Compute $v_H[n^{(k)}]$ and $v_{\text{xc}}[n^{(k)}]$, build $v_{\text{eff}}^{(k)}$.
3. **Solve KS equations:** Diagonalize the KS Hamiltonian $\hat{h}_{KS} = -\frac{1}{2}\nabla^2 + v_{\text{eff}}^{(k)}$ to get orbitals $\{\phi_i^{(k+1)}\}$.
4. **New density:** Build $n^{(k+1)} = \sum_i |\phi_i^{(k+1)}|^2$.
5. **Convergence check:** If $\|n^{(k+1)} - n^{(k)}\| <$ threshold, stop. Otherwise, mix $n^{(k+1)}$ with previous iterates (Pulay DIIS, Broyden mixing) and return to step 2.

4.5.1 Mixing and convergence

Naive SCF iterations often do not converge: small changes in the density can produce large changes in v_{xc}, leading to oscillations. Practical convergence requires density mixing schemes. The Pulay DIIS (Direct Inversion in the Iterative Subspace) method constructs the next iterate as a linear combination of previous densities minimizing a residual, and is the standard approach in most DFT codes. Charge sloshing in metallic systems requires additional techniques such as Kerker preconditioning.

4.6 Physical Interpretation of KS Eigenvalues

The KS eigenvalues ϵ_i are often interpreted as single-particle energies, but their physical meaning requires careful discussion.

4.6.1 Koopmans'-like theorem for the HOMO

For the highest occupied molecular orbital (HOMO), there is an exact statement: the KS HOMO eigenvalue ϵ_{HOMO} equals the negative of the exact ionization potential I_P:

$$\epsilon_{HOMO} = -I_P = -[E_0(N-1) - E_0(N)] \tag{4.20}$$

This follows from the exact asymptotic decay of the density (Eq. 1.25) and is an exact result for the *exact* v_{xc}. For approximate functionals, this relation is violated.

4.6.2 KS eigenvalues and band structures

For all other occupied orbitals and for unoccupied orbitals, the KS eigenvalues have no rigorous physical interpretation as electron removal or addition energies. Nevertheless, in practice:

- KS band structures of metals (where the HOMO eigenvalue is well-defined) agree remarkably well with photoemission experiments.
- KS gaps between the HOMO and LUMO systematically *underestimate* the fundamental gap (see Chapter 7 on the derivative discontinuity).
- KS band structures provide useful qualitative information about orbital character and topology.

4.6.3 The KS system is fictitious

The KS system is a mathematical construct: the KS orbitals are not the true single-particle states of the interacting system, and the KS eigenvalues are not exact excitation energies. The only property of the KS system that is physically meaningful is the ground-state density $n(\mathbf{r})$ and, from it, the ground-state energy and other ground-state observables.

4.7 Spin-Polarized Kohn–Sham DFT

For magnetic systems, or in the presence of a magnetic field, it is necessary to work with the spin-polarized formulation. The ground state is characterized by two spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$, with total density $n = n_\uparrow + n_\downarrow$ and magnetization density $m = n_\uparrow - n_\downarrow$.

The spin-DFT HK theorem states that both $n(\mathbf{r})$ and $m(\mathbf{r})$ are required to determine the ground state in the presence of a magnetic field. The energy functional becomes $E[n_\uparrow, n_\downarrow]$, and the exchange-correlation functional depends on both spin channels:

$$E_{xc}[n_\uparrow, n_\downarrow] = E_{xc}[n, m] \tag{4.21}$$

The spin-KS equations become:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}^\sigma(\mathbf{r})\right]\phi_i^\sigma = \epsilon_i^\sigma \phi_i^\sigma, \quad \sigma \in \{\uparrow, \downarrow\} \tag{4.22}$$

with $v_{\text{eff}}^\sigma = v_{\text{ext}} + v_H + v_{\text{xc}}^\sigma$ and:

$$v_{\text{xc}}^\sigma(\mathbf{r}) = \frac{\delta E_{xc}[n_\uparrow, n_\downarrow]}{\delta n_\sigma(\mathbf{r})} \tag{4.23}$$

The spin-scaling relation links the spin-polarized and spin-unpolarized exchange:

$$E_x[n_\uparrow, n_\downarrow] = \frac{1}{2}\left[E_x^{\text{unp}}[2n_\uparrow] + E_x^{\text{unp}}[2n_\downarrow]\right] \tag{4.24}$$

This exact identity is used to construct spin-polarized LDA and GGA functionals from their spin-unpolarized counterparts.

4.8 Summary of the Kohn–Sham Framework

Key Points: The Kohn–Sham Framework

1. KS-DFT maps the interacting many-body problem onto a set of non-interacting single-particle equations—but the effective potential v_{eff} contains all many-body effects via v_{xc}.
2. The KS equations are exact: all approximations enter only through $E_{xc}[n]$.
3. The SCF procedure iterates to self-consistency; convergence requires density mixing.
4. KS eigenvalues have rigorous physical meaning only for the HOMO (equals $-I_P$).
5. KS-DFT formally scales as $O(N^3)$ (dominated by matrix diagonalization); linear-scaling methods exist for large systems.
6. Spin-DFT extends the formalism to magnetic systems using $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$.

Chapter 5

The Exchange-Correlation Energy: Exact Properties and Jacob's Ladder

5.1 The Central Challenge: Approximating E_{xc}

The exchange-correlation functional $E_{xc}[n]$ is an exact but unknown functional of the density. All approximations in KS-DFT are approximations to E_{xc}. The central challenge of functional development is to construct approximations that:

(i) Satisfy as many exact constraints as possible.

(ii) Are computationally tractable.

(iii) Are accurate for the classes of systems one cares about.

Perdew has described the hierarchy of functionals as *Jacob's Ladder of DFT*, with higher rungs incorporating more ingredients and satisfying more constraints, at greater computational cost. The rungs are:

Rung	**Name**	**Ingredients**
1	LDA	$n(\mathbf{r})$
2	GGA	$n(\mathbf{r}), \|\nabla n\|$
3	meta-GGA	$n, \|\nabla n\|, \tau$ or $\nabla^2 n$
4	Hybrids	above + E_x^{HF} (nonlocal)
5	Double hybrids	above + MP2 correlation

5.2 Exact Constraints on E_{xc}

Before discussing approximations, it is essential to know what the exact E_{xc} satisfies. These exact constraints serve as targets for approximation.

5.2.1 Separation into exchange and correlation

The exact E_{xc} can be partitioned as:

$$E_{xc}[n] = E_x[n] + E_c[n] \tag{5.1}$$

where the exchange energy:

$$E_x[n] = \langle \Phi[n] | \hat{V}_{ee} | \Phi[n] \rangle - J[n] \tag{5.2}$$

is the Fock exchange energy of the KS Slater determinant $\Phi[n]$, and the correlation energy E_c contains everything else.

5.2.2 Scaling relations

Under uniform coordinate scaling $n_\lambda(\mathbf{r}) = \lambda^3 n(\lambda\mathbf{r})$ (which preserves normalization):

$$T_s[n_\lambda] = \lambda^2 T_s[n] \tag{5.3}$$

$$E_x[n_\lambda] = \lambda E_x[n] \tag{5.4}$$

$$E_c[n_\lambda] \geq \lambda E_c[n] \quad (\lambda > 1) \tag{5.5}$$

$$E_c[n_\lambda] \leq \lambda E_c[n] \quad (\lambda < 1) \tag{5.6}$$

The scaling of E_x (linear in λ) is exact and provides a strong constraint on exchange functionals.

5.2.3 Non-positivity of exchange

The exchange energy is always negative:

$$E_x[n] \leq 0 \tag{5.7}$$

This follows from the fact that exchange is the Fock self-interaction of a single Slater determinant.

5.2.4 Non-positivity of correlation

The correlation energy also satisfies:

$$E_c[n] \leq 0 \tag{5.8}$$

This follows from the variational principle: $F[n] = \min_{\Psi\to n}\langle\Psi|\hat{T}+\hat{V}_{ee}|\Psi\rangle \leq \langle\Phi[n]|\hat{T}+\hat{V}_{ee}|\Phi[n]\rangle = T_s[n] + J[n] + E_x[n]$, so $E_c = F - T_s - J - E_x \leq 0$.

5.2.5 The Lieb–Oxford bound

An important lower bound on E_{xc} is the Lieb–Oxford inequality:

$$E_{xc}[n] \geq E_x[n] \geq -C_{LO}\int n^{4/3}\,d\mathbf{r}, \quad C_{LO} \approx 1.68 \tag{5.9}$$

(The exact constant is bounded by $1.44 \leq C_{LO} \leq 1.68$; the value 1.68 is an upper bound.) This bound constrains the strength of exchange-correlation functionals and is used to fix parameters in LDA and meta-GGA functionals.

5.2.6 One-electron self-interaction freedom

For any one-electron density $n_1(\mathbf{r}) = |\phi|^2$ (normalized to 1):

$$E_{xc}[n_1] + J[n_1] = 0 \tag{5.10}$$

This states that for a single electron, the exchange exactly cancels the self-Hartree energy, leaving zero net self-interaction. This is an exact constraint that most approximate functionals violate, leading to the *self-interaction error* (SIE), discussed in Chapter **??**.

5.2.7 The uniform electron gas limit

Any reasonable functional must reduce to the HEG result when applied to a uniform density:

$$E_{xc}[n = \text{const}] = \int n\,\epsilon_{xc}^{\text{HEG}}(n)\,d\mathbf{r} \tag{5.11}$$

5.3 Local Density Approximation (LDA)

Local Density Approximation

The **LDA** treats each infinitesimal volume element as a piece of the HEG at the local density $n(\mathbf{r})$:

$$E_{xc}^{LDA}[n] = \int n(\mathbf{r})\, \epsilon_{xc}^{\text{HEG}}(n(\mathbf{r}))\, d\mathbf{r} \tag{5.12}$$

where $\epsilon_{xc}^{\text{HEG}}(n)$ is the exchange-correlation energy per electron of the HEG at density n.

5.3.1 LDA exchange

The exchange part is known analytically from Dirac (1930):

$$\epsilon_{x}^{\text{HEG}}(n) = -\frac{3}{4}\left(\frac{3n}{\pi}\right)^{1/3} \tag{5.13}$$

giving:

$$E_{x}^{LDA}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \int n^{4/3}\, d\mathbf{r} \tag{5.14}$$

5.3.2 LDA correlation

The correlation part is obtained from quantum Monte Carlo simulations of the HEG by Ceperley and Alder (1980). Several analytical parameterizations exist. The most commonly used are:

Vosko–Wilk–Nusair (VWN, 1980): An accurate rational interpolation:

$$\epsilon_c^{VWN}(r_s) = \frac{A}{2}\left\{\ln\frac{r_s}{X(r_s)} + \frac{2b}{Q}\arctan\frac{Q}{2r_s^{1/2}+b} - \frac{bx_0}{X(x_0)}\left[\ln\frac{(r_s^{1/2}-x_0)^2}{X(r_s)} + \frac{2(2x_0+b)}{Q}\arctan\frac{Q}{2r_s^{1/2}+b}\right]\right\} \tag{5.15}$$

where $X(x) = x^2 + bx + c$, $Q = \sqrt{4c - b^2}$, and the parameters A, b, c, x_0 are fitted to QMC data.

