

Chapter 2

Density Functional Theory

Density Functional Theory (DFT) is a *ground-state* theory in which the emphasis is on the *charge density* as the relevant physical quantity. DFT has proved to be highly successful in describing structural and electronic properties in a vast class of materials, ranging from atoms and molecules to simple crystals to complex extended systems (including glasses and liquids). Furthermore DFT is computationally simple. For these reasons DFT has become a common tool in *first-principles* calculations aimed at describing – or even predicting – properties of molecular and condensed matter systems.

2.1 The Hohenberg-Kohn Theorem

Let us consider a system of N interacting (spinless) electrons under an external potential $V(\mathbf{r})$ (usually the Coulomb potential of the nuclei). If the system has a nondegenerate ground state, it is obvious that there is only one ground-state charge density $n(\mathbf{r})$ that corresponds to a given $V(\mathbf{r})$. In 1964 Hohenberg and Kohn demonstrated the opposite, far less obvious result: there is only one external potential $V(\mathbf{r})$ that yields a given ground-state charge density $n(\mathbf{r})$. The demonstration is very simple and uses a *reductio ad absurdum* argument.

Let us consider a many-electron Hamiltonian $H = T + U + V$, with ground state wavefunction Ψ . T is the kinetic energy, U the electron-electron interaction, V the external potential. The charge density $n(\mathbf{r})$ is defined as

$$n(\mathbf{r}) = N \int |\Psi(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N)|^2 d\mathbf{r}_2 \dots d\mathbf{r}_N. \quad (2.1)$$

Let us consider now a *different* Hamiltonian $H' = T + U + V'$ (V and V' do not differ simply by a constant: $V - V' \neq \text{const.}$), with ground state wavefunction Ψ' . Let us assume that the ground state charge densities are the same: $n[V] = n'[V']$. The following inequality holds:

$$E' = \langle \Psi' | H' | \Psi' \rangle < \langle \Psi | H' | \Psi \rangle = \langle \Psi | H + V' - V | \Psi \rangle \quad (2.2)$$

that is,

$$E' < E + \int (V(\mathbf{r}) - V'(\mathbf{r}))n(\mathbf{r})d\mathbf{r}. \quad (2.3)$$

The inequality is strict because Ψ and Ψ' are different, being eigenstates of different Hamiltonians. By reversing the primed and unprimed quantities, one obtains an absurd result. This demonstrates that no two different potentials can have the same charge density.

A subtle point about the existence of the potential corresponding to a given ground state charge density (the *v-representability* problem), and the various extensions of the Hohenberg and Kohn theorem, are discussed in the specialized literature.

A straightforward consequence of the first Hohenberg and Kohn theorem is that the ground state energy E is also uniquely determined by the ground-state charge density. In mathematical terms E is a *functional* $E[n(\mathbf{r})]$ of $n(\mathbf{r})$. We can write

$$E[n(\mathbf{r})] = \langle \Psi | T + U + V | \Psi \rangle = \langle \Psi | T + U | \Psi \rangle + \langle \Psi | V | \Psi \rangle = F[n(\mathbf{r})] + \int n(\mathbf{r})V(\mathbf{r})d\mathbf{r} \quad (2.4)$$

where $F[n(\mathbf{r})]$ is a *universal* functional of the charge density $n(\mathbf{r})$ (and *not* of $V(\mathbf{r})$). For this functional a variational principle holds: the ground-state energy is *minimized* by the ground-state charge density. In this way, DFT exactly reduces the N -body problem to the determination of a 3-dimensional function $n(\mathbf{r})$ which minimizes a functional $E[n(\mathbf{r})]$. Unfortunately this is of little use as $F[n(\mathbf{r})]$ is not known.

2.2 The Kohn-Sham equations

One year later, Kohn and Sham (KS) reformulated the problem in a more familiar form and opened the way to practical applications of DFT. The system of interacting electrons is mapped on to an auxiliary system of non-interacting electrons having the same ground state charge density $n(\mathbf{r})$. For a system of non-interacting electrons the ground-state charge density is representable as a sum over one-electron orbitals (the *KS orbitals*) $\psi_i(\mathbf{r})$:

$$n(\mathbf{r}) = 2 \sum_i |\psi_i(\mathbf{r})|^2, \quad (2.5)$$

where i runs from 1 to $N/2$ if we assume double occupancy of all states, and the KS orbitals are the solutions of the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V_{KS}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (2.6)$$

(m is the electron mass) obeying orthonormality constraints:

$$\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}. \quad (2.7)$$

The existence of a unique potential $V_{KS}(\mathbf{r})$ having $n(\mathbf{r})$ as its ground state charge density is a consequence of the Hohenberg and Kohn theorem, which holds irrespective of the form of the electron-electron interaction U .

