Density Functional Theory
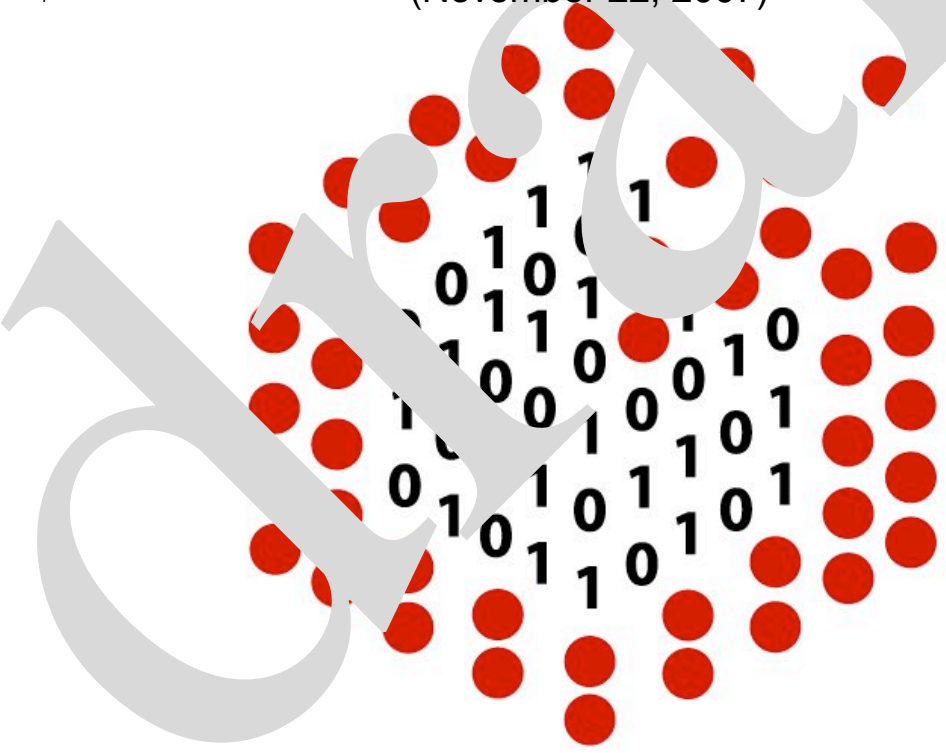
a different elementary introduction

Stefano Baroni

`baroni@sissa.it`

Scuola Internazionale Superiore di Studi Avanzati and
DEMOCRITOS National Simulation Center, Trieste, Italy

(November 22, 2007)

Notes of the lectures given at the school on
Electronic structure calculations and correlated materials
Les Houches, August 27 – September 8, 2006

Contents

1   Introduction

Density Functional Theory (DFT) has emerged as a *de facto* standard for simulating and modeling the properties of materials at the atomic ("*nano*") scale. In spite of all the successes met in forty years of DFT practice, some of its mathematical foundations are still poorly understood, and the limits of current approximations to it (not all of which can be systematically improved, not even in principle) barely begin to show in a clear way. The aim of the present short course is to provide an overview of the conceptual foundations of the theory, as well as a glimpse of its successes and failures, with an eye on the prospects to overcome its present limitations.

In a nutshell, DFT is a technique that allows to describe the properties of a system of interacting electrons in terms of its ground-state electron charge-density distribution, rather than of the external potential acting upon it. In a way that will be made more clear in Sec. 3, this change of variables closely parallels the switch between conjugate variables (volume and pressure, entropy and temperature, particle number and chemical potential, etc.) in elementary thermodynamics. The mathematics of thermodynamics is rooted in the properties of *Legendre transforms*. The possibility of formulating DFT in terms of Legendre transforms hardly comes as a surprise, and has in fact popped up in the literature ever since the early eighties [1, 2, 3, 4, 5, 6]. What is surprising is the minor attention generally paid to such a formulation which is in fact largely overlooked. I do not know if this lack of attention is due to some fundamental mathematical flaws in the arguments. I do not believe so. In any event, I do believe that the analogies between DFT and thermodynamics serve the purpose of demystifying the *magic* of DFT and shed light on some deep, fundamental aspects of the theory. Furthermore, these analogies are *beautiful* (or at least, so they look to me) and lend themselves to the introduction of some general mathematical concepts which are ubiquitous in theoretical physics and hardly properly taught in curricular courses.

My lectures will be divided in three rather independent parts. In the first, I will introduce and briefly discuss some basic mathematical concepts, such as *functionals*, *variational principles*, a generalized version of the *Hellmann-Feynman theorem* [7, 8], *Legendre transforms*, etc. The second part will be devoted to an introduction of DFT along the lines sketched above. No attempt will be made at mathematical rigor (nor would I be able to even try!). Rather, I will try to highlight the analogies between the mathematical structure of DFT and that of elementary thermodynamics, without entering (nor even mentioning, most of the times!) the many mathematical subtleties that arise in basing DFT entirely on such similarities. In the last part, I will present a (very limited) sample of applications of DFT to materials modeling, and try to contrast its many successes with some of the failures which still need to be overcome.