Perdew–Zunger (PZ81, 1981): Simpler piecewise parameterization:

$$\epsilon_c^{PZ}(r_s) = \begin{cases} A\ln r_s + B + Cr_s\ln r_s + Dr_s & r_s < 1 \\ \frac{\gamma}{1+\beta_1\sqrt{r_s}+\beta_2 r_s} & r_s \geq 1 \end{cases} \tag{5.16}$$

5.3.3 Why does LDA work?

At first sight, it seems surprising that the LDA—which treats real inhomogeneous systems as locally homogeneous—works at all for atoms and molecules, where the density varies enormously. The reason is a *cancellation of errors*: LDA satisfies the sum rule on the XC hole (see Chapter 6) exactly (because it uses the HEG XC hole, which satisfies the sum rule). Even though the LDA XC hole has the wrong shape for inhomogeneous systems, it has the right integral, and E_{xc} depends only on the spherically-averaged coupling-constant-integrated hole. This cancellation makes LDA far more accurate than its simplicity suggests.

5.3.4 Failures of LDA

Despite its surprising effectiveness, LDA has systematic failures:

- **Overbinding:** LDA overbinds molecules and solids; bond lengths are typically 1–2% too short; atomization energies are 10–30% too large.
- **Band gaps:** LDA underestimates semiconductor band gaps by 30–100%.
- **Strongly correlated systems:** LDA fails for Mott insulators (e.g., FeO, NiO).

- **Dispersion:** LDA misses long-range van der Waals interactions.
- **Self-interaction error:** LDA violates Eq. (5.10).
- **Negative ions:** LDA sometimes fails to bind excess electrons in anions.

5.4 Generalized Gradient Approximation (GGA)

The LDA uses only the local density. The natural next step is to include the gradient ∇n to capture the spatial variation of the density.

5.4.1 Gradient expansion approximation (GEA)

The first attempt at gradient corrections was the systematic gradient expansion of the exchange-correlation energy:

$$E_{xc}^{GEA}[n] = \int n\, \epsilon_{xc}^{\text{HEG}}(n)\, d\mathbf{r} + \int C_{xc}(n) \frac{|\nabla n|^2}{n^{4/3}}\, d\mathbf{r} + \cdots \tag{5.17}$$

derived from a density-matrix gradient expansion. However, the GEA performs *worse* than LDA for many systems, because it violates the sum rule on the XC hole.

5.4.2 The GGA form and the reduced gradient

The GGA remedies the failures of the GEA by ensuring that the XC hole satisfies important sum rules. The general form is:

$$E_{xc}^{GGA}[n] = \int n(\mathbf{r})\, \epsilon_{xc}^{\text{HEG}}(n(\mathbf{r}))\, F_{xc}(n, s)\, d\mathbf{r} \tag{5.18}$$

where F_{xc} is an enhancement factor and s is the *reduced (dimensionless) density gradient*:

$$s = \frac{|\nabla n|}{2(3\pi^2)^{1/3} n^{4/3}} = \frac{|\nabla n|}{2k_F n} \tag{5.19}$$

s measures the local inhomogeneity in units of the local Fermi wavevector. It is small in metallic regions (nearly uniform density) and large at nuclear cusps, tails, and bond midpoints.

5.4.3 The PBE functional

The Perdew–Burke–Ernzerhof (PBE, 1996) functional is the most widely used GGA. It is derived non-empirically by satisfying known exact constraints. The exchange part:

$$E_x^{PBE} = \int n\, \epsilon_x^{\text{HEG}}(n)\, F_x^{PBE}(s)\, d\mathbf{r} \tag{5.20}$$

with the enhancement factor:

$$F_x^{PBE}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \tag{5.21}$$

where $\kappa = 0.804$ (from the Lieb–Oxford bound) and $\mu = 0.21951$ (from the gradient expansion of exchange). For the exchange, $F_x^{PBE}(0) = 1$ (reproduces LDA at $s = 0$) and $F_x^{PBE} \to 1 + \kappa$ as $s \to \infty$.

The PBE correlation:

$$E_c^{PBE} = \int n \left[\epsilon_c^{HEG}(r_s, \zeta) + H(r_s, \zeta, t)\right] d\mathbf{r} \tag{5.22}$$

where $\zeta = (n_\uparrow - n_\downarrow)/n$ is the relative spin polarization, $t = |\nabla n|/(2\phi k_s n)$ is another dimensionless gradient (with $k_s = \sqrt{4k_F/\pi}$ the Thomas–Fermi screening wavevector and ϕ a spin-scaling factor), and:

$$H = \frac{e^2}{a_0}\gamma\phi^3 \ln\left\{1 + \frac{\beta}{\gamma}\frac{t^2(1+At^2)}{1+At^2+A^2t^4}\right\} \tag{5.23}$$

with $A = (\beta/\gamma)[e^{-\epsilon_c^{HEG}/(\gamma\phi^3)} - 1]^{-1}$, $\beta \approx 0.066725$, and $\gamma \approx 0.031091$.

Exact constraints satisfied by PBE:

(i) Correct uniform gas limit ($s \to 0$, $t \to 0$).
(ii) Correct slowly-varying limit (recovers GEA to leading order).
(iii) Spin-scaling relation Eq. (4.24).
(iv) Lieb–Oxford bound on exchange.
(v) Correct $H \to -\epsilon_c^{HEG}$ as $t \to \infty$ (rapidly varying limit cancels correlation).
(vi) Correct high-density (scaling) limit: $E_c \to$ const as $\lambda \to \infty$.

5.4.4 The BLYP functional

The Becke–Lee–Yang–Parr (BLYP) functional combines:

- **Becke88 exchange (B88):** $F_x^{B88}(s) = 1 + \delta\beta s^2/(1 + 6\beta s \sinh^{-1} s)$ with $\beta = 0.0042$, fitted to exact exchange energies of noble gas atoms.
- **LYP correlation:** A non-local correlation functional derived from the Colle-Salvetti correlation energy formula, expressed in terms of n, ∇n, and $\nabla^2 n$.

BLYP is highly accurate for thermochemistry of organic molecules (mean absolute error ~ 3 kcal/mol) but is empirically parametrized, lacking the non-empirical grounding of PBE.

5.4.5 PBEsol

The PBEsol functional modifies PBE by restoring the GEA for exchange ($\mu_{\text{sol}} = 10/81$ from the second-order gradient expansion, replacing $\mu_{\text{PBE}} = 0.21951$) and a revised correlation, improving description of solids and surfaces while sacrificing some accuracy for molecules.

5.5 Meta-GGA Functionals

Meta-GGAs add to the GGA ingredients the *kinetic energy density*:

$$\tau(\mathbf{r}) = \frac{1}{2}\sum_i^{\text{occ}} |\nabla\phi_i(\mathbf{r})|^2 \tag{5.24}$$

or the Laplacian of the density $\nabla^2 n$. The kinetic energy density τ carries additional information about the orbital structure not contained in n alone.

5.5.1 Physical significance of τ

The ratio:

$$\alpha(\mathbf{r}) = \frac{\tau(\mathbf{r}) - \tau_W(\mathbf{r})}{\tau^{HEG}(\mathbf{r})} \tag{5.25}$$

where $\tau_W = |\nabla n|^2/(8n)$ is the von Weizsäcker kinetic energy density and $\tau^{HEG} = \frac{3}{10}(3\pi^2)^{2/3}n^{5/3}$ is the HEG value, provides a local measure of correlation strength:

- $\alpha \approx 0$: single-bond region (one orbital dominates, like in H_2 at equilibrium)
- $\alpha \approx 1$: slowly varying (metallic, free-electron-like) region
- $\alpha \gg 1$: weakly overlapping atomic densities (van der Waals region)

This discrimination capability makes meta-GGAs significantly more flexible than GGAs.

5.5.2 The SCAN functional

The Strongly Constrained and Appropriately Normed (SCAN) functional by Sun, Ruzsinszky, and Perdew (2015) is currently the best non-empirical meta-GGA. It satisfies all 17 known exact constraints that a meta-GGA can satisfy, organized around the isoelectronic neon series, the jellium surface, and the uniform gas.

The SCAN exchange uses an interpolation between the single-orbital ($\alpha \approx 0$) and slowly-varying ($\alpha \approx 1$) regimes:

$$F_x^{SCAN}(s,\alpha) = h_x^1(s) + f_x(\alpha)\left[h_x^0(s) - h_x^1(s)\right] \tag{5.26}$$

where h_x^0 is the exchange enhancement for $\alpha = 0$ (single orbital), h_x^1 is for $\alpha = 1$ (slowly varying), and $f_x(\alpha)$ is an interpolation function. The specific forms of h_x^0, h_x^1, and f_x are chosen to satisfy the appropriate constraints.

Key advantages of SCAN over PBE:

- Improved lattice constants and bulk moduli of solids.
- Better hydrogen-bond geometries in water.
- Improved description of both covalent and non-covalent interactions.
- Better band gaps (though still underestimated).

5.5.3 r^{2}SCAN

The regularized-restored SCAN (r^{2}SCAN, 2020) by Furness et al. addresses numerical instabilities in SCAN arising from its construction near $\alpha = 0$. r^{2}SCAN introduces regularization parameters η and τ_r to smooth the functional while preserving almost all exact constraints. It is now the recommended meta-GGA for production calculations.

5.6 Hybrid Functionals and the Adiabatic Connection

5.6.1 The adiabatic connection

A rigorous justification for mixing exact exchange comes from the *adiabatic connection* formalism. Consider a family of Hamiltonians parameterized by $\lambda \in [0,1]$:

$$\hat{H}_\lambda = \hat{T} + \lambda \hat{V}_{ee} + \hat{V}_{ext}^\lambda \tag{5.27}$$

where $\hat{V}_{ext}^\lambda$ is chosen so that the ground-state density is $n_0(\mathbf{r})$ for all λ. At $\lambda = 0$, this is the KS non-interacting system; at $\lambda = 1$, it is the fully interacting system.

The exchange-correlation energy is exactly:

$$\boxed{E_{xc}[n] = \int_0^1 U_{xc}^\lambda[n]\, d\lambda} \tag{5.28}$$

where $U_{xc}^\lambda = \langle \Psi_\lambda | \hat{V}_{ee} | \Psi_\lambda \rangle - J[n]$ is the potential energy contribution to E_{xc} at coupling strength λ.

At $\lambda = 0$: $U_{xc}^{\lambda=0} = E_x[n]$ (exact exchange of the KS Slater determinant). At $\lambda = 1$: $U_{xc}^{\lambda=1} = V_{ee}[n] - J[n]$ (the full non-classical *e*-*e* interaction).