2.3 Kohn-Sham equations and the variational principle

The problem is now to determine $V_{KS}(\mathbf{r})$ for a given $n(\mathbf{r})$. This problem is solved by considering the variational property of the energy. For an arbitrary variation of the $\psi_i(\mathbf{r})$,

under the orthonormality constraints of Eq. (2.7), the variation of E must vanish. This translates into the condition that the functional derivative (see appendix) with respect to the ψ_i of the constrained functional

$$E' = E - \sum_{ij} \lambda_{ij} \left(\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} - \delta_{ij} \right), \quad (2.8)$$

where λ_{ij} are Lagrange multipliers, must vanish:

$$\frac{\delta E'}{\delta \psi_i^*(\mathbf{r})} = \frac{\delta E'}{\delta \psi_i(\mathbf{r})} = 0. \quad (2.9)$$

It is convenient to rewrite the energy functional as follows:

$$E = T_s[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] + \int n(\mathbf{r}) V(\mathbf{r}) d\mathbf{r}. \quad (2.10)$$

The first term is the kinetic energy of *non-interacting* electrons:

$$T_s[n(\mathbf{r})] = -\frac{\hbar^2}{2m} 2 \sum_i \int \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) d\mathbf{r}. \quad (2.11)$$

The second term (called the Hartree energy) contains the electrostatic interactions between clouds of charge:

$$E_H[n(\mathbf{r})] = \frac{e^2}{2} \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (2.12)$$

The third term, called the *exchange-correlation energy*, contains all the remaining terms: our ignorance is hidden there. The logic behind such procedure is to subtract out easily computable terms which account for a large fraction of the total energy.

Using

$$\frac{\delta n(\mathbf{r})}{\delta \psi_i^*(\mathbf{r}')} = \psi_i(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \quad (2.13)$$

and the formulae given in the appendix, one finds

$$\frac{\delta T_s}{\delta \psi_i^*(\mathbf{r})} = -\frac{\hbar^2}{2m} 2 \sum_i \nabla^2 \psi_i(\mathbf{r}), \quad (2.14)$$

$$\frac{\delta E_H}{\delta \psi_i^*(\mathbf{r})} = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \psi_i(\mathbf{r}) \quad (2.15)$$

and finally

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V_H(\mathbf{r}) + V_{xc}[n(\mathbf{r})] + V(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \sum_j \lambda_{ij} \psi_j(\mathbf{r}) \quad (2.16)$$

where we have introduced a *Hartree potential*

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (2.17)$$

and an *exchange-correlation potential*

$$V_{xc}[n(\mathbf{r})] = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}. \quad (2.18)$$

The Lagrange multiplier λ_{ij} are obtained by multiplying both sides of Eq.2.16 by $\psi_k^*(\mathbf{r})$ and integrating:

$$\lambda_{ik} = \int \psi_k^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 + V_H(\mathbf{r}) + V_{xc}[n(\mathbf{r})] + V(\mathbf{r}) \right) \psi_i(\mathbf{r}) d\mathbf{r}. \quad (2.19)$$

For an insulator, whose states are either fully occupied or completely empty, it is always possible to make a subspace rotation in the space of ψ 's (leaving the charge density invariant). We finally get the KS equations:

$$(H_{KS} - \epsilon_i) \psi_i(\mathbf{r}) = 0, \quad (2.20)$$

where $\lambda_{ij} = \delta_{ij} \epsilon_j$ and the operator H_{KS} , called KS Hamiltonian, is defined as

$$H_{KS} = -\frac{\hbar^2}{2m} \nabla^2 + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + V(\mathbf{r}) \equiv -\frac{\hbar^2}{2m} \nabla^2 + V_{KS}(\mathbf{r}) \quad (2.21)$$

and is related to the functional derivative of the energy:

$$\frac{\delta E}{\delta \psi_i^*(\mathbf{r})} = H_{KS} \psi_i(\mathbf{r}). \quad (2.22)$$