3

I wish the students will enjoy these lectures (and learn from them!) no less than I enjoyed their preparation (and learned from it!).

1.1 Disclaimer

I spent quite some time writing these notes, but the time to leave for Les Houches inexorably came before I could complete them (or at least, complete them the way I had wished in the first place). A few hours before takeoff, these notes still lack any reference to actual applications, which I wish I could include. All of the sections need revision to some extent. This is particularly so for the last few sections where this need is explicitly stated (whereas for the others it is simply to be understood!). The literature is also way incomplete and inhomogeneous.

In spite of all these limitations, I wish you will find these notes useful, but I would like to urge you to treat them as what they are: the draft of a work still to be completed. Do not blindly trust each and every statement you find in these pages (nor should you do so for any, even most respectable, scientific text). If you find typos, errors, inconsistencies, or obscure statements, please send me your suggestions and complaints: I will incorporate them in the (hopefully) soon-to-be-completed version. Meanwhile, you may want to check from time to time for up-to-date versions at

http://www.democritos.it/~baroni

2 Mathematical preliminaries

2.1 What is a functional?

A real-valued function, $F : \Omega \mapsto \mathbb{R}$, is a binary relation that associates to every element $\omega \in \Omega$ a well defined real number, $y = F(\omega)$. Suppose that Ω is some *functional space*, $\mathcal{F}$, so that its elements are real-valued functions of a real argument, $f : \mathbb{R} \mapsto \mathbb{R}$. Elements of $\mathcal{F}$ will be called *test functions*. For instance, $\mathcal{F}$ could be the (linear) space of polynomials up to some order n, or the set of functions integrable on some domain, etc. Then a *function* $F : \mathcal{F} \mapsto \mathbb{R}$ is called a *functional*. Trivial examples of functionals are:

1. $F[f] = f(x_0)$ (the value that the test function has at a given point, x_0).

2. $F[f] = \int_a^b f(x)dx$ (the integral of the test function over a given interval).

3. $F[f] = \int_a^b f'(x)^2 dx$ (the integral of the squared derivative of the test function over a given interval).

4. $F[f] = \int_a^b \mu(f(x))g(x)dx$ (μ and g are here generic real-valued functions), $\cdots$

5. $\cdots$ or any other rule that would allow one to univocally associate a real number to a test function, $f(x)$.

Complex-valued ($F : \mathcal{F} \mapsto \mathbb{C}$) or vector-valued ($F : \mathcal{F} \mapsto \mathbb{R}^n$) functionals are obvious generalizations of the concept of real-valued functionals, and so are multivariate test functions.

2.1.1 Functional derivatives

Let us suppose that the test functions, $f \in \mathcal{F}$, are defined on a finite interval $I = [a, b]$. Then, under rather general smoothness conditions, each one of them can be well represented by the set of values it has on a discrete mesh: $\{x_i = a + \frac{i}{N}(b - a)\} \subset I$, for $i = 1, 2, \ldots N$: $\{f_i = f(x_i)\}$. Such a discrete representation of real-valued functions is common practice in scientific computing where it is used, *e.g.*, to calculate numerical derivatives and integrals, to solve differential equations, etc. When a discrete representation of test functions is adopted, a *functional* is simply represented by an ordinary function of many variables: $F[f] \simeq \bar{F} : \mathbb{R}^N \mapsto \mathbb{R}$. In this sense, a functional may be simply thought as a function of infinitely many variables, $f(x)$, labeled by a continuous index, x.

¹ Let us now suppose that $\bar{F}$ is differentiable, and let us consider its Taylor
² expansion:

$$\bar{F}(f + \Delta) = \bar{F}(f) + \sum_i \Delta_i \bar{F}_i' + \mathcal{O}\left(\Delta^2\right), \tag{1}$$

³ where $\Delta_i = \Delta(x_i)$ and $\bar{F}_i' = \frac{\partial \bar{F}}{\partial f_i}$ are the partial derivatives of $\bar{F}$. The sum
⁴ appearing in Eq. (1) is a discrete approximation of the integral of $\Delta(x)$ times
⁵ a function that we *define* as the *functional derivative* of F, $F'(x) \equiv \frac{\delta F}{\delta f(x)}$:

$$\int_a^b F'(x)\Delta(x)dx \approx h \sum_i \Delta_i F'(x_i) + \mathcal{O}\left(n^{-1}\right), \tag{2}$$