5.6.2 Half-and-half mixing

The simplest model for the adiabatic connection integrand: if U_{xc}^λ is linear in λ (connecting E_x at $\lambda = 0$ to $2E_{xc}^{LDA} - E_x^{LDA}$ at $\lambda = 1$), integration gives the Becke half-and-half functional (1993):

$$E_{xc}^{H\&H} = \frac{1}{2}E_x^{HF} + \frac{1}{2}E_{xc}^{LDA} \tag{5.29}$$

5.6.3 The B3LYP functional

The three-parameter Becke functional (B3LYP) is the most widely used hybrid in quantum chemistry:

$$E_{xc}^{B3LYP} = E_{xc}^{LDA} + a_0(E_x^{HF} - E_x^{LDA}) + a_x(E_x^{B88} - E_x^{LDA}) + a_c(E_c^{LYP} - E_c^{VWN}) \tag{5.30}$$

with empirically fitted parameters $a_0 = 0.20$, $a_x = 0.72$, $a_c = 0.81$. B3LYP gives chemical accuracy ($\lesssim 3$ kcal/mol) for thermochemistry of main-group organic molecules.

5.6.4 PBE0

The non-empirical PBE0 hybrid uses a mixing of 25% exact exchange motivated by the adiabatic connection and perturbation theory (Adamo and Barone, 1999):

$$E_{xc}^{PBE0} = \frac{1}{4}E_x^{HF} + \frac{3}{4}E_x^{PBE} + E_c^{PBE} \tag{5.31}$$

The 25% mixing fraction can be justified by noting that exact exchange in the high-density limit contributes a fraction $\sim 1/4$ from second-order perturbation theory (Görling–Levy theory).

5.6.5 Range-separated hybrids

Standard hybrids use a fixed fraction of exact exchange at all interelectronic distances. Range-separated hybrids split the Coulomb interaction into short-range (SR) and long-range (LR) parts using the error function:

$$\frac{1}{r_{12}} = \underbrace{\frac{\operatorname{erfc}(\omega r_{12})}{r_{12}}}_{\text{short-range}} + \underbrace{\frac{\operatorname{erf}(\omega r_{12})}{r_{12}}}_{\text{long-range}} \tag{5.32}$$

where ω is the range-separation parameter (typically 0.2–0.5 a_0^{-1}).

HSE (Heyd–Scuseria–Ernzerhof): Uses exact exchange in the short range only:

$$E_{xc}^{HSE} = \frac{1}{4}E_x^{HF,SR}(\omega) + \frac{3}{4}E_x^{PBE,SR}(\omega) + E_x^{PBE,LR}(\omega) + E_c^{PBE} \tag{5.33}$$

The truncation of exact exchange at long range makes HSE computationally cheaper than PBE0 for periodic systems (where the HF exchange has slow convergence with **k**-point sampling).

CAM-B3LYP, ωB97X-D: Long-range corrected hybrids that increase the fraction of exact exchange at large r_{12}. These are essential for describing charge-transfer excitations in TD-DFT (Chapter 11).

5.6.6 Computational cost of hybrid functionals

The Fock exchange requires computing the exchange integrals:

$$(ij|kl) = \int\int \frac{\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\phi_k^*(\mathbf{r}')\phi_l(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' \tag{5.34}$$

The formal scaling of the Fock exchange is $O(N^4)$ with system size, compared to $O(N^3)$ for local and semi-local functionals. Various techniques (RI/density fitting, resolution-of-identity) reduce this to effectively $O(N^3)$ with a larger prefactor.

5.7 Double Hybrids

Double hybrids (Grimme, 2006) go one step higher by adding a fraction of MP2 correlation:

$$E_{xc}^{DH} = (1 - a_x)E_x^{DFA} + a_x E_x^{HF} + (1 - a_c)E_c^{DFA} + a_c E_c^{MP2} \tag{5.35}$$

The MP2 correlation energy:

$$E_c^{MP2} = -\frac{1}{4}\sum_{ijab} \frac{|(ia|jb) - (ib|ja)|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{5.36}$$

The B2-PLYP functional ($a_x = 0.53$, $a_c = 0.27$) achieves near-CCSD(T) accuracy for thermochemistry. The price is $O(N^5)$ scaling.

Chapter 6

The Exchange-Correlation Hole

6.1 The Pair Density and the XC Hole

The exchange-correlation energy has a deep physical interpretation in terms of the *exchange-correlation hole*. Understanding the XC hole is crucial for developing intuition about what makes a good functional.

Recall the pair density $n^{(2)}(\mathbf{r}, \mathbf{r}')$ from Eq. (1.21). For an uncorrelated system, the pair density would factorize: $n^{(2)}(\mathbf{r}, \mathbf{r}') \to n(\mathbf{r})n(\mathbf{r}')$ as correlations are switched off. The deviation from this uncorrelated form defines the pair correlation function $g(\mathbf{r}, \mathbf{r}')$:

$$n^{(2)}(\mathbf{r}, \mathbf{r}') = n(\mathbf{r})\, n(\mathbf{r}')\, g(\mathbf{r}, \mathbf{r}') \tag{6.1}$$

The XC Hole

The **exchange-correlation hole** $h_{xc}(\mathbf{r}, \mathbf{r}')$ is defined through:

$$n^{(2)}(\mathbf{r}, \mathbf{r}') = n(\mathbf{r})[n(\mathbf{r}') + h_{xc}(\mathbf{r}, \mathbf{r}')] \tag{6.2}$$

That is:

$$h_{xc}(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}')[g(\mathbf{r}, \mathbf{r}') - 1] \tag{6.3}$$

Physically, $h_{xc}(\mathbf{r}, \mathbf{r}')$ is the depletion of electron density at position $\mathbf{r}'$ due to the presence of an electron at $\mathbf{r}$. The hole represents the reduction in electron density around any given electron due to (i) quantum statistics (Pauli exclusion → exchange hole) and (ii) Coulomb repulsion (→ correlation hole).

6.2 Exact Sum Rules on the XC Hole

The XC hole satisfies several exact, important constraints.

6.2.1 Normalization sum rule

Integrating Eq. (1.22) gives:

$$\boxed{\int h_{xc}(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' = -1 \quad \text{for all } \mathbf{r}} \tag{6.4}$$

The XC hole always integrates to exactly -1. This is the most important constraint: the electron has excluded exactly one electron (itself, plus correlation effects) from its surroundings.

6.2.2 Separation into exchange and correlation holes

Write $h_{xc} = h_x + h_c$, where:

$$h_x(\mathbf{r}, \mathbf{r}') = -\frac{|\gamma(\mathbf{r}, \mathbf{r}')|^2}{2n(\mathbf{r})} \tag{6.5}$$

$$h_c(\mathbf{r}, \mathbf{r}') = h_{xc}(\mathbf{r}, \mathbf{r}') - h_x(\mathbf{r}, \mathbf{r}') \tag{6.6}$$

The exchange hole is the Fermi hole from antisymmetry:

$$\int h_x(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' = -1 \quad \Rightarrow \quad \int h_c(\mathbf{r}, \mathbf{r}')\, d\mathbf{r}' = 0 \tag{6.7}$$

The correlation hole integrates to zero: correlation redistributes electrons but does not change their total count.

6.2.3 Non-positivity of the exchange hole

$$h_x(\mathbf{r}, \mathbf{r}') \leq 0 \quad \text{for all } \mathbf{r}, \mathbf{r}' \tag{6.8}$$

This follows directly from Eq. (6.5) since $h_x \propto -|\gamma|^2 \leq 0$.

6.2.4 On-top values

At $\mathbf{r}' = \mathbf{r}$ (the "on-top" hole):

$$h_x(\mathbf{r}, \mathbf{r}) = -n(\mathbf{r})/2 \tag{6.9}$$

for a spin-unpolarized system. The exchange hole at the electron's own position equals $-n/2$ (half the density, because the exchange hole acts on one spin at a time).

6.3 E_{xc} as the Coulomb Interaction with the XC Hole

The exchange-correlation energy has a beautiful physical interpretation:

$$\boxed{E_{xc}[n] = \frac{1}{2} \int n(\mathbf{r}) \left[\int \frac{h_{xc}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right] d\mathbf{r}} \tag{6.10}$$

This states that E_{xc} is the Coulomb interaction energy between an electron at $\mathbf{r}$ and its surrounding XC hole $h_{xc}(\mathbf{r}, \mathbf{r}')$. The sum rule (Eq. 6.4) says the hole contains -1 electron, so E_{xc} is the electrostatic interaction of an electron with the -1 charge of its own hole.

6.3.1 The coupling-constant averaged hole

The adiabatic connection formula (Eq. 5.28) can be rewritten in terms of the coupling-constant averaged hole:

$$E_{xc}[n] = \frac{1}{2} \int n(\mathbf{r}) \left[\int \frac{\bar{h}_{xc}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \right] d\mathbf{r} \tag{6.11}$$

where $\bar{h}_{xc}(\mathbf{r}, \mathbf{r}') = \int_0^1 h_{xc}^{\lambda}(\mathbf{r}, \mathbf{r}')\, d\lambda$ is the coupling-constant averaged XC hole.

This makes explicit why the LDA works well despite its seemingly crude approximation: the XC energy depends not on the hole itself, but on its *system-averaged, coupling-constant-averaged, spherical average*. Many errors in the shape of the LDA hole cancel when integrated over angles and coupling constants.

6.3.2 Spherical averaging

Define the system-averaged hole:

$$\bar{h}_{xc}(u) = \frac{1}{N} \int n(\mathbf{r}) \frac{1}{4\pi} \int \bar{h}_{xc}(\mathbf{r}, \mathbf{r} + \mathbf{u}) \, d\hat{\mathbf{u}} \, d\mathbf{r} \tag{6.12}$$

The XC energy is then:

$$E_{xc}[n] = N \cdot 2\pi \int_0^\infty u \, \bar{h}_{xc}(u) \, du \tag{6.13}$$

This is remarkable: E_{xc} depends only on the one-dimensional function $\bar{h}_{xc}(u)$, not on the full three-dimensional hole. This explains why relatively simple functionals that correctly describe the spherically averaged, coupling-constant-averaged hole can be accurate.

6.4 The LDA Hole

The LDA replaces the true XC hole with that of the HEG at the local density. The LDA exchange hole is known analytically:

$$h_x^{LDA}(\mathbf{r}, u) = -\frac{9n(\mathbf{r})}{2} \left[\frac{j_1(k_F u)}{k_F u} \right]^2 \tag{6.14}$$

where j_1 is the spherical Bessel function of order 1, $k_F = (3\pi^2 n)^{1/3}$. This hole:

- Is spherically symmetric around the reference electron (an artifact of the HEG).
- Satisfies $h_x^{LDA}(r, u = 0) = -n/2$ (correct on-top value).
- Satisfies the normalization sum rule $\int h_x^{LDA} \, d^3u = -1$ (exact).
- Is always non-positive.

The LDA XC hole is too deep and too short-ranged for molecules, but its spherical average is remarkably accurate, explaining LDA's unexpected success.

Chapter 7

Janak's Theorem, Derivative Discontinuity, and the Band Gap Problem

7.1 Janak's Theorem

In the KS formalism with integer occupation numbers, the total energy changes discretely as electrons are added or removed. Janak (1978) extended the formalism to allow *fractional* occupation numbers $f_i \in [0, 1]$ and proved an exact theorem.