2.4 Local Density Approximation for the exchange-correlation energy

The KS equations are somewhat reminiscent of the Hartree-Fock (HF) equations. Both are derived from a variational principle: the minimization of the energy functional for the latter, of the energy for a single Slater determinant wavefunction for the former. Both are self-consistent equations for one-electron wavefunctions. In the HF equations the *exchange* term appears in the place of the exchange-correlation potential of KS equations:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V_H(\mathbf{r}) + V(\mathbf{r}) \right) \psi_i(\mathbf{r}) - e^2 \sum_{j, \parallel} \int \frac{\psi_j(\mathbf{r}) \psi_j^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}') d\mathbf{r}' = \epsilon_i \psi_i(\mathbf{r}) \quad (2.23)$$

where the sum over j extends only to states with parallel spins. Traditionally, one defines the *correlation energy* as the difference between the HF and the real energy. The name “exchange-correlation” in DFT reflects such tradition¹.

The exchange term in the HF equations is a *nonlocal* operator – one acting on a function ϕ as $(V\phi)(\mathbf{r}) = \int V(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') d\mathbf{r}'$, and is quite difficult to compute. In earlier calculations, done with primitive computer machinery (or even *without* any computer machinery), the Slater’s local approximation (see Sec.(1.3.1) was often used:

$$(V\phi)(\mathbf{r}) = V_x(\mathbf{r}) \phi(\mathbf{r}), \quad V_x(\mathbf{r}) = -\frac{3e^2}{2\pi} \left[3\pi^2 n(\mathbf{r}) \right]^{1/3}. \quad (2.24)$$

¹Note however that the exchange-correlation energy of DFT is not exactly the sum of HF exchange plus correlation energy: see Sec.(2.6)

Within this approximation, the only difference between Kohn-Sham and Hartree-Fock equations is the form of the exchange-correlation, or exchange only, potential.

We still don't have a reasonable estimate for the exchange-correlation energy $E_{xc}[n(\mathbf{r})]$ and potential. Kohn and Sham introduced, as early as 1965, the Local Density Approximation (LDA): in the spirit of Slater's local exchange, they approximated the functional with a *function* of the local density $n(\mathbf{r})$:

$$E_{xc}[n(\mathbf{r})] = \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r}, \quad \frac{\delta E_{xc}}{\delta n(\mathbf{r})} \equiv \mu_{xc}(n(\mathbf{r})) = \left(\epsilon_{xc}(n) + n \frac{d\epsilon_{xc}(n)}{dn} \right)_{n=n(\mathbf{r})} \quad (2.25)$$

and they used for $\epsilon_{xc}(n(\mathbf{r}))$ the same dependence upon the density as for the homogeneous electron gas. The latter is unknown, except at the HF level (see Eq.(1.23)), but approximate forms have been known for a long time, going back to Wigner (1931). Highly accurate results from Quantum Monte-Carlo techniques were found by Ceperley and Alder and parameterized by Perdew and Zunger with a simple analytical form:

$$\begin{aligned} \epsilon_{xc}(n) &= -0.9164/r_s - 0.2846/(1 + 1.0529\sqrt{r_s} + 0.3334r_s), & r_s \geq 1 \\ &= -0.9164/r_s - 0.0960 + 0.0622 \ln r_s - 0.0232r_s + 0.0040r_s \ln r_s, & r_s \leq 1. \end{aligned} \quad (2.26)$$

Here r_s , defined in Eq.(1.9), is in Bohr radii and ϵ_{xc} is in Ry. The first term is the HF exchange contribution, Eq.(1.23), the remaining terms are the correlation energy. Several other expressions for ϵ_{xc} can be found in the literature. All forms yield very similar results in condensed-matter calculations, which is not surprising, since all parameterizations are very similar in the range of r_s applicable for solid-state phenomena.

2.5 Successes and failures of DFT

DFT, even in the simplest LDA approximation, turns out to be much more successful than expected. Especially for solids, LDA is computationally much simpler than HF with the true exchange potential and no more complex than Slater's local exchange approximation. Yet, LDA yields results that compare well to HF results, even in atoms and molecules – highly inhomogeneous systems for which an approximation based on the homogeneous electron gas would hardly look appropriate. The best results are however obtained in solids, whose structural and vibrational properties are in general well described: the correct crystal structure is usually found to have the lowest energy; bond lengths, bulk moduli, phonon frequencies are accurate within a few percent.

One may wonder why LDA is so successful, given its resemblance with the not-so-praised Slater approximation to HF. One reason is somewhat fortuitous: LDA contains a fair amount of error compensation between the exchange and correlation parts. A deeper reason is explained in Sec.(2.6): LDA grants a good description of the spherical term of the so-called “exchange-correlation hole”.