⁶ where $h = \frac{b-a}{N}$. By comparing Eq. (1) with Eq. (2), we arrive at the conclusion
⁷ that:

$$F'(x_i) \approx \frac{1}{h}\bar{F}_i'. \tag{3}$$

⁸ Eq. (3) is sufficient in many cases to calculate functional derivatives. Let us for
⁹ instance calculate the derivative of the functional in Example 4 of Sec. 2.1:

$$\bar{F} = h \sum_i \mu(f_i)g_i, \tag{4}$$

$$\bar{F}_i' = h\mu'(f_i)g_i. \tag{5}$$

¹⁰ By using Eq. (3) we readily obtain:

$$F'(x) = \mu'(f(x))g(x). \tag{6}$$

¹¹ When the definition of the functional contains derivatives of the test function,
¹² such as in Example 3 of Sec. 2.1, this procedure can also be applied by discretizing
¹³ the functional with an appropriate finite-difference approximation of the derivative
¹⁴ ($f'(x_i) \approx \frac{f_i - f_{i-1}}{h}$, or the like). However, it is often more convenient in these cases
¹⁵ to use a definition of the functional derivative in terms of a functional Taylor
¹⁶ expansion. Let f and g be two test functions and ϵ a real number. The functional
¹⁷ derivative, $F'(x) = \frac{\delta F}{\delta f(x)}$, can be *defined* by the relation:

$$F[f + \epsilon g] = F[f] + \epsilon \int F'(x)g(x)dx + \mathcal{O}\left(\epsilon^2\right). \tag{7}$$

¹⁸ Let us apply this definition to the functional of Example 3 of Sec. 2.1, and calcu-
¹⁹ late its derivative. In order to simplify the algebra, we suppose that the functional
²⁰ is defined over a space of test functions that vanish at the end points of the in-
²¹ terval $[a, b]$ (this would correspond to the usual definition of the kinetic-energy

functional in quantum mechanics):

$$
\begin{aligned}
F[f + \epsilon g] &= \int_a^b (f'(x) + \epsilon g'(x))^2 dx \\
&= F[f] + 2\epsilon \int_a^b f'(x)g'(x)dx + \mathcal{O}\left(\epsilon^2\right) \\
&= F[f] + 2\epsilon \left[g(x)f'(x)\big|_a^b - \int_a^b g(x)f''(x)dx \right].
\end{aligned}
\tag{8}
$$

Using the boundary conditions, $g(a) = g(b) = 0$, we readily see that $F'(x) = -2f''(x)$.

All in all, a useful rule of thumb for calculating functional derivatives is the following:

$$
\frac{\delta F}{\delta f(y)} = \lim_{\epsilon \to 0} \frac{F[f(x) + \epsilon\delta(x - y)] - F[f(x)]}{\epsilon}.
\tag{9}
$$

Notice that y is the actual argument of the functional derivative (which is a function of a real argument itself), whereas x is the dummy argument of the argument of the functional, $f(x)$.

2.2 Variational principles

Many fundamental laws of nature can be formulated in terms of *variational principles* which assert that the variables describing the state of a system (*state variables*) minimize some appropriate functions (the *state functions*). Such functions may depend on one or more parameters which describe the interaction of the system with the external world (*control parameters*). It is often useful to understand how the state variables, as well as the value of the state function at its minimum, depend on control parameters. Both the state variables and the control parameters may be *functions* (and the state function would then be a *functional*), but for simplicity in the following discussion we will assume that they are simply real numbers.

insert a few examples here

Let us indicate by x and v the state variable and external parameter, respectively, and let $f(x, v)$ be the state function whose minima determine the physical states of the system as functions of v. If $f(x, v)$ has a minimum with respect to x, its position is determined by the equation:

$$
\frac{\partial f(x, v)}{\partial x} = 0.
\tag{10}
$$

If the function f is *convex* with respect to x, *i.e.* $\frac{\partial^2 f(x,v)}{\partial x^2} > 0$ all over its domain of definition, the solution of Eq. (10), if any, is unique (see Sec. 2.3.2). If f were

₁ concave, rather than convex, the same would hold for its maximum. Let us now
₂ indicate the solution of Eq. (10) by $x(v)$. The value of the minimum as a function
₃ of v defines a function:

$$F(v) = \min_x f(x, v) \tag{11}$$
$$\equiv f(x(v), v) \tag{12}$$

₄ that describes the physical states of the system corresponding to different values
₅ of the control parameter v.

2.2.1 The (generalized) Hellmann-Feynman theorem

₇ It is often useful to study the derivative of the value of the minimum, $F(v)$,
₈ with respect to the control parameter v. For instance, the ground-state energy
₉ of a molecule can be thought as the minimum of the Schrödinger functional
₁₀ $\langle \Psi | \hat{H} | \Psi \rangle / \langle \Psi | \Psi \rangle$ with respect to the wave-function, Ψ, where $\hat{H}$ is the Hamil-
₁₁ tonian. In the Born-Oppenheimer approximation, the molecular Hamiltonian de-
₁₂ pends parametrically on the nuclear coordinates, $\hat{H} = \hat{H}(\mathbf{R})$, and the derivatives
₁₃ of the minimum of the functional would yield in this case the forces acting on in-
₁₄ dividual atoms. The expectation value of the Hamiltonian depends on nuclear po-
₁₅ sitions *explicitly*, because the molecular Hamiltonian does so through the external
₁₆ potential, and *implicitly*, through the ground-state wave-function which obviously
₁₇ depends on nuclear coordinates. The Hellmann-Feynman (HF) theorem [7, 8]
₁₈ states that the forces acting on atoms are the expectation values of the derivatives
₁₉ of the external potential, and hence the implicit dependence of the energy on the
₂₀ nuclear coordinates does not contribute to the derivative. The validity of this *the-*
₂₁ *orem* is not limited to the quantum-chemical framework in which it was originally
₂₂ formulated. The HF theorem states instead a general property of the solution of
₂₃ any variational problem.
₂₄ Let us calculate the derivative of $F(v)$ using Eq. (12):

$$F'(v) = \left.\frac{\partial f(x, v)}{\partial x}\right|_{x=x(v)} x'(v) + \left.\frac{\partial f(x, v)}{\partial v}\right|_{x=x(v)} \tag{13}$$
$$= \left.\frac{\partial f(x, v)}{\partial v}\right|_{x=x(v)}, \tag{14}$$