Janak's Theorem

For the total KS energy $E(\{f_i\})$ with fractional occupations, the partial derivative with respect to the i-th occupation number equals the i-th KS eigenvalue:

$$\frac{\partial E}{\partial f_i} = \epsilon_i[\{f_j\}] \tag{7.1}$$

This holds for any fixed external potential.

Proof. The total energy is:

$$E = \sum_j f_j \langle \phi_j | -\frac{1}{2}\nabla^2 | \phi_j \rangle + \int v_{\text{ext}} n d\mathbf{r} + J[n] + E_{xc}[n] \tag{7.2}$$

Taking $\partial/\partial f_i$ at fixed v_{ext}, and using the KS equations to simplify $\partial E/\partial \phi_j^* \cdot \partial \phi_j^*/\partial f_i = 0$ (since each ϕ_j satisfies its own equation), one obtains:

$$\begin{aligned}\frac{\partial E}{\partial f_i} &= \langle \phi_i | -\frac{1}{2}\nabla^2 | \phi_i \rangle + \int [v_{\text{ext}} + v_H + v_{\text{xc}}] \frac{\partial n}{\partial f_i} d\mathbf{r} \\ &= \langle \phi_i | -\frac{1}{2}\nabla^2 | \phi_i \rangle + \int v_{\text{eff}} |\phi_i|^2 d\mathbf{r} = \langle \phi_i | \hat{h}_{KS} | \phi_i \rangle = \epsilon_i \quad \square\end{aligned} \tag{7.3}$$

□

7.1.1 Implications of Janak's theorem

Integrating Janak's theorem from $f_i = 0$ to $f_i = 1$:

$$E(f_i = 1) - E(f_i = 0) = \int_0^1 \epsilon_i(f)\, df \approx \epsilon_i(f_i = 1/2) \tag{7.4}$$

This gives a DFT analog of Koopmans' theorem: the orbital energy at half-occupation is approximately equal to the ionization energy or electron affinity. This is the basis of the Slater transition state method for excitation energies.

7.2 Fractional Electron Number and Piecewise Linearity

7.2.1 The grand canonical ensemble

For a system in contact with an electron reservoir, the ground-state energy as a function of electron number N (which can be non-integer in the grand canonical sense) satisfies:

$$E(N) = (1-f)\,E(\lfloor N\rfloor) + f\,E(\lfloor N\rfloor + 1), \quad f = N - \lfloor N\rfloor \tag{7.5}$$

That is, $E(N)$ is *piecewise linear* in N between integers.

Piecewise Linearity (Perdew, Parr, Levy, Balduz, 1982)

The exact ground-state energy of a system coupled to an electron reservoir is piecewise linear in the electron number, with a slope discontinuity at integer N:

$$\left.\frac{\partial E}{\partial N}\right|_{N^-} = -I_P, \qquad \left.\frac{\partial E}{\partial N}\right|_{N^+} = -A \tag{7.6}$$

where $I_P = E(N-1) - E(N)$ is the ionization potential and $A = E(N) - E(N+1)$ is the electron affinity.

7.2.2 The fundamental gap

The *fundamental gap* E_g is:

$$E_g = I_P - A = [E(N-1)-E(N)] - [E(N)-E(N+1)] = E(N+1)+E(N-1)-2E(N) \tag{7.7}$$

This is also the curvature of $E(N)$: it measures how much $E(N)$ deviates from linearity between the integers $N-1$, N, and $N+1$.

7.3 The Derivative Discontinuity of v_{xc}

The exact $v_{\text{xc}}(\mathbf{r}) = \delta E_{xc}/\delta n$ undergoes a rigid upward shift as N crosses an integer:

$$v_{\text{xc}}^{+}(\mathbf{r}) = v_{\text{xc}}^{-}(\mathbf{r}) + \Delta_{xc} \tag{7.8}$$

where the superscripts ± denote limits from above and below integer N, and Δ_{xc} is a spatially constant (system-wide) positive shift in the XC potential. Sham and Schlüter (1983) and Perdew and Levy (1983) showed:

$$E_g = \Delta_{KS} + \Delta_{xc} \tag{7.9}$$

where:

$$\Delta_{KS} = \epsilon_{LUMO} - \epsilon_{HOMO} \tag{7.10}$$

is the KS eigenvalue gap (HOMO-LUMO gap), and Δ_{xc} is the derivative discontinuity of E_{xc}.

7.4 The DFT Band Gap Problem

Most approximate exchange-correlation functionals (LDA, GGA, meta-GGA) are *continuous* as a function of N: their v_{xc} has no derivative discontinuity ($\Delta_{xc}^{approx} = 0$). Therefore:

$$E_g^{approx} = \Delta_{KS}^{approx} < E_g^{exact} = \Delta_{KS}^{exact} + \Delta_{xc} \tag{7.11}$$

The missing Δ_{xc} accounts for a substantial fraction of the band gap underestimation:

- For Si: $\Delta_{KS}^{LDA} \approx 0.52$ eV, $E_g^{exp} = 1.17$ eV, so $\Delta_{xc} \approx 0.65$ eV.
- For GaAs: $\Delta_{KS}^{LDA} \approx 0.42$ eV, $E_g^{exp} = 1.52$ eV.
- For Ge: $\Delta_{KS}^{LDA} < 0$ eV (predicted metallic!), $E_g^{exp} = 0.74$ eV.

7.4.1 Remedies for the band gap problem

Several approaches can recover the correct band gap:

(i) **Hybrid functionals:** Exact exchange provides a partial Δ_{xc}. HSE and PBE0 improve gaps substantially. However, they are still too small for wide-gap insulators.

(ii) **Optimally tuned range-separated hybrids:** Tune ω such that the HOMO eigenvalue equals $-I_P$ (ionization potential tuning), restoring piecewise linearity.

(iii) **DFT+U:** Adds a Hubbard-U term for localized states (Chapter 8).

(iv) **Many-body perturbation theory (GW):** Computes the quasiparticle gap using the GW self-energy, including the full frequency-dependent response (Chapter 12).

(v) **Scissor correction:** Simply shift the conduction bands by a constant; this is not predictive but useful for post-processing optical spectra.

7.5 Delocalization Error and Strong Correlation

Closely related to the derivative discontinuity is the *delocalization error* of approximate functionals. Because LDA and GGA give a smooth, concave $E(N)$ curve (rather than piecewise linear), they over-delocalize electrons. Consequences include:

- Over-delocalization of electrons in symmetric charge-transfer complexes (e.g., H_2^+ at stretched geometry).
- Systematic errors in reaction barrier heights (too low by $\sim$ 5–10 kcal/mol).
- Incorrect prediction of metallic behavior for Mott insulators.

Conversely, some hybrid functionals over-localize due to excessive exact exchange—this is the *static correlation error*. The interplay between delocalization error (favoring delocalized, metallic states) and static correlation error (favoring over-localized, spin-symmetry-broken states) is a central challenge in functional development.

Chapter 8

Self-Interaction Error and Strong Correlation: DFT+U and Beyond

8.1 The Self-Interaction Problem

8.1.1 Definition

For a single electron in orbital ϕ_1, the exact energy satisfies:

$$J[n_1] + E_{xc}[n_1] = 0, \quad n_1 = |\phi_1|^2 \tag{8.1}$$

The Hartree energy $J[n_1]$ is the spurious Coulomb repulsion of the electron with its own density. For the exact functional, E_{xc} exactly cancels this. For approximate functionals (LDA, GGA), this cancellation is incomplete, leaving a residual *self-interaction error* (SIE):

$$\epsilon^i_{SIE} = J[n_i] + E_{xc}[n_i] \neq 0 \tag{8.2}$$

8.1.2 Physical consequences

- **Orbital over-delocalization:** Each electron sees a spurious repulsion from itself, effectively lowering the energy of delocalized states relative to localized ones. Electrons spread out more than they should.
- **Wrong asymptotic potential:** The true effective potential $v_{\text{eff}} \to -1/r$ as $r \to \infty$ (the electron at large distance sees $N-1$ electrons). The LDA/GGA potential decays as e^{-r} because the SIE causes $v_H + v_{\text{xc}}$ to partly cancel, but the cancellation is imperfect, leaving a potential that decays too quickly. This affects ionization energies, Rydberg states, and electron affinities.
- **Band gap underestimation:** A deeper reason for the band gap problem.
- **Barriers and charge transfer:** Too-low reaction barriers, wrong charge distribution in donor-acceptor complexes.

8.1.3 Perdew-Zunger Self-Interaction Correction (PZ-SIC)

Perdew and Zunger (1981) proposed correcting the SIE orbital-by-orbital:

$$E_{xc}^{PZ-SIC} = E_{xc}^{DFA}[n_\uparrow, n_\downarrow] - \sum_{i\sigma} \left[J[n_{i\sigma}] + E_{xc}^{DFA}[n_{i\sigma}, 0]\right] \tag{8.3}$$

The sum over $i\sigma$ subtracts the spurious self-interaction of each orbital. PZ-SIC satisfies Eq. (8.1) by construction. However:

- PZ-SIC is not unitary invariant: the result depends on the choice of localized orbitals. One must minimize over all unitary rotations of the occupied space.
- PZ-SIC improves ionization energies, band gaps, and charge-transfer properties but can degrade bond lengths and atomization energies.

8.2 Strongly Correlated Systems: Mott Insulators

8.2.1 The Mott insulator problem

Transition metal oxides such as FeO, CoO, NiO, and MnO are experimentally large-gap insulators ($E_g \sim$ 3–7 eV), yet LDA and GGA predict them to be metals or small-gap semiconductors. The reason is that LDA/GGA cannot describe the strong Coulomb correlations between partially filled d or f shells that localize electrons and open the gap.

In a Mott insulator, the on-site Coulomb repulsion U (the energy cost of having two electrons on the same site) exceeds the bandwidth W. When $U \gg W$, electron hopping is suppressed and the system is insulating. LDA underestimates U and overestimates W (due to SIE), incorrectly predicting delocalized metallic behavior.

8.2.2 The Hubbard model

The essential physics is captured by the Hubbard model:

$$\hat{H}_{Hub} = -t \sum_{\langle ij \rangle \sigma} (c^\dagger_{i\sigma} c_{j\sigma} + \text{h.c.}) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \tag{8.4}$$

The bandwidth is $W \sim 2zt$ (where z is coordination number), and the insulating state appears when $U/W \gtrsim 1$ (Mott transition).

8.2.3 DFT+U

The DFT+U method (Anisimov, Zaanen, Andersen, 1991; Dudarev et al., 1998) adds a penalty functional to DFT that explicitly corrects the treatment of localized d/f orbitals.

The Dudarev simplified form (rotationally invariant DFT+U):

$$\boxed{E_{DFT+U} = E_{DFT} + \frac{U_{eff}}{2} \sum_{I,\sigma} \text{Tr}\left[\hat{n}^{I\sigma}(1 - \hat{n}^{I\sigma})\right]} \tag{8.5}$$

where $U_{eff} = U - J$ (combining the Hubbard U and Hund's exchange J), and $\hat{n}^{I\sigma}_{mm'} = \sum_k f^\sigma_k \langle \phi^\sigma_k | \hat{P}^{Im} | \phi^\sigma_k \rangle$ is the occupation matrix of the d/f subspace at site I.