LDA also has some well-known serious problems. Some can be avoided by using better functionals, some others have a deeper and more fundamental nature.

2.5.1 Gradient-corrected functionals

The accuracy of LDA is often considered satisfactory in condensed-matter systems, but it is much less so in atomic and molecular physics, for which highly accurate experimental

data are available. Also, LDA badly overestimates ($\sim 20\%$ and more) cohesive energies and bond strengths in molecules and solids, and as a consequence bond lengths are often underestimated.

Such problems are mostly corrected by the introduction of *gradient corrections*. The exchange-correlation functional is written as a function of the local density *and* of the local gradient of the density, usually as an “enhancement factor” F_{xc} multiplying the homogeneous electron results:

$$E_{xc}[n(\mathbf{r})] = \int \epsilon_{xc}(n(\mathbf{r})F_{xc}(n(\mathbf{r}), |\nabla n(\mathbf{r})|)n(\mathbf{r}))d\mathbf{r}. \quad (2.27)$$

The enhancement factor is written in terms of r_s and of a dimensionless reduced density gradient $s(\mathbf{r})$:

$$F_{xc}(n(\mathbf{r}), |\nabla n(\mathbf{r})|) \rightarrow F_{xc}(r_s, s), \quad s(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|}{2k_F n(\mathbf{r})}. \quad (2.28)$$

Gradient-corrected functionals are the simplest extension of LDA to inhomogeneous systems one can think of. Their adoption has been quite slow: earlier functionals, based on the so-called “gradient expansion”, turned out to yield mixed results, but a later generation of “generalized gradient approximations”, GGA, performed much better and eventually found widespread acceptance. GGA functionals are known to satisfy some known conditions that the exact functional should satisfy as well. They yield much better atomic energies and binding energies than LDA, at a modest additional computational cost. In particular, they yield a good description of the Hydrogen bond, thus opening the way to calculations for systems, such as water, in which Hydrogen bonds play a crucial role (LDA is unusable for those systems: the Hydrogen bond in LDA is way too strong).

2.5.2 Weakly-bonded and strongly correlated systems

The weak van der Waals (vdW) forces between closed-shells systems, that are responsible for the physisorption, are however still beyond the reach of simple approximations to DFT. The vdW (or dispersive) interactions have a *nonlocal* character: they are due to charge fluctuations on one system, inducing a dipole on the other. This phenomenon is absent by construction from LDA, as well as from any functional based on the *local* density $n(\mathbf{r})$ and on its *local* derivatives $|\nabla n(\mathbf{r})|$.² Let us consider a system formed by two fragments a and b with non-overlapping charge densities, $n_a(\mathbf{r})$ and $n_b(\mathbf{r})$. For such system, $V_{xc}(n(\mathbf{r})) = V_{xc}(n_a(\mathbf{r})) + V_{xc}(n_b(\mathbf{r}))$, i.e. in absence of overlap there is no interaction energy between the two fragments (assuming that the two systems are neutral and the charge densities spherical; otherwise there will be the usual electrostatic interactions, which however do not include van der Waals interactions).

As a matter of facts, LDA overestimates the attractive potential coming from the overlap of the tails of the charge density. As a consequence, closed-shell systems in LDA are actually bound with binding energies and binding distances in apparent agreement with experimental results. This is a fictitious result (and the dependence on the separation distance is wrong) that disappears if better-behaved gradient-corrected functionals are used (thus leading to the funny case of a better approximation yielding an apparently

²Although the exchange-correlation potential is a local potential of \mathbf{r} , i.e. it is described by a function $V(\mathbf{r})$, its functional dependence upon $n(\mathbf{r})$ is in principle nonlocal, as it is for instance in the Hartree potential.

worse result than a less good approximation). This situation is quite embarrassing: vdW-bonded systems can be better studied with classical molecular dynamics and interatomic potentials, but in many interesting systems (e.g.: molecules on surfaces, enzymatic sites in biological systems) whose electronic structure must be studied with DFT, vdW interactions are also present and important. Currently, there are two kinds of possible approaches: i) adding semi-empirical vdW interactions, or ii) using a *nonlocal* exchange-correlation functional accounting for vdW interactions as well. The former is a simple and inexpensive solution but not really "first-principle". The latter is computationally much heavier, but recent breakthroughs suggest that it may become soon a viable option.