₂₅ where the first term in Eq. (13) vanishes because of the extremum condition,
₂₆ Eq. (10). In a nutshell, the HF theorem states that the derivative of the extrema
₂₇ of a function with respect to the control parameters upon which it may depend is
₂₈ only determined by the explicit dependence on them, and not by the implicit one.

2.2.2 Constrained minimization and the method of Lagrange multipliers

Let us suppose that a function, $f(X)$, depending on N variables, $X = \{x_1, x_2, \cdots x_N\}$, has to be minimized subject to the constraint that P auxiliary functions have preassigned values: $g_i(X) = \gamma_i$, with $i = 1, \cdots P$. Each one of these equations is satisfied on a manifold of dimension $N - 1$. The dimension of the intersection of all the P such manifolds (one for each constraint) is $N - P$. The linear manifold orthogonal at a point X to this intersection space has dimension P, and points belonging to it are linear combinations of the gradients of the P constraining functions at that point: $\sum_i \lambda_i \frac{\partial g_i}{\partial X}$. The condition that $f(X)$ is extremum subject to the constraint that X is varied within the intersection manifold amounts to requesting that the gradient of f belongs to the orthogonal manifold:

$$\frac{\partial f}{\partial X} = \sum_i \lambda_i \frac{\partial g_i}{\partial X}. \tag{15}$$

Eq. (15) can be lumped together with the P equations of the constraints by requiring that the auxiliary function:

$$F(X, \Lambda; \Gamma) = f(X) - \sum_i \lambda_i (g_i(X) - \gamma_i), \tag{16}$$

where $\Lambda = \{\lambda_1, \lambda_2 \cdots \lambda_P\}$ and $\Gamma = \{\gamma_1, \gamma_2 \cdots \gamma_P\}$, is stationary with respect to *unconstrained variations* of the $N + P$ variables X and Λ. Variation of the λ's gives the P equations of the constraints, whereas variation of the x's gives the parallelity condition, (15). Note that the stationary point is not in general a minimum nor a maximum, but just a saddle point. The resulting extremum is a function of the Γ constants which appear in the variational problem as parameters:

$$\Phi(\Gamma) = \operatorname*{extr}_{\{X,\Lambda\}} F(X, \Lambda; \Gamma). \tag{17}$$

The Hellmann-Feynman theorem and the definition of the F function, Eq. (16), show that the partial derivatives of the extremum with respect to the constraining constants are the corresponding Lagrange multipliers:

$$\frac{\partial \Phi(\Gamma)}{\partial \gamma_i} = \lambda_i. \tag{18}$$

2.3 Legendre transforms

Let us consider a function $F(v)$ which we suppose to be concave ($F''(v) < 0$; convexity would also do—the important thing is that F'' does not change sign). The auxiliary function

$$g(v, p) = F(v) - vp \tag{19}$$

has an extremum with respect to v for

$$F'(v) = p. \tag{20}$$

This extremum is a maximum because

$$\frac{\partial^2 g}{\partial v^2} = F''(v) < 0, \tag{21}$$

by virtue of the concavity of F. Concavity also implies that the solution of Eq. (20), if any, is unique. This property is easily seen by graphic inspection (see Fig. 1), and it will be stated more formally in the case of multivariate transforms in Sec. 2.3.2.

The value of the maximum of $g(v, p)$ with respect to v defines a function of p,

$$G(p) = \max_v \big(F(v) - vp\big) \tag{22}$$
$$\equiv F(v(p)) - pv(p), \tag{23}$$

where $v(p)$ is the solution of Eq. (20), which is called the *Legendre transform* of $F(v)$. The function f depends on p as a parameter, hence the derivative of $G(p)$ is given by the HF theorem as the value of the partial derivative of $f(v, p)$, calculated at the maximum (see Sec. 2.2.1):

$$G'(p) = \left.\frac{\partial g(v, p)}{\partial p}\right|_{v=v(p)}$$
$$= -v. \tag{24}$$

It follows that the Legendre transform of a concave function is convex, because $G''(p) = -\frac{dv}{dp} = -1/F''(v) > 0$.