Physical interpretation

The added term penalizes fractional occupations: $\text{Tr}[\hat{n}(1 - \hat{n})] = \sum_m n_m(1 - n_m)$ is maximized when $n_m = 1/2$ (fully fractional) and minimized (zero) when $n_m \in \{0, 1\}$. The penalty drives occupations toward integer values, restoring the piecewise linearity of $E(N)$ for the correlated subspace.

The DFT+U effective potential acting on the d/f subspace is:

$$V^{I\sigma}_{mm'} = \frac{\delta E_{DFT+U}}{\delta n^{I\sigma}_{m'm}} = \frac{U_{eff}}{2}(\delta_{mm'} - 2n^{I\sigma}_{mm'}) \tag{8.6}$$

This shifts the energy of *occupied* d/f orbitals (with $n_m \to 1$) down by $\sim -U_{eff}/2$ and *unoccupied* orbitals (with $n_m \to 0$) up by $\sim +U_{eff}/2$, opening a gap of approximately U_{eff}.

Choice of U

The value of U_{eff} is not uniquely determined. Approaches include:

(i) Fitting to experimental properties (gap, magnetic moment).
(ii) Linear-response theory (Cococcioni and de Gironcoli, 2005): compute U from the response of the system to a small perturbation of the occupation matrix.
(iii) Constrained-RPA (cRPA): compute the screened Coulomb interaction in the low-energy subspace.
(iv) ACBN0 (pseudo-hybrid): compute U from the occupations themselves.

8.2.4 DFT+U+J

The full Liechtenstein form retains U and J separately:

$$E_{DFT+U+J} = E_{DFT} + \frac{1}{2} \sum_{I,\sigma\neq\sigma'} \sum_{mm'} \left[U \, n^{I\sigma}_{mm} \, n^{I\sigma'}_{m'm'} - J \, n^{I\sigma}_{mm'} \, n^{I\sigma'}_{mm'} \right] \tag{8.7}$$

The J term corrects Hund's exchange, improving magnetic moments in addition to the gap.

8.3 DFT+DMFT

For systems where DFT+U is insufficient (materials near the Mott transition, heavy-fermion systems, actinides), Dynamical Mean-Field Theory (DMFT) can be combined with DFT. The DFT+DMFT approach:

1. Performs a DFT calculation for the crystal.
2. Constructs Wannier functions for the correlated subspace.
3. Solves the DMFT impurity problem (using, e.g., continuous-time quantum Monte Carlo) to obtain the local self-energy $\Sigma(\omega)$.
4. Updates the KS Green's function and iterates to self-consistency.

DMFT captures dynamical correlations (frequency-dependent self-energy), orbital polarization, and the Mott transition with unprecedented accuracy. It is computationally expensive ($O(N^3)$ per DMFT iteration, with many iterations needed) but currently the state-of-the-art for correlated materials.

Chapter 9

Dispersion Interactions and van der Waals Forces

9.1 The Physical Origin of Dispersion

Van der Waals (vdW) or London dispersion forces arise from instantaneous quantum mechanical fluctuations in the electron density. At large separation R between two neutral atoms A and B, the asymptotic interaction energy is:

$$E_{disp} \xrightarrow{R\to\infty} -\frac{C_6}{R^6} - \frac{C_8}{R^8} - \frac{C_{10}}{R^{10}} - \cdots \tag{9.1}$$

The C_6 coefficient is related to the dynamic polarizabilities $\alpha_A(i\omega)$ and $\alpha_B(i\omega)$ via the London/Casimir–Polder formula:

$$C_6^{AB} = \frac{3}{\pi}\int_0^\infty \alpha_A(i\omega)\alpha_B(i\omega)\,d\omega \tag{9.2}$$

This $-C_6/R^6$ interaction is a pure correlation effect: it vanishes in the Hartree–Fock or KS exchange approximation. More precisely, it arises from *long-range* electron-electron correlations between electrons on different fragments, which local and semi-local DFT functionals cannot capture because their exchange-correlation kernels are local in space.

9.2 Why LDA/GGA Fail for Dispersion

The XC energy in LDA and GGA can be written as:

$$E_{xc}^{LDA/GGA} = \int f(n(\mathbf{r}), |\nabla n(\mathbf{r})|)\,d\mathbf{r} \tag{9.3}$$

Such a functional cannot capture interactions between spatially separated, non-overlapping fragments, because when two atoms are far apart, their densities do not overlap and the functional reduces to the sum of atomic contributions. There is no term coupling atom A's density fluctuations to atom B's. Consequently, LDA and GGA give interaction energies that decay exponentially at large separations, while the true interaction is $-C_6/R^6$.

In practice:

- LDA significantly overbinds van der Waals dimers (due to exchange "glue" at short range mimicking attraction).
- GGA (PBE) correctly gives near-zero binding for rare-gas dimers but misses the $-C_6/R^6$ tail entirely.
- Without dispersion corrections, DFT cannot describe molecular crystals, stacked DNA bases, graphene intercalation, or protein folding.

9.3 Grimme's DFT-D Corrections

Grimme's dispersion corrections add a semi-empirical $-C_6/R^6$ term to the DFT energy.

9.3.1 DFT-D2 (Grimme, 2006)

$$E_{DFT-D2} = E_{DFT} - s_6 \sum_{i<j} \frac{C_6^{ij}}{R_{ij}^6} f_{damp}(R_{ij}) \tag{9.4}$$

where s_6 is a global scaling parameter (functional-dependent), $C_6^{ij} = \sqrt{C_6^i C_6^j}$ is the geometric mean of atomic C_6 coefficients (tabulated), and f_{damp} is a damping function that avoids double-counting at short range:

$$f_{damp}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/s_R R_0^{ij} - 1)}} \tag{9.5}$$

DFT-D2 improved description of π-stacking and hydrogen bonding, but its C_6 coefficients are fixed and do not respond to the chemical environment.

9.3.2 DFT-D3 (Grimme, Antony, Ehrlich, Krieg, 2010)

DFT-D3 introduced coordination-number-dependent C_6 coefficients, significantly improving transferability. The pair potential includes both C_6 and C_8 terms:

$$E_{D3} = -\sum_{i<j} \left[\frac{s_6 C_6^{ij}}{R_{ij}^6 + R_0^6} + \frac{s_8 C_8^{ij}}{R_{ij}^8 + R_0^8} \right] \tag{9.6}$$

The D3(BJ) variant (Becke-Johnson damping) uses a physically motivated damping to a finite value at $R \to 0$, rather than to zero, better representing the short-range behavior.

Three-body Axilrod-Teller-Muto (ATM) terms can be added:

$$E_{3-body} = s_9 \sum_{i<j<k} \frac{C_9^{ijk}(3\cos\theta_i \cos\theta_j \cos\theta_k + 1)}{(R_{ij} R_{jk} R_{ik})^3} \tag{9.7}$$

9.4 Tkatchenko–Scheffler (TS) and Many-Body Dispersion (MBD)

The Tkatchenko-Scheffler method (2009) computes C_6 coefficients from the Hirshfeld partitioned density:

$$C_6^{AB} = \frac{2C_{6,\text{free}}^A C_{6,\text{free}}^B}{\frac{\alpha_0^B}{\alpha_0^A} C_{6,\text{free}}^A + \frac{\alpha_0^A}{\alpha_0^B} C_{6,\text{free}}^B} \cdot \left(\frac{V_A^{Hirsh}}{V_A^{free}} \right)^2 \left(\frac{V_B^{Hirsh}}{V_B^{free}} \right)^2 \tag{9.8}$$

where V^{Hirsh} is the Hirshfeld volume. This makes C_6 dependent on the chemical environment.

The Many-Body Dispersion (MBD) method (Tkatchenko, DiStasio, Car, Scheffler, 2012) goes beyond pairwise additivity by treating the coupled quantum harmonic oscillator model for a collection of atoms, capturing three-body and higher-order dispersion effects exactly within the model. This is essential for extended systems such as molecular crystals, nanostructures, and biomolecules where cooperative effects are significant.

9.5 Non-Local Density Functionals: vdW-DF

Dion, Rydberg, Schroder, Langreth, and Lundqvist (2004) derived a fully non-local correlation functional from first principles:

$$E_c^{nl}[n] = \frac{1}{2} \iint n(\mathbf{r})\, \phi(\mathbf{r}, \mathbf{r}')\, n(\mathbf{r}')\, d\mathbf{r}\, d\mathbf{r}' \tag{9.9}$$

where the kernel $\phi(\mathbf{r}, \mathbf{r}')$ depends on $n(\mathbf{r})$, $n(\mathbf{r}')$, and $|\mathbf{r}-\mathbf{r}'|$. This kernel is derived from the local dielectric model and captures the physics of frequency-dependent polarizabilities. The vdW-DF functional:

$$E_{xc}^{vdW-DF} = E_x^{GGA} + E_c^{LDA} + E_c^{nl} \tag{9.10}$$

satisfies the $-C_6/R^6$ asymptotic behavior exactly within the model. Various improvements (vdW-DF2, optB86b-vdW, rev-vdW-DF2) have refined the exchange functional and kernel to improve accuracy.

A major practical advantage of non-local vdW functionals over DFT-D approaches is that they are derived from the density non-locally without empirical parameters for specific element pairs, making them more transferable.

9.6 The D3 and vdW-DF Compared

Property	**DFT-D3**	**vdW-DF**
Derivation	Semi-empirical	Non-empirical
Pairwise	Yes (+ 3-body)	No (truly non-local)
Computational cost	$O(N^2)$ pair sum	$O(N \log N)$ FFT
Asymptotic C_6	Correct by construction	Correct within model
Accuracy	Excellent for small molecules	Good for extended systems

For typical applications:

- **Molecular systems, benchmark datasets:** PBE-D3(BJ) or B3LYP-D3(BJ)
- **Solids, surfaces, layered materials:** PBE-D3 or rev-vdW-DF2
- **Biomolecules:** MBD or vdW-DF-cx

Chapter 10

Numerical Implementation of DFT

10.1 Overview of a DFT Calculation

A practical DFT calculation requires:

1. A representation of the KS orbitals $\{\phi_i\}$ in a basis set.
2. A method for computing the Hartree and XC potentials on a grid.
3. An algorithm for solving the generalized eigenvalue problem.
4. An SCF iteration scheme for convergence.
5. A treatment of the core electrons (all-electron or pseudopotential).

Different DFT codes make different choices at each step. Here we describe the main approaches.

10.2 Basis Sets

10.2.1 Plane waves

In periodic systems (crystals), Bloch's theorem dictates that the KS orbitals have the form $\phi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r})$, where $u_{n\mathbf{k}}$ is lattice-periodic. The lattice-periodic part can be expanded in a Fourier series:

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G})\, e^{i\mathbf{G}\cdot\mathbf{r}} \tag{10.1}$$

where $\mathbf{G}$ are reciprocal lattice vectors. The expansion is truncated at a kinetic energy cutoff:

$$\frac{1}{2}|\mathbf{k}+\mathbf{G}|^2 \leq E_{cut} \tag{10.2}$$

The number of plane waves is $N_{PW} \sim E_{cut}^{3/2}\Omega/\pi^2$ where Ω is the unit cell volume.