Another field in which simple DFT approximations often fail is in strongly correlated materials. These systems contain localized, atomic-like electronic states, typically originating from *d* or *f* atomic states, together with delocalized, band-like states originating from *s* and *p* states. The correct description of both set of states is problematic: LDA and GGA may easily produce an incorrect occupancy of the localized states, leading to an incorrect description of all properties of the material under study. Typically, one finds localized states that are less localized than they should be, with fractional instead of integer occupations. A famous example of failure is NiO: an insulator with a simple structure, that simple DFT approximations consistently predict to be a metal. A similar behavior is displayed by K_4C_{60} , even if no *d* or *f* states are present: in this case the electron correlation on the C_{60} molecule is responsible. Currently two approaches yield better results: i) the so-called "DFT+U", in which a Coulomb repulsion term, U , responsible for the peculiar behavior of highly correlated materials, is added to a DFT calculation, and ii) *hybrid* functionals, i.e. containing an admixture of HF exchange. The former approach is computationally cheap but not very rigorous; the latter is quite expensive in solids.

2.5.3 The problem of self-interaction

A large number of problems in simple DFT approximations comes from *self-interaction*, i.e. the interaction of an electron with the field it generates. In HF, self-interaction cancels exactly, by construction. A simple test for the presence of self-interaction is provided by one-electron systems, for which the sum of exchange-correlation and Hartree potentials and energy should be exactly zero. The cancellation is actually quite good: in the H atom, within 5%, but not good enough to be taken for granted. Self-interaction mostly affects finite systems, or systems containing localized electrons, while its effect is vanishing for delocalized electronic states in extended systems (solids). In finite systems the presence of self-interaction is reflected in an incorrect long-range behavior of the potential felt by an electron. For an atom, we should have $V_{xc}(r) \rightarrow -1/r$ for $r \rightarrow \infty$, but simple functionals yield instead a potential that decays exponentially.

A remarkable example of the effects of self-interaction is the Al defect in SiO_2 . In SiO_2 , four O atoms are bound to a Si atom, forming a tetrahedron; each O is shared between two such tetrahedra. The strong bonding involves the sp^3 hybrid Si orbital and the $2p$ O orbitals. The presence of Al introduces a "hole" in the electronic structure. In DFT, one finds that the "hole" is evenly distributed on the Al-O bonds and that the local structure is still tetrahedral. Experiments show however that the hole is localized on one of the Al-O bonds, that is longer and weaker than the three others, thus yielding a local distorted structure around Al. The reason for this failure – affecting both LDA and GGA – has been tracked to self-interaction.

Several recipes for “Self-Interaction Correction” (SIC) are known, but their theoretical foundation is not completely convincing and their usage often meets numerical trouble, so their usefulness is still a matter of debate. More advanced functionals (e.g hybrid functionals) seem to perform better in this respect.

2.5.4 The “band gap problem”

One would like very much to be able to calculate one-electron energies having the meaning of removal (or addition) energies, as for a non interacting system (in the language of many-body theory, *quasiparticle* energies). If one electron in the state v is removed from the system, $E_N - E_{N-1} = \epsilon_v$, where E_N is the energy of the system with N electrons. If one electron is added to the system in the state c , $E_{N+1} - E_N = \epsilon_c$. The difference between the largest addition energy and the smallest removal energy defines the energy band gap: $E_g = \epsilon_c - \epsilon_v = E_{N+1} + E_{N-1} - 2E_N$. In solids this is the onset of the continuum of optical transitions, if the gap is direct (if the lowest empty state and the highest filled state have the same \mathbf{k} vector). From atomic and molecular physics, the highest occupied and lowest unoccupied states are respectively called HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied MO), while addition and removal energy are respectively called *electron affinity*, A , and *ionization potential*, I .

In HF the one-electron energies have the meaning of removal (or addition) energies for extended systems (Koopman’s theorem). If the world were described by single Slater determinants, the difference between the LUMO and HOMO one-electron HF energies would yield the real energy gaps in solids (neglecting polarization effects, i.e. the change in the one-electron states upon addition or removal of an electron). As a matter of facts, experience shows that HF (with the true exchange potential) seriously overestimates the band gap in solid.

In DFT, the one-electron energies have acquired a rather bad reputation, mostly due to the failure of KS band gaps (that is: calculated as the difference between LUMO and HOMO KS energies) to reproduce with an acceptable accuracy the true band gap in solids: gaps in DFT are strongly underestimated (up to 50%). This problem is present in both LDA and GGA and is common to HF with Slater’s local approximation to exchange.