The Legendre transform of a Legendre transform equals the original function:

$$\begin{aligned} H(v) &= G(p(v)) - p(v)G'(p(v)) \\ &= G + pv \\ &= F(v). \end{aligned} \tag{25}$$

Convexity of the Legendre transform, together with Eq. (24), implies that the original function satisfies a variational principle analogous to Eq. (22), in terms of its Legendre transform:

$$F(v) = \min_p \big(G(p) + pv\big). \tag{26}$$

The uniqueness of the solution of Eq. (20) implies that there is a one-to-one correspondence between the v and p variables which can therefore be both assumed
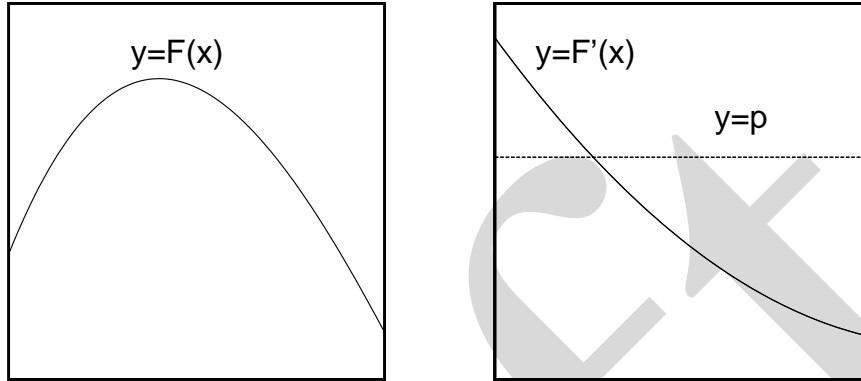
10

Figure 1: The left panel features a concave function ($F''(x) < 0$), the right panel its first derivative (full line). The solution of the equation $F'(x) = p$ (dotted line), if any, is unique.

1 as state variables of the system. The variables v and p are said to be *conjugate* to
2 each other.

3 Let us now assume that the function F depends on a control parameter λ:
4 $F = F(v, \lambda)$. The Legendre transform of F will also depend on λ:

$$G(p, \lambda) = F(v(p, \lambda), \lambda) - pv(p, \lambda), \tag{27}$$

5 where $v(p, \lambda)$ indicates the λ-dependent solution of Eq. (20). Because the Leg-
6 endre obeys a variational principle, Eq. (22), and because of the HF theorem (see
7 Sec. 2.2.1), the partial derivative of G with respect to the control parameter λ has
8 only contributions from the explicit dependence of $g(v, p, \lambda)$—see Eq. (19)—upon
9 λ:

$$\frac{\partial G(p, \lambda)}{\partial \lambda} = \frac{\partial F(v, \lambda)}{\partial \lambda}. \tag{28}$$

10 Eq. (28) shows that G as a function of p depends on external parameters the same
11 way as F as a function of v. For this reason, F and G are considered to be *natural*
12 functions of v and p, respectively.
13 insert here a few examples

2.3.1 Multivariate Lagendre transforms

15 The concept of Legendre transform can be readily generalized to functions of
16 many variables (*multivariate* Legendre transforms). Let F(V) be a concave func-
17 tion of N variables, $V = \{v_1, v_2, \cdots v_N\}$, with $F''_{ij} \equiv \frac{\partial^2 F}{\partial v_i \partial v_j} < 0$ (here and in the
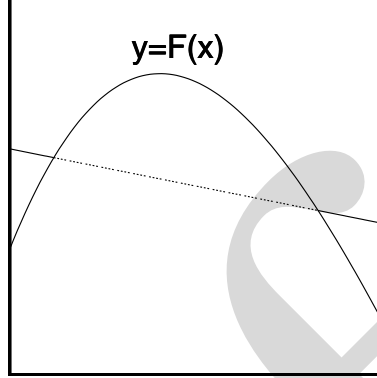
11

Figure 2: A concave function, $y = F(x)$, can intercept a straight line in two points at most. The value of the function in between the two intercepts is larger than the value of the linear function represented by the straight line. This condition can be expressed in general as: $F\big(\alpha x_1 + (1 - \alpha)x_2\big) > \alpha F(x_1) + (1 - \alpha)F(x_2)$ for $0 < \alpha < 1$. This condition naturally extends to concave functions of many variables.

following "<" and ">", as referred to matrices and operators, indicate negative and positive definiteness, respectively). The function $f(V, P) = F(V) - \sum_i v_i p_i$ has a maximum for

$$F_i' \equiv \frac{\partial F}{\partial v_i} = p_i. \tag{29}$$

The value of the maximum, as a function of $P = \{p_1, p_2, \cdots p_N\}$, is the Legendre transform of F:

$$G(P) = F(V(P)) - \sum_i p_i v_i(P), \tag{30}$$

where the v's are functions of the p's, as defined by the solution of Eq. (29).