Advantages: Systematic convergence (increase E_{cut}), no basis-set superposition error (BSSE), simple implementation of forces via Hellmann-Feynman theorem, exact treatment of periodic boundary conditions.

Disadvantages: Require large E_{cut} for all-electron calculations (core oscillations need many plane waves); typically used with pseudopotentials. Vacuum regions needed for non-periodic systems (supercell approach).

10.2.2 Localized (Gaussian) basis sets

In molecular quantum chemistry, orbitals are expanded in atom-centered Gaussian functions:

$$\chi_\mu(\mathbf{r}) = N_\mu\, Y_l^m(\hat{\mathbf{r}} - \hat{\mathbf{R}}_A)\, r^{n-1}\, e^{-\alpha_\mu|\mathbf{r}-\mathbf{R}_A|^2} \tag{10.3}$$

where Y_l^m are spherical harmonics and α_μ is the Gaussian exponent. The KS orbitals are expanded as:

$$\phi_i(\mathbf{r}) = \sum_\mu c_{\mu i} \chi_\mu(\mathbf{r}) \tag{10.4}$$

This is the LCAO (linear combination of atomic orbitals) approach. All integrals (overlap, kinetic, nuclear attraction, Coulomb) can be computed analytically for Gaussians.

Standard Gaussian basis sets: STO-3G (minimal), 6-31G* (split-valence with polarization), cc-pVDZ/TZ/QZ (Dunning correlation-consistent), def2-SVP/TZVP/QZVP (Ahlrichs), aug-cc-pVTZ (with diffuse functions for anions).

Advantages: Efficient for molecules, exact analytic integrals, small basis size for comparable accuracy.

Disadvantages: Basis-set superposition error (BSSE) requires counterpoise correction; incomplete basis requires extrapolation; less systematic convergence than plane waves.

10.2.3 Numerical atomic orbitals (NAO)

Codes such as FHI-aims and SIESTA use numerical radial functions:

$$\chi_\mu(\mathbf{r}) = \frac{u_{nl}(r)}{r} Y_l^m(\hat{\mathbf{r}}) \tag{10.5}$$

where $u_{nl}(r)$ is a numerical function confined to a sphere of radius r_{cut}. NAOs combine the localization of Gaussians with more flexibility to represent the true atomic orbitals, and can be tuned to system-independent "tight" or "light" settings.

10.2.4 Projector augmented wave (PAW) method

The PAW method (Blöchl, 1994) provides an all-electron result with plane-wave efficiency. It introduces a linear transformation $|\tilde{\phi}\rangle \to |\phi\rangle$ between smooth (tilde, pseudized) and true all-electron functions:

$$|\phi_i\rangle = |\tilde{\phi}_i\rangle + \sum_I \sum_a (|\phi_a^I\rangle - |\tilde{\phi}_a^I\rangle)\langle \tilde{p}_a^I | \tilde{\phi}_i\rangle \tag{10.6}$$

where $|\phi_a^I\rangle$ are all-electron partial waves, $|\tilde{\phi}_a^I\rangle$ are smooth partial waves, and $\langle \tilde{p}_a^I|$ are projectors centered on atom I. PAW recovers the exact all-electron wavefunction from a smooth pseudo-wavefunction, making it both accurate and efficient.

10.3 Pseudopotentials

For heavy elements, the KS equations must handle core electrons whose wavefunctions oscillate rapidly near the nucleus (to be orthogonal to each other). These oscillations require large E_{cut} in a plane-wave expansion. Pseudopotentials replace the true nuclear potential plus core electrons with an effective potential that reproduces the correct behavior of valence electrons outside a core radius r_c.

10.3.1 Norm-conserving pseudopotentials (NCPP)

Designed so that the pseudo-wavefunction $\tilde{\phi}$ equals the all-electron ϕ outside r_c and has the same norm inside:

$$\int_0^{r_c} |\tilde{\phi}(r)|^2 r^2 \, dr = \int_0^{r_c} |\phi(r)|^2 r^2 \, dr \tag{10.7}$$

The norm-conservation ensures correct scattering properties at the reference energy and good transferability. However, NCPPs require relatively high E_{cut} for first-row elements and transition metals.

10.3.2 Ultrasoft pseudopotentials (USPP)

Vanderbilt (1990) relaxed the norm-conservation condition by introducing augmentation charges:

$$n(\mathbf{r}) = |\tilde{\phi}(\mathbf{r})|^2 + \sum_I \sum_{mn} Q^I_{mn}(\mathbf{r}) \langle\tilde{\phi}|\tilde{p}^I_m\rangle\langle\tilde{p}^I_n|\tilde{\phi}\rangle \tag{10.8}$$

where Q^I_{mn} are augmentation functions that restore the correct all-electron density inside r_c. USPPs allow much lower E_{cut} (by a factor $\sim$ 4–8) at the cost of a more complex formalism (generalized eigenvalue problem instead of standard).

10.4 The Generalized Eigenvalue Problem

In a non-orthogonal basis (Gaussians, NAOs), the KS equations become a generalized eigenvalue problem:

$$\mathbf{HC} = \mathbf{SC}\,\boldsymbol{\epsilon} \tag{10.9}$$

where $H_{\mu\nu} = \langle\chi_\mu|\hat{h}_{KS}|\chi_\nu\rangle$, $S_{\mu\nu} = \langle\chi_\mu|\chi_\nu\rangle$, $C_{\mu i}$ are the expansion coefficients, and $\boldsymbol{\epsilon}$ are the KS eigenvalues. This is solved by Cholesky decomposition of $\mathbf{S}$, transformation to an orthogonal basis, and standard diagonalization.

For large systems ($N > 1000$ atoms), direct diagonalization (scaling as $O(N^3)$) becomes the bottleneck. Linear-scaling $O(N)$ DFT methods avoid diagonalization by working with the density matrix directly, using its exponential localization in insulators.

10.5 k-Point Sampling and the Brillouin Zone

For periodic systems, the total energy per unit cell requires integrating over the Brillouin zone (BZ):

$$E = \int_{BZ} \sum_n f_{n\mathbf{k}}\, \epsilon_{n\mathbf{k}} \frac{d^3k}{(2\pi)^3} \tag{10.10}$$

In practice, this integral is replaced by a discrete sum over special $\mathbf{k}$-points. The Monkhorst-Pack scheme generates a uniform mesh $\{k_1, k_2, k_3\} = \{n_1/N_1, n_2/N_2, n_3/N_3\}$ in reduced coordinates, which is the standard approach for insulators and semiconductors.

For metals, the Fermi surface requires special treatment (Methfessel-Paxton smearing, Marzari-Vanderbilt cold smearing, tetrahedron method) to handle the discontinuity in occupation numbers.

10.6 Forces, Stress, and Geometry Optimization

The Hellmann-Feynman theorem gives exact atomic forces from the DFT total energy without numerical differentiation:

$$\mathbf{F}_I = -\frac{\partial E}{\partial \mathbf{R}_I} = -\left\langle \Psi_0 \middle| \frac{\partial \hat{H}}{\partial \mathbf{R}_I} \middle| \Psi_0 \right\rangle \tag{10.11}$$

For plane-wave DFT, there are additional Pulay corrections when using a non-complete basis (but for a complete plane-wave basis the Pulay terms vanish). Forces are used in molecular dynamics and geometry optimization.

Geometry optimization proceeds by moving atoms along the negative gradient of E (steepest descent, BFGS quasi-Newton, conjugate gradient), until $|\mathbf{F}|$ is below a threshold (typically 0.01–0.05 eV/Å).

Chapter 11

Time-Dependent DFT (TD-DFT)

11.1 Motivation: Beyond Ground States

Ground-state DFT accesses only ground-state properties (energy, density, geometry, forces). Many important physical phenomena—optical absorption, photochemistry, fluorescence, electro-optical properties—require a description of *excited states*. Time-Dependent DFT (TD-DFT) extends the DFT framework to time-dependent external fields and provides access to excitation energies and optical spectra.

11.2 The Runge–Gross Theorem

The foundation of TD-DFT is the Runge–Gross theorem (1984), the time-dependent analog of the Hohenberg–Kohn theorem.

Runge–Gross Theorem (1984)

For a system of interacting electrons evolving from a fixed initial state Ψ_0 under a time-dependent external potential $v(\mathbf{r}, t)$, the time-dependent electron density $n(\mathbf{r}, t)$ uniquely determines $v(\mathbf{r}, t)$ (up to a spatially uniform function of time). Equivalently, there is a one-to-one correspondence:

$$v(\mathbf{r}, t) \overset{1:1}{\longleftrightarrow} n(\mathbf{r}, t) \tag{11.1}$$

given the initial state.

Proof sketch. Expand $v(\mathbf{r}, t)$ and $v'(\mathbf{r}, t)$ in Taylor series around t_0: $v(\mathbf{r}, t) = \sum_k v_k(\mathbf{r})(t-t_0)^k/k!$. Assume $v \neq v' + c(t)$, so there exists a first k_0 at which $v_{k_0}(\mathbf{r}) \neq v'_{k_0}(\mathbf{r}) + c$. From the equation of motion of the current density $\mathbf{j}$:

$$\frac{\partial \mathbf{j}}{\partial t} = -n\nabla(v - v') + \cdots \tag{11.2}$$

Using the continuity equation $\partial n/\partial t = -\nabla \cdot \mathbf{j}$, one shows:

$$\left.\frac{\partial^{k_0+1}(n - n')}{\partial t^{k_0+1}}\right|_{t_0} = -n_0\nabla^2(v_{k_0} - v'_{k_0}) + \cdots \tag{11.3}$$

If $\nabla^2(v_{k_0} - v'_{k_0}) \neq 0$, then $n \neq n'$, establishing the map. A regularity condition (potentials expandable in Taylor series) is required. □ □

The Runge-Gross theorem establishes that $n(\mathbf{r}, t)$ determines $v(\mathbf{r}, t)$ given the initial state, and hence determines the time-evolving wavefunction and all observables. The energy is no longer the central quantity; instead the density (or current density) plays the fundamental role.

11.3 The Time-Dependent Kohn–Sham Equations

By analogy with the ground-state case, the time-dependent density can be reproduced by a non-interacting KS system:

$$i\frac{\partial \phi_i(\mathbf{r},t)}{\partial t} = \left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r},t)\right]\phi_i(\mathbf{r},t) \tag{11.4}$$

with:

$$v_{\text{eff}}(\mathbf{r},t) = v_{\text{ext}}(\mathbf{r},t) + v_H[n](\mathbf{r},t) + v_{\text{xc}}[n](\mathbf{r},t) \tag{11.5}$$

The time-dependent XC potential $v_{\text{xc}}[n](\mathbf{r},t)$ is now a functional of the time-dependent density and the initial state. In the *adiabatic approximation*:

$$v_{\text{xc}}^{adiab}[n](\mathbf{r},t) = v_{\text{xc}}^{GS}[n(\cdot,t)](\mathbf{r}) \tag{11.6}$$

one uses the ground-state functional evaluated at the instantaneous density. This ignores memory effects (retardation in the XC response) and is the standard approximation used in almost all TD-DFT calculations.