One may say that after all DFT is a ground state theory, and that KS eigenvalues and eigenvectors are purely mathematical quantities without any physical meaning, but it wouldn’t be a very satisfactory answer. It can be demonstrated that in *exact* DFT, $I = -\epsilon_{HOMO}$ holds. In finite systems ionization potentials and electron affinities can be calculated as energy differences between the ground state and a state with one electron added or removed. In extended systems this is of course not possible.

The reason for the infamous “band gap problem” lies in the dependence of the exact energy functional upon the number of electrons and in the inability of approximate functionals to reproduce it. In the next section we consider the extension of DFT to a fractionary number of electrons.

2.5.5 The discontinuity of exchange-correlation potential

The basic variational property of the density functional can be expressed by the stationary condition

$$\frac{\delta}{\delta n(\mathbf{r})} \left(E - \mu \left(\int n(\mathbf{r}) d\mathbf{r} - N \right) \right) = 0 \quad (2.29)$$

where μ is a Lagrange multiplier and N an integer number. The formulation of DFT can be extended to noninteger number of particles $N + \omega$ ($\omega > 0$) via the following definition:

$$E[n(\mathbf{r})] = F_{frac}[n(\mathbf{r})] + \int V(\mathbf{r})n(\mathbf{r})d\mathbf{r} \quad (2.30)$$

and

$$F_{frac}[n(\mathbf{r})] = \min \text{tr}\{D(T + U)\}, \quad D = (1 - \omega)|\Psi_N\rangle\langle\Psi_N| + \omega|\Psi_{N+1}\rangle\langle\Psi_{N+1}| \quad (2.31)$$

where the minimum must be searched on all density matrices D that yield the prescribed density $n(\mathbf{r})$. It is easily verified that integration of $n(\mathbf{r})$ over all space yields $N + \omega$ electrons. With this definition the variational principle, Eq. 2.29, is defined for any number of electrons and yields the Euler equations

$$\frac{\delta E}{\delta n(\mathbf{r})} = \mu \quad (2.32)$$

and that μ is really the *chemical potential*: if we call E_N the energy at the ground state for N electrons, one has

$$\mu(N) = \frac{\partial E_N}{\partial N}. \quad (2.33)$$

There is an obvious problem if we consider $\mu(N)$ a continuous function of N for all values of N . Consider two neutral isolated atoms: in general, they will have two different values for μ . As a consequence the total energy of the two atoms will be lowered by a charge transfer from the atom at a higher chemical potential to the one at lower chemical potential.

In reality there is no paradox, because the E_N curve is not continuous. If we write down explicitly $E_{N+\omega}$, we find that both energy and minimizing charge density at fractionary number of electrons are simply a linear interpolation between the respective values at the end points with N and $N + 1$ electrons:

$$E_{N+\omega} = (1 - \omega)E_N + \omega E_{N+1}, \quad n_{N+\omega}(\mathbf{r}) = (1 - \omega)n_N(\mathbf{r}) + \omega n_{N+1}(\mathbf{r}) \quad (2.34)$$

with obvious notations. The interesting and far-reaching consequence is that there is a discontinuity of the chemical potential $\mu(N)$ and of the functional derivative $\delta E/\delta n(\mathbf{r})$ at integer N . This is an important and essential characteristic of the exact energy functional that simply reflects the discontinuity of the energy spectrum.

Coming back to our paradox: for an atom with nuclear charge Z , ionization potential $I(Z)$ and electron affinity $A(Z)$ in the ground state,

$$\mu(N) = -I(Z) \quad Z - 1 < N < Z \quad (2.35)$$

$$= -A(Z) \quad Z < N < Z + 1. \quad (2.36)$$

For a system of two neutral atoms with nuclear charges X and Y , in which ω electrons are transferred from the first to the second atom:

$$\mu(\omega) = \mu(0) + I(Y) - A(X) \quad -1 < \omega < 0 \quad (2.37)$$

$$= \mu(0) + I(X) - A(Y) \quad 0 < \omega < 1. \quad (2.38)$$

Since the largest A (3.62 eV, for Cl) is still smaller than the smallest I (3.89 eV, for Cs), the neutral ground state is stable.

2.5.6 Band gaps and discontinuity of exchange-correlation potential

A consequence of the results of the previous section is that the true band gap of a solid, $E_g = I - A$, can be written as

$$E_g = -\mu(N - \delta) + \mu(N + \delta) = \left. \frac{\delta E}{\delta n(\mathbf{r})} \right|_{N+\delta} - \left. \frac{\delta E}{\delta n(\mathbf{r})} \right|_{N-\delta} \quad (2.39)$$

with $\delta \rightarrow 0$.