2.3.2 Uniqueness of the Legendre transform

As it was the case for univariate transforms, in the multivariate case concavity (or convexity) guarantees the uniqueness of the solution of Eq. (29) and, hence, of the Legendre transform. One possible demonstration of this proceeds via *reductio ad absurdum*. Let us suppose that Eq. (29) has two solutions, $V_1 \neq V_2$, and let us show that this contradicts the hypothesis that $F(V)$ is concave. Concavity requires that:

$$F(\alpha V_1 + (1 - \alpha)V_2) > \alpha F(V_1) + (1 - \alpha)F(V_2) \tag{31}$$

12

1 for any $0 < \alpha < 1$ (see Fig. 2). Summing this inequality with the equivalent one
2 obtained by interchanging V_1 with V_2, one obtains:

$$F\left(V_1 + \alpha \Delta V\right) + F\left(V_2 - \alpha \Delta V\right) > F\left(V_1\right) + F\left(V_2\right), \tag{32}$$

3 where $\Delta V = V_2 - V_1$. We now expand this inequality up to second order in α.
4 The zero-order terms cancel on both sides; first-order terms vanish because by
5 hypothesis $F_i'\left(V_1\right) = F_i'\left(V_2\right)$; higher-order terms give:

$$\sum_{ij} \left(\left.\frac{\partial F}{\partial v_i \partial v_j}\right|_{V=V_1} + \left.\frac{\partial F}{\partial v_i \partial v_j}\right|_{V=V_0} \right) \Delta v_i \Delta v_j + \mathcal{O}(\alpha) > 0. \tag{33}$$

6 For small enough α, the term $\mathcal{O}(\alpha)$ can be neglected, and Eq. (33) contradicts the
7 fact that the Hessian matrix of a concave function is negative definite.

13

3 Density-functional theory

3.1 The Schrödinger "potential" functional

Much of elementary quantum mechanics can be derived from the Ralyleigh-Ritz variational principle which states that the Schrödinger functional,

$$E_S[\Psi] = \langle \Psi | \hat{H} | \Psi \rangle, \tag{34}$$

is stationary around the eigenfunctions of the Schrödinger equation, subject to the constraint that the eigenfunctions be normalized:

$$\frac{\delta E_S}{\delta \Psi} = 0, \langle \Psi | \Psi \rangle = 1 \quad \Rightarrow \quad \hat{H}\Psi = \epsilon\Psi, \tag{35}$$

where the energy eigenvalue, ϵ, is the Lagrange multiplier corresponding to the normalization constraint. In order to keep the notation as simple as possible, we assume that the wave-functions are real, which is always permitted whenever the Hamiltonian is time-reversal invariant. The ground-state wave-function, Ψ_0, is the global (constrained) minimum of the Schrödinger functional.

Let us now specialize to a system of N interacting electrons (Hartree atomic units, $\hbar = m = e = 1$, will be used throughout):

$$\hat{H} = -\frac{1}{2} \sum_i \frac{\partial^2}{\partial \mathbf{r}_i^2} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V(\mathbf{r}_i). \tag{36}$$

The first two terms in Eq. (36) are the kinetic-energy and electron-electron (*e-e*) interaction operators which will be indicated by $\hat{K}$ and $\hat{W}$, respectively, while the third term is the external (one-electron) potential, $\hat{V}$.

The very existence of a ground state for the Hamiltonian of many-electron systems, Eq. (36), indicates that its energy is a *functional* of the external potential:

$$E^\circ[V] = \min_\Psi E_S[\Psi; V], \quad \langle \Psi | \Psi \rangle = 1, \tag{37}$$

where the parametric dependence of the Schrödinger functional, Eq. (34), upon the external potential, V, has been explicitly indicated. The functional of Eq. (37) will be called the *potential functional*. Its actual evaluation would of course require the solution of the many-body Schrödinger equation:

$$\left[-\frac{1}{2} \sum_i \frac{\partial^2}{\partial \mathbf{r}_i^2} + \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V(\mathbf{r}_i) \right] \Psi(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N)$$
$$= E^\circ[V]\, \Psi(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N). \tag{38}$$

14

¹ In order to proceed further, let us introduce the *electron charge-density distri-*
² *bution*, $\rho(\mathbf{r})$, as the ground-state expectation value of the electron charge-density
³ operator, $\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$:

$$
\begin{aligned}
\rho(\mathbf{r}) &= \langle \Psi | \hat{\rho}(\mathbf{r}) | \Psi \rangle && (39) \\
&= N \int |\Psi(\mathbf{r}, \mathbf{r}_2, \cdots \mathbf{r}_N)|^2 d\mathbf{r}_2 \cdots d\mathbf{r}_N. && (40)
\end{aligned}
$$

⁴ In order to simplify the notation, in Eq. (40), as well as in the rest of these notes,
⁵ electrons will be assumed to be spinless, unless otherwise explicitly stated. Notice
⁶ that the argument of $\rho(\mathbf{r})$, $\mathbf{r}$, is a parameter which labels the $\hat{\rho}$ operator at different
⁷ points of space (strictly speaking, $\hat{\rho}$ is a collection of Hermitean operators, one
⁸ for each point in the 3D space). In terms of the electron density distribution, the
⁹ Schrödinger functional reads:

$$
\begin{aligned}
E_S[\Psi; V] &= \langle \Psi | \hat{K} + \hat{W} | \Psi \rangle + \int V(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} && (41) \\
&\equiv F_S[\Psi] + \int V(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}. && (42)
\end{aligned}
$$