11.4 Linear Response TD-DFT and the Casida Equation

For weak perturbations, one can work in the frequency domain using linear response theory. The density response $\delta n(\mathbf{r},\omega)$ to a perturbation $\delta v_{\text{ext}}(\mathbf{r},\omega)$ is:

$$\delta n(\mathbf{r},\omega) = \int \chi(\mathbf{r},\mathbf{r}';\omega)\,\delta v_{\text{ext}}(\mathbf{r}',\omega)\,d\mathbf{r}' \tag{11.7}$$

The exact density-density response function χ is related to the non-interacting KS response function χ_s via the Dyson equation:

$$\chi(\mathbf{r},\mathbf{r}';\omega) = \chi_s(\mathbf{r},\mathbf{r}';\omega) + \int\int \chi_s(\mathbf{r},\mathbf{r}_1;\omega)\left[\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} + f_{xc}(\mathbf{r}_1,\mathbf{r}_2;\omega)\right]\chi(\mathbf{r}_2,\mathbf{r}';\omega)\,d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.8}$$

where the XC kernel f_{xc} is:

$$f_{xc}(\mathbf{r},\mathbf{r}';\omega) = \frac{\delta v_{\text{xc}}[n](\mathbf{r},\omega)}{\delta n(\mathbf{r}',\omega)} \tag{11.9}$$

In the adiabatic approximation, f_{xc} is frequency-independent: $f_{xc}^{adiab}(\mathbf{r},\mathbf{r}') = \delta^2 E_{xc}/\delta n(\mathbf{r})\delta n(\mathbf{r}')$.

11.4.1 The Casida equations

Casida (1995) transformed the Dyson equation into a matrix eigenvalue problem for excitation energies Ω_I and oscillator strengths f_I:

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}^* & \mathbf{A}^*\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} = \Omega\begin{pmatrix}\mathbf{1} & \mathbf{0}\\ \mathbf{0} & -\mathbf{1}\end{pmatrix}\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} \tag{11.10}$$

where the matrix elements are indexed by occupied-virtual orbital pairs (ia):

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + (ia|jb) + (ia|f_{xc}|jb) \tag{11.11}$$

$$B_{ia,jb} = (ia|bj) + (ia|f_{xc}|bj) \tag{11.12}$$

The Tamm-Dancoff approximation (TDA) sets $\mathbf{B} = 0$, simplifying the problem to a standard Hermitian eigenvalue problem:

$$\mathbf{AX} = \Omega\mathbf{X} \tag{11.13}$$

The oscillator strength for each excitation I is:

$$f_I = \frac{2}{3}\Omega_I \sum_{\alpha\in\{x,y,z\}} |\langle I|\hat{r}_\alpha|0\rangle|^2 \tag{11.14}$$

and the optical absorption spectrum is the sum $\sum_I f_I\,\delta(\omega - \Omega_I)$.

11.5 Accuracy and Limitations of TD-DFT

11.5.1 Successes

- Valence excitations in organic molecules: errors typically 0.2–0.5 eV.
- Computational cost: $O(N^2)$–$O(N^3)$, far cheaper than CASPT2 or EOM-CCSD.
- Semi-quantitative optical spectra of large chromophores, transition metal complexes.

11.5.2 Known failures

(i) **Charge-transfer excitations:** Standard LDA/GGA kernels give drastically wrong (too low by 1–2 eV) CT excitation energies because the adiabatic XC kernel lacks the $-1/r$ discontinuity at large separation. Range-separated hybrids (CAM-B3LYP) largely fix this.

(ii) **Rydberg excitations:** The LDA/GGA potential decays too fast at large r, giving a compressed Rydberg series. Asymptotic correction (LB94, SAOP) or range-separated hybrids improve this.

(iii) **Doubly excited states:** The adiabatic kernel cannot describe states with dominant double-excitation character, because the exact f_{xc} must have poles at double-excitation frequencies but the adiabatic approximation is frequency-independent.

(iv) **Conical intersections:** The topology of potential energy surfaces near conical intersections is wrong in adiabatic TD-DFT; S_1/S_0 crossings are lifted to avoided crossings.

(v) **Excited-state geometry optimization:** Applicable in TDA approximation, but less reliable than CASSCF/CASPT2 for strongly correlated excited states.

Chapter 12

Beyond Standard DFT: GW, BSE, and the Path to Spectroscopy

12.1 Many-Body Perturbation Theory: Framework

When accurate quasiparticle energies (band gaps, photoemission spectra) are needed beyond the KS eigenvalues, one turns to Many-Body Perturbation Theory (MBPT). The central object is the single-particle Green's function:

$$G(\mathbf{r},\mathbf{r}';\omega) = \sum_n \frac{\phi_n(\mathbf{r})\phi_n^*(\mathbf{r}')}{\omega - \epsilon_n - \Sigma_n + i\eta\,\mathrm{sgn}(\epsilon_n - \mu)} \tag{12.1}$$

where Σ_n is the quasiparticle self-energy correction to the KS eigenvalue ϵ_n. The quasiparticle energies E_n^{QP} are the poles of G:

$$E_n^{QP} = \epsilon_n^{KS} + Z_n\langle\phi_n|\Sigma(E_n^{QP}) - v_{\mathrm{xc}}|\phi_n\rangle \tag{12.2}$$

where $Z_n = [1 - \partial\mathrm{Re}\,\Sigma/\partial\omega|_{\epsilon_n}]^{-1}$ is the quasiparticle renormalization factor.

12.2 The GW Approximation

Hedin (1965) derived a self-consistent set of equations for the self-energy, which when approximated as:

$$\Sigma(\mathbf{r},\mathbf{r}';\omega) = \frac{i}{2\pi}\int G(\mathbf{r},\mathbf{r}';\omega+\omega')\,W(\mathbf{r},\mathbf{r}';\omega')\,d\omega' \tag{12.3}$$

defines the GW approximation. Here W is the screened Coulomb interaction:

$$W(\mathbf{r},\mathbf{r}';\omega) = \int \epsilon^{-1}(\mathbf{r},\mathbf{r}'';\omega)\,v(\mathbf{r}''-\mathbf{r}')\,d\mathbf{r}'' \tag{12.4}$$

where ϵ is the dielectric function. The dielectric function is computed within the Random Phase Approximation (RPA):

$$\epsilon(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r}-\mathbf{r}') - \int v(\mathbf{r}-\mathbf{r}'')\chi_s(\mathbf{r}'',\mathbf{r}';\omega)\,d\mathbf{r}'' \tag{12.5}$$

In practice, one performs a single-shot G_0W_0 calculation: use KS orbitals and eigenvalues to construct G_0 and W_0, evaluate $\Sigma = iG_0W_0$, and compute quasiparticle corrections perturbatively. The G_0W_0 approach (starting from LDA or PBE KS states) gives band gaps accurate to $\sim$ 0.1–0.3 eV for many semiconductors and insulators.

12.2.1 Accuracy of GW

- G_0W_0@LDA: Good for *sp* semiconductors (Si, GaAs, ZnO), slight underestimation of gaps.
- G_0W_0@PBE0 or @HSE: Improved starting point for wide-gap insulators and strongly correlated systems.
- Self-consistent GW (scGW or quasiparticle self-consistent GW, QSGW): Better for correlated systems, ionization energies; computationally demanding.

12.3 The Bethe–Salpeter Equation (BSE)

Optical spectra involve coupled electron-hole excitations (excitons). The Bethe–Salpeter equation (BSE) describes the two-particle electron-hole Green's function L:

$$L = L_0 + L_0 K L \tag{12.6}$$

where $L_0 = G \cdot G$ (free electron-hole propagator) and $K = v - W$ is the electron-hole interaction kernel (bare exchange v plus screened direct interaction $-W$). The resulting eigenvalue problem has the same structure as the Casida equation (Eq. 11.10) but uses quasiparticle energies and the screened Coulomb kernel W instead of f_{xc}. The BSE gives exciton binding energies and optical spectra of semiconductors with ~ 0.1 eV accuracy, capturing strongly bound excitons (e.g., LiF: $E_b \approx 1$ eV) that TD-DFT completely misses.

Chapter 13

Current Frontiers: Machine Learning Functionals and Emerging Methods

13.1 Machine Learning Exchange-Correlation Functionals

The limitation of traditional DFT functional development is that approximations are constrained to simple functional forms (local, semi-local, single nonlocal term). Machine learning offers a radically different approach: train a flexible model on high-quality reference data, with or without explicit enforcement of exact constraints.

13.1.1 Neural network functionals

Snyder et al. (2012) demonstrated that neural networks can learn the kinetic energy density functional $T_s[n]$ for 1D model systems from exact data. This opened the door to machine learning the full $E_{xc}[n]$.

Key strategies include:

(i) **Learn E_{xc} directly from densities:** Represent $n(\mathbf{r})$ as a set of features (e.g., on a grid or in a local descriptor) and train a neural network to map features to E_{xc}.

(ii) **Learn v_{xc} from densities and reference potentials:** Invert high-quality wavefunctions (CCSD(T) or FCI) to obtain the exact v_{xc}, then train a model.

(iii) **Constraint-satisfying ML:** Build neural networks that by construction satisfy scaling relations, sum rules, and the Lieb-Oxford bound.

13.1.2 DeepMind DM21

The DM21 functional by Kirkpatrick et al. (Nature, 2021) is a neural network XC functional trained on:

1. Fractional electron number ($N \pm \delta$) data to enforce piecewise linearity.
2. Fractional spin data to reduce static correlation error.
3. Standard thermochemical databases (atomization energies, reaction barriers).

DM21 achieves near-CCSD(T) accuracy on a wide range of molecules and, importantly, explicitly enforces piecewise linearity—addressing the fundamental delocalization error of conventional functionals. Its inputs are similar to a meta-hybrid GGA (local density, gradients, kinetic energy density, HF exchange density), but with many more parameters fitted by neural networks.

13.1.3 Challenges for ML functionals

- **Transferability:** Neural networks may fail for systems not in the training set.
- **Efficiency:** Evaluating the XC potential requires computing $\delta E_{xc}/\delta n$ by backpropagation, which is slower than traditional functionals.
- **Interpretability:** Hard to understand what physics the network has learned.
- **Consistency:** Forces and stresses computed from ML functionals require careful implementation to ensure energy conservation.

13.2 Ensemble DFT for Excited States

Gross, Oliveira, and Kohn (1988) extended the HK theorem to ensembles of states. For a weighted ensemble $\hat{\Gamma} = \sum_k w_k |\Psi_k\rangle\langle\Psi_k|$ with weights $w_1 \geq w_2 \geq \cdots \geq 0$, $\sum_k w_k = 1$, the ensemble density:

$$n^{\mathbf{w}}(\mathbf{r}) = \sum_k w_k n_k(\mathbf{r}) \tag{13.1}$$

uniquely determines the external potential (generalized HK theorem). The ensemble energy is:

$$E^{\mathbf{w}}[n^{\mathbf{w}}] = \sum_k w_k E_k \tag{13.2}$$

For an ensemble of the ground state and first excited state with weights $(1 - w, w)$:

$$E_1 = \frac{E^{\mathbf{w}} - (1-w)E_0}{w} \tag{13.3}$$

This provides a rigorous framework for excited-state DFT. In practice, accurate ensemble XC functionals for excited states are still under development.