Let us substitute to $E[n(\mathbf{r})]$ the explicit KS form, Eq.2.10. The Hartree and external potential terms of the functional will yield no discontinuity and no contribution to E_g . Only the kinetic and exchange-correlation terms may have a discontinuity and contribute to E_g .

For a non interacting system, only the kinetic term contributes, and the gap is exactly given by the KS gap:

$$E_g^{KS} = \left. \frac{\delta T_s}{\delta n(\mathbf{r})} \right|_{N+\delta} - \left. \frac{\delta T_s}{\delta n(\mathbf{r})} \right|_{N-\delta} = \epsilon_{LUMO} - \epsilon_{HOMO}. \quad (2.40)$$

We remark that even the kinetic energy of non interacting electrons, considered as a functional of the density, must have a discontinuous derivative when crossing an integer number of electrons. This is one reason why it is so difficult to produce explicit functionals of the charge density for T_s that are able to yield good results: no simple functional form will yield the discontinuity, but this is needed in order to get the correct energy spectrum.

For the interacting system:

$$E_g = \left. \frac{\delta T_s}{\delta n(\mathbf{r})} \right|_{N+\delta} - \left. \frac{\delta T_s}{\delta n(\mathbf{r})} \right|_{N-\delta} + \left. \frac{\delta E_{xc}}{\delta n(\mathbf{r})} \right|_{N+\delta} - \left. \frac{\delta E_{xc}}{\delta n(\mathbf{r})} \right|_{N-\delta} = E_g^{KS} + E_g^{xc}. \quad (2.41)$$

Note that the kinetic term is evaluated at the same charge density as for the non interacting system, so it coincides with the KS gap.

In conclusion: the KS gaps are not, by construction, equal to the true gap, because they are missing a term (E_g^{xc}) coming from the discontinuity of derivatives of the exchange-correlation functional. This is absent by construction from any current approximated functional (be it LDA or gradient-corrected or more complex). There is some evidence that this missing term is responsible for a large part of the band gap problem, at least in common semiconductors.

2.6 *Adiabatic continuation formula and the exchange-correlation hole

The exchange-correlation energy can be recast into a form that sheds some light on the unexpected success of LDA and gives a possible path for the production of better functionals. One considers a system in which the Coulomb interaction between electrons is adiabatically switched on:

$$U_\lambda = \lambda \frac{e^2}{2} \sum_{i,j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} = \lambda U \quad (2.42)$$

where λ is a parameter that goes from $\lambda = 0$, for the noninteracting system, to $\lambda = 1$, for the true interacting system. The charge density is forced to remain equal to the charge density of the interacting system:

$$n_\lambda(\mathbf{r}) = n(\mathbf{r}), \quad (2.43)$$

while the potential V_λ will depend on λ . At $\lambda = 0$ the potential is nothing but the KS potential: and the energy functional at $\lambda = 0$ has the simple form:

$$E_0 = T_s[n(\mathbf{r})] + \int n(\mathbf{r})V_{KS}(\mathbf{r})d\mathbf{r}. \quad (2.44)$$

The following step is to write the energy functional for the true interacting system as an integral of the derivative with respect to λ :

$$E_1 = E_0 + \int_0^1 \frac{dE_\lambda}{d\lambda} d\lambda. \quad (2.45)$$

The derivative can be simply expressed using the Hellmann-Feynman theorem:

$$\frac{dE_\lambda}{d\lambda} = \langle \Psi_\lambda | \frac{\partial H}{\partial \lambda} | \Psi_\lambda \rangle \quad (2.46)$$

(see section on Hellmann-Feynman forces for the demonstration). Explicitly:

$$\frac{dE_\lambda}{d\lambda} = \langle \Psi_\lambda | U | \Psi_\lambda \rangle + \langle \Psi_\lambda | \frac{\partial V_\lambda}{\partial \lambda} | \Psi_\lambda \rangle. \quad (2.47)$$

By performing the integration, one finally finds

$$E_{xc} = \frac{1}{2} \int \frac{f_{xc}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} n(\mathbf{r}) d\mathbf{r} d\mathbf{r}' \quad (2.48)$$

where $f_{xc}(\mathbf{r}, \mathbf{r}')$ is the *exchange-correlation hole*: the charge missing around a point \mathbf{r} due to exchange effects (Pauli antisymmetry) and to Coulomb repulsion. The exchange-correlation hole obeys the sum rule

$$\int f_{xc}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' = -1. \quad (2.49)$$