¹⁰ Eq. (42) shows that the Schrödinger functional depends on the external potential,
¹¹ $V(\mathbf{r})$, through a linear coupling to the electron charge density. The functional
¹² derivative of $E_S[\Psi; V]$ with respect to V is ρ:

$$
\frac{\delta E_S}{\delta V(\mathbf{r})} = \rho(\mathbf{r}). \qquad (43)
$$

¹³ As the potential functional is defined in terms of a variational principle, Eq. (37),
¹⁴ the Hellmann-Feynman theorem can be used to equate the functional derivative of
¹⁵ E° with respect to V to the partial (functional) derivative of E_S with respect to V:
¹⁶

$$
\frac{\delta E^\circ}{\delta V(\mathbf{r})} = \frac{\delta E_S}{\delta V(\mathbf{r})} = \rho(\mathbf{r}). \qquad (44)
$$

¹⁷ The above considerations hold unambiguously whenever the ground state is
¹⁸ non-degenerate. In case of degeneracy, the correspondence $V \mapsto \rho$ is not univo-
¹⁹ cal. Univocality is restored by considering $E^\circ[V]$ as the zero-temperature limit
²⁰ of the system's free energy. In this case, the relevant *ground-state electron den-*
²¹ *sity distribution* is the average of the distributions corresponding to the various
²² components of the ground-state multiplet, and Eq. (44) still holds.

²³ 3.1.1 Concavity of the potential functional

²⁴ The definition of the potential functional in terms of a variational principle, Eq.
²⁵ (37), naturally entails concavity [1]. Let $V_\alpha = \alpha V_1 + (1 - \alpha) V_0$, and $\hat{H}_\alpha =$

$\alpha \hat{H}_1 + (1 - \alpha)\hat{H}_0$ and Ψ_α° the corresponding Hamiltonian and ground-state wave-function. One has that:

$$
\begin{aligned}
E^\circ[V^\alpha] = \langle \Psi_\alpha^\circ | \hat{H}_\alpha | \Psi_\alpha^\circ \rangle &= \alpha \langle \Psi_\alpha^\circ | \hat{H}_1 | \Psi_\alpha^\circ \rangle + (1 - \alpha) \langle \Psi_\alpha^\circ | \hat{H}_0 | \Psi_\alpha^\circ \rangle \\
&> \alpha E^\circ[V_1] + (1 - \alpha) E^\circ[V_0],
\end{aligned}
\tag{45}
$$

where the inequality of the second line comes from the fact that the Hamiltonian expectation values in the previous line are calculated with respect to a "wrong" wave-function.

It is instructive to derive the same result as a property of the second (functional) derivative of E°. According to Eq. (44), the second derivative of E° with respect to V is the derivative of ρ with respect to V, *i.e.* the density linear-response function, or polarizability, χ:

$$
\frac{\delta^2 E^\circ}{\delta V(\mathbf{r}) \delta V(\mathbf{r}')} = \frac{\delta \rho(\mathbf{r})}{\delta V(\mathbf{r}')}
\tag{46}
$$

$$
\equiv \chi(\mathbf{r}, \mathbf{r}').
\tag{47}
$$

The polarizability, as a linear operator, is negative definite. To see this, let us make a short detour to perturbation theory and consider the Hamiltonian:

$$
\hat{H}_\lambda = \hat{H}_0 + \lambda \hat{V}',
\tag{48}
$$

where $\hat{V}'$ is a local operator: $\hat{V}' = \sum_i V'(\mathbf{r}_i)$. The Hellmann-Feynman theorem states that:

$$
\frac{dE_\lambda}{d\lambda} = \langle \Psi_\lambda | \hat{V}' | \Psi_\lambda \rangle
\tag{49}
$$

$$
= \int \rho_\lambda(\mathbf{r}) V'(\mathbf{r}) d\mathbf{r}.
\tag{50}
$$

By differentiating Eq. (50), one obtains:

$$
\frac{d^2 E_\lambda}{d\lambda^2} = \int \rho_\lambda'(\mathbf{r}) V'(\mathbf{r}) d\mathbf{r},
\tag{51}
$$

where:

$$
\rho_\lambda'(\mathbf{r}) = \frac{\partial \rho_\lambda(\mathbf{r})}{\partial \lambda}
\tag{52}
$$

$$
= \int \chi(\mathbf{r}, \mathbf{r}') V'(\mathbf{r}') d\mathbf{r}'.
\tag{53}
$$

By inserting Eq. (53) into a Taylor expansion of the ground-state energy in powers of λ, we can express the second-order correction to the energy in terms of the polarizability:

$$
E^{(2)} = \frac{1}{2} \lambda^2 \int \chi(\mathbf{r}, \mathbf{r}') V'(\mathbf{r}) V'(\mathbf{r}') d\mathbf{r} d\mathbf{r}'.
\tag{54}
$$

The second-order correction to the ground-state energy, $E^{(2)}$, can also be obtained from standard perturbation theory:

$$E^{(2)} = \lambda^2 \sum_n \frac{|\langle \Psi_n | \hat{V}' | \Psi_0 \rangle|^2}{\epsilon_0 - \epsilon_n}. \tag{55}$$

Whenever the ground state is non degenerate all the denominators of Eq. (55) are strictly negative and the second order correction to the ground-state energy is thus strictly negative. By comparing Eq. (54) with the sign of Eq. (55), we conclude that the polarizability is negative definite, and the potential functional is therefore *concave* [1, 9].