13.3 The Random Phase Approximation and Beyond

The Random Phase Approximation (RPA) for the correlation energy provides a non-local, dispersion-inclusive correlation functional. Starting from the adiabatic connection:

$$E_c^{RPA} = \frac{1}{2}\int_0^\infty \frac{d\omega}{2\pi} \mathrm{Tr}\left\{\ln\left(1 - v\chi_s^{i\omega}\right) + v\chi_s^{i\omega}\right\} \tag{13.4}$$

RPA naturally includes van der Waals interactions, describes metallic screening, and avoids the SIE problem. Its main failures are: overcorrection of short-range correlation, self-correlation error for electron gases, and $O(N^4)$ formal scaling. Beyond-RPA corrections (SOSEX, rPT2, ACFDT-RPA+) address these issues progressively.

13.4 Density Matrix Functional Theory (DMFT/RDMFT)

Reduced density matrix functional theory (RDMFT) uses the 1-RDM $\gamma(\mathbf{r}, \mathbf{r}')$ as the basic variable instead of just the diagonal $n(\mathbf{r})$. The Hohenberg-Kohn-like theorem for the 1-RDM (Gilbert, 1975) states that the ground-state energy is a functional of γ. Since the kinetic energy is $T = -\frac{1}{2}\int \nabla^2\gamma(\mathbf{r}, \mathbf{r}')|_{\mathbf{r}'=\mathbf{r}}\, d\mathbf{r}$ (exact, not approximated), RDMFT sidesteps the kinetic energy functional problem entirely. The 2-electron reduced density matrix is approximated as a functional of γ (e.g., Müller, Goedecker-Umrigar, Piris functionals). RDMFT can in principle describe strongly correlated systems and fractional occupation, but the N-representability conditions on γ are complex to enforce.

13.5 Developments in Periodic DFT: Dielectric Properties and DFPT

Density Functional Perturbation Theory (DFPT) computes the linear response of the KS system to external perturbations (phonons, electric fields, strain). For a perturbation λ, the response of the density:

$$\frac{\partial n(\mathbf{r})}{\partial \lambda} = 2\sum_{i}^{occ} \phi_i(\mathbf{r}) \frac{\partial \phi_i(\mathbf{r})}{\partial \lambda} \tag{13.5}$$

can be computed self-consistently through the first-order KS equations. DFPT gives access to phonon dispersions (full Brillouin zone without supercells), dielectric tensors, Born effective charges, Raman and infrared spectra, and electron-phonon coupling constants—all at $O(N^3)$ cost with no additional empirical parameters.

Chapter 14

Practical Aspects and Best Practices

14.1 Choosing the Right Functional

There is no universal functional. The choice depends on the property of interest:

Property	**Recommended functional**
Ground-state geometries (molecules)	PBE, BLYP, B3LYP, M06-2X
Thermochemistry (organic, main group)	B3LYP-D3, M06-2X, ωB97X-D
Reaction barriers	M06-2X, ωB97X-D, double hybrids
Solids, surfaces (lattice constants)	PBE, PBEsol, SCAN, r^{2}SCAN
Band gaps (semiconductors)	HSE06, PBE0, G_0W_0@PBE
Band gaps (strongly correlated)	LDA+U, G_0W_0@PBE0, QSGW
van der Waals interactions	PBE-D3(BJ), rev-vdW-DF2, MBD
Optical spectra (valence)	TD-B3LYP, TD-CAM-B3LYP
Optical spectra (charge transfer)	TD-CAM-B3LYP, BSE@G_0W_0
Magnetic properties	PBE (with spin-polarization), PBE+U
Ionization energies, electron affinities	Tuned ω-range-separated hybrids

14.2 Convergence and Quality Control

A DFT calculation is only as good as its convergence. The key parameters to converge are:

(i) **Basis set:** Increase E_{cut} (plane waves) or basis size (Gaussians/NAOs) until total energy changes are below 1 meV/atom.

(ii) **k-point sampling:** For metals, use dense Monkhorst-Pack meshes; for insulators, $4\times4\times4$ or finer. Check convergence of total energy.

(iii) **SCF convergence:** Typical thresholds: 10^{-6} Hartree in energy, 10^{-5} electrons in density.

(iv) **Force convergence:** For geometry optimization, < 0.02 eV/Å per atom.

(v) **Smearing:** Use the smallest smearing width that avoids convergence difficulties; extrapolate to zero smearing.

14.3 Error Analysis and Benchmarking

Every DFT calculation should be accompanied by an honest assessment of expected errors:

- Compare with known experimental data or high-level quantum chemical reference calculations on related molecules.
- Report uncertainties: LDA overbinds by $\sim$1 eV/bond; GGA atomization errors $\sim$0.3 eV/bond; hybrids $\sim$0.1 eV/bond.
- Use multiple functionals and check for qualitative agreement on key conclusions.
- For band gaps: always note whether the gap is the KS gap (and thus underestimated) or has been corrected.

Chapter 15

Summary: The DFT Landscape

15.1 The Flow of DFT Theory

We close with a synthesis of the main conceptual thread:

1. The exact ground state of N electrons in potential v satisfies the many-body Schrödinger equation, but the wavefunction is intractable for $N \gtrsim 10$.
2. **Hohenberg-Kohn:** The ground-state density $n(\mathbf{r})$ uniquely determines v and hence all ground-state properties. This justifies the density as the fundamental variable, and defines the universal functional $F[n]$.
3. **Kohn-Sham:** Replace the interacting problem with a non-interacting system that has the same density, computed exactly via orbital equations. All unknowns are in $E_{xc}[n]$.
4. **Jacob's Ladder:** Systematically improve E_{xc} from LDA → GGA → meta-GGA → hybrids → double hybrids, adding more non-locality and satisfying more exact constraints at each rung.
5. **Corrections:** Add dispersion (D3, MBD, vdW-DF), strong-correlation ($+U$, DMFT), and self-interaction corrections as needed.
6. **Extensions:** TD-DFT for optical spectra; GW+BSE for quasiparticle gaps and excitons; DFPT for phonons and dielectric properties.
7. **Future:** Machine-learning functionals that learn the exact XC hole from high-quality reference data while satisfying exact constraints.

The Central Equation of DFT

All of ground-state DFT reduces to finding the self-consistent solution of the Kohn-Sham equations:

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}\right] \phi_i = \epsilon_i \phi_i, \quad n = \sum_i |\phi_i|^2 \tag{15.1}$$

This is exact given the exact E_{xc}. All of DFT methodology is the art of approximating E_{xc}.

Bibliography

[1] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Phys. Rev.* **136**, B864 (1964).

[2] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Phys. Rev.* **140**, A1133 (1965).

[3] L. H. Thomas, "The calculation of atomic fields," *Math. Proc. Cambridge Philos. Soc.* **23**, 542 (1927).

[4] E. Fermi, "Un metodo statistico per la determinazione di alcune proprietà dell'atome," *Rend. Accad. Naz. Lincei* **6**, 602 (1927).

[5] P. A. M. Dirac, "Note on Exchange Phenomena in the Thomas Atom," *Math. Proc. Cambridge Philos. Soc.* **26**, 376 (1930).

[6] M. Levy, "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem," *Proc. Natl. Acad. Sci.* **76**, 6062 (1979).

[7] E. H. Lieb, "Density functionals for Coulomb systems," *Int. J. Quantum Chem.* **24**, 243 (1983).

[8] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," *Phys. Rev. B* **45**, 13244 (1992).

[9] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," *Phys. Rev. Lett.* **77**, 3865 (1996).

[10] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations," *Can. J. Phys.* **58**, 1200 (1980).

[11] D. M. Ceperley and B. J. Alder, "Ground State of the Electron Gas by a Stochastic Method," *Phys. Rev. Lett.* **45**, 566 (1980).

[12] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," *Phys. Rev. B* **23**, 5048 (1981).

[13] J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly Constrained and Appropriately Normed Semilocal Density Functional," *Phys. Rev. Lett.* **115**, 036402 (2015).

[14] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, "Accurate and Numerically Efficient r^{2}SCAN Meta-Generalized Gradient Approximation," *J. Phys. Chem. Lett.* **11**, 8208 (2020).

[15] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *J. Chem. Phys.* **98**, 5648 (1993); P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).

[16] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened Coulomb potential," *J. Chem. Phys.* **118**, 8207 (2003).

[17] S. Grimme, "Semiempirical GGA-type density functional constructed with a long-range

dispersion correction," *J. Comput. Chem.* **27**, 1787 (2006).

[18] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," *J. Chem. Phys.* **132**, 154104 (2010).

[19] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, "Van der Waals Density Functional for General Geometries," *Phys. Rev. Lett.* **92**, 246401 (2004).

[20] J. F. Janak, "Proof that $\partial E/\partial n_i = \varepsilon_i$ in density-functional theory," *Phys. Rev. B* **18**, 7165 (1978).

[21] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, "Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy," *Phys. Rev. Lett.* **49**, 1691 (1982).

[22] L. J. Sham and M. Schlüter, "Density-Functional Theory of the Energy Gap," *Phys. Rev. Lett.* **51**, 1888 (1983).

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

[24] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study," *Phys. Rev. B* **57**, 1505 (1998).

[25] E. Runge and E. K. U. Gross, "Density-Functional Theory for Time-Dependent Systems," *Phys. Rev. Lett.* **52**, 997 (1984).

[26] M. E. Casida, "Time-dependent density functional response theory for molecules," in *Recent Advances in Density Functional Methods*, ed. D. P. Chong (World Scientific, 1995), p. 155.

[27] L. Hedin, "New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem," *Phys. Rev.* **139**, A796 (1965).

[28] N. D. Mermin, "Thermal Properties of the Inhomogeneous Electron Gas," *Phys. Rev.* **137**, A1441 (1965).

[29] P. E. Blöchl, "Projector augmented-wave method," *Phys. Rev. B* **50**, 17953 (1994).

[30] D. Vanderbilt, "Soft self-consistent pseudopotentials in a generalized eigenvalue formalism," *Phys. Rev. B* **41**, 7892 (1990).

[31] J. Kirkpatrick, B. McMorrow, D. H. P. Turban et al., "Pushing the frontiers of density functionals by solving the fractional electron problem," *Science* **374**, 1385 (2021).

[32] A. Tkatchenko, R. A. DiStasio, Jr., R. Car, and M. Scheffler, "Accurate and Efficient Method for Many-Body van der Waals Interactions," *Phys. Rev. Lett.* **108**, 236402 (2012).

[33] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, 2nd ed. (Cambridge University Press, 2020).

[34] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed. (Wiley-VCH, 2001).

[35] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer, 1990).

[36] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, 1989).

[37] K. Burke, "Perspective on density functional theory," *J. Chem. Phys.* **136**, 150901 (2012).

[38] A. D. Becke, "Perspective: Fifty years of density-functional theory in chemical physics," *J. Chem. Phys.* **140**, 18A301 (2014).