The exchange-correlation hole is related to the pair correlation function $g(\mathbf{r}, \mathbf{r}')$, giving the probability to find an electron in \mathbf{r}' if there is already one in \mathbf{r} . Its exact definition is:

$$f_{xc}(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}') \int_0^1 (g_\lambda(\mathbf{r}, \mathbf{r}') - 1) d\lambda \quad (2.50)$$

where $g_\lambda(\mathbf{r}, \mathbf{r}')$ is the pair correlation function the system having the electron-electron interaction multiplied by λ , Eq.(2.42). In homogeneous systems $f_{xc}(\mathbf{r}, \mathbf{r}')$ and $g(\mathbf{r}, \mathbf{r}')$ are well known and studied functions. It has been shown that in inhomogeneous systems LDA does not give a good approximation for $f_{xc}(\mathbf{r}, \mathbf{r}')$. However LDA yields a very good approximation for its spherical part $f_{xc}(\mathbf{r}, s)$:

$$\tilde{f}_{xc}(\mathbf{r}, s) = \int f_{xc}(\mathbf{r}, \mathbf{r} + s\hat{r}) \frac{d\hat{r}}{4\pi}. \quad (2.51)$$

It is easily shown the Eq.2.48 depends only on the spherical part of the exchange-correlation hole:

$$E_{xc} = \frac{1}{2} \int \frac{\tilde{f}_{xc}(\mathbf{r}, s)}{s} n(\mathbf{r}) d\mathbf{r} ds. \quad (2.52)$$

This explains at least partially the good performances of LDA. The above procedure is a good starting point in the search for better functional, via better modeling of the exchange-correlation hole. It also shows that the exact exchange-correlation energy contains a contribution coming from the difference between the true many-body kinetic energy $\langle \Psi | T | \Psi \rangle$ and the kinetic energy $T_s[n(\mathbf{r})]$ of non-interacting electrons, in addition to HF exchange plus correlation energy.

2.7 *The exact exchange-correlation potential from many-body theory

Many-body perturbation theory yields the following exact solution for the many-body problem:

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) - \epsilon_i \right) \psi_i(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}', \epsilon_i) \psi_i(\mathbf{r}') d\mathbf{r}' = 0 \quad (2.53)$$

where the *self-energy* $\Sigma(\mathbf{r}, \mathbf{r}', \epsilon)$ is a complex, nonlocal, energy-dependent operator, the $\psi_i(\mathbf{r})$ and ϵ_i have the physical meaning of quasiparticle states and energies. The energies ϵ_i are also complex and their imaginary part is related to the lifetime of the state.

Both DFT and many-body perturbation theory are exact on the ground state (and the latter also on excited states). This implies

$$n(\mathbf{r}) = \int \text{Im} G_{DFT}(\mathbf{r}, \mathbf{r}, \epsilon) d\epsilon = \int \text{Im} G(\mathbf{r}, \mathbf{r}, \epsilon) d\epsilon \quad (2.54)$$

where $G(\mathbf{r}, \mathbf{r}', \epsilon)$ is the Green's function of the system, $G_{DFT}(\mathbf{r}, \mathbf{r}', \epsilon)$ is the same in DFT, and the integration extends to the energies of occupied states. The Dyson equation must also apply between G and G_{DFT} :

$$G = G_{DFT} + G_{DFT} (\Sigma - V_{xc}) G. \quad (2.55)$$

By combining the above equations, one finally gets the following result:

$$\text{Im} \int [G_{DFT} (\Sigma - V_{xc}) G]_{\mathbf{r}=\mathbf{r}'} = 0. \quad (2.56)$$

This equation can be used to deduce the exact exchange-correlation potential. Practical many-body perturbation theory calculations are very difficult but not impossible. Some test calculations on simple systems have shown that the LDA V_{xc} is a good approximation to the true V_{xc} .

2.8 Bibliography

Density Functional Theory of Atoms and Molecules, R.G. Parr and W. Yang (Oxford University Press, New York, 1989).

General introduction to DFT

Density Functional Theory, R.M. Dreizler and E.K.U. Gross, Springer-Verlag, Berlin (1990).

More formal and deeper approach to DFT

R.O. Jones and O. Gunnarson, *Rev. Mod. Phys.* **61**, 689 (1989).

Performances of DFT and LDA in atoms.