3.2 The Hohenberg-Kohn density functional

3.2.1 The first Hohenberg-Kohn theorem

In Sec. 2.3.2 we have seen that, whenever a function $F(X)$ is concave (or convex, for that matter), the mapping between its independent variables and its partial derivatives, $X \mapsto \frac{\partial F}{\partial X}$ is invertible. When applied to the potential functional of Sec. 3.1, this property amounts to the *first Hohenberg-Kohn theorem* (HK) [10] which states that the mapping $V \mapsto \rho$ is invertible, *i.e.* the ground-state electron-density distributions of two systems with a same number of electrons in the field of two different external potentials (*i.e.* such that their difference is not a constant) are necessarily different.

The potential functional, of Sec. 3.1 can be written as:

$$E^\circ[V] = \langle \Psi_{[V]} | \hat{K} + \hat{W} | \Psi_{[V]} \rangle + \int \rho_{[V]}(\mathbf{r}) V(\mathbf{r}) d\mathbf{r}, \tag{56}$$

where $\Psi_{[V]}$ is the solution of Eq. (38), resulting from the variational principle of Eq. (37). Using Eq. (44), the Legendre transform of $E^\circ[V]$ reads:

$$F[\rho] = \langle \Psi_{[\rho]} | \hat{K} + \hat{W} | \Psi_{[\rho]} \rangle, \tag{57}$$

where $\Psi_{[\rho]}$ is the ground-state wave-function generated by the external potential corresponding to ρ, which is unique according to the first HK theorem. $F[\rho]$ is thus a *universal* functional of the density (*i.e* independent of the external potential V) and it is called the *Hohenberg-Kohn functional*.

3.2.2 The Levy-Lieb density functional

Let us now define the Levy-Lieb (LL) density functional as the minimum of the kinetic plus *e-e* interaction operators over the set of many-body wave-functions

which correspond to a same given density, $\rho(\mathbf{r})$:

$$F_{LL}[\rho] = \min_{\Psi}\langle\Psi|\hat{K} + \hat{W}|\Psi\rangle; \quad \langle\Psi|\hat{\rho}(\mathbf{r})|\Psi\rangle = \rho(\mathbf{r}). \tag{58}$$

This constrained variational problem can be solved using the method of Lagrange multipliers of Sec. 2.2.2:

$$\delta\left(\langle\Psi|\hat{K} + \hat{W}|\Psi\rangle + \int V(\mathbf{r})[\langle\Psi|\hat{\rho}(\mathbf{r})|\Psi\rangle - \rho(\mathbf{r})]d\mathbf{r}\right) = 0, \tag{59}$$

where $V(\mathbf{r})$ plays the role of a (functional) Lagrange multiplier. As usual, variation with respect to the Lagrange multipliers gives the equations for the constraints [see Eq. (58)], whereas variation with respect to Ψ gives:

$$\left(\hat{K} + \hat{W} + \int V(\mathbf{r})\hat{\rho}(\mathbf{r})d\mathbf{r}\right)\Psi = 0, \tag{60}$$

which is the Schrödinger equation for a system of interacting electrons in the external potential $V(\mathbf{r})$ whose arbitrary reference energy has been chosen so as to make the eigenvalue vanish. This shows that, whenever the minimum of Eq. (58) exists, the value of the LL functional equals the expectation value of $\hat{K} + \hat{T}$ evaluated over the ground-state wave-function that has ρ as density distribution, and hence $F_{LL}[\rho] = F_{HK}[\rho]$. The minimum may not exist for some $\rho(\mathbf{r})$. In this case, the HK functional is not defined, which means that $\rho(\mathbf{r})$ is not the ground-state charge density of any physical system ($\rho(\mathbf{r})$ is said in this case to be not *v-representable*).

The conditions that a density has to fulfill in order to be v-representable are poorly known. The concavity of the polarizability, Sec. 3.1.1, can be used to show that if a density is v-representable, any normalized function sufficiently close to it is also v-representable [9]. When a density is not v-representable, the LL functional can be equally defined as the *infimum* of $\hat{K} + \hat{W}$, rather than the minimum, provided that ρ is at least 'N-*representable*' (*i.e.* it can be expressed by Eq. (40) for some antisymmetric wave-function, Ψ, not necessarily the ground state of any physical system).

3.2.3 The second Hohenberg-Kohn theorem

According to the results of Sec. 2.3, the original *potential* functional, $E^\circ[V]$, can be obtained as the *inverse Legendre transform* of the density functional, $F[\rho]$. In particular, using the variational principle expressed by Eq. (26), one has:

$$E^\circ[V] = \min_{\rho}\left(F[\rho] + \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}\right) \tag{61}$$

$$= F\left[\rho_{[V]}\right] + \int V(\mathbf{r})\rho_{[V]}(\mathbf{r})d\mathbf{r}, \tag{62}$$

18

1 where $\rho_{[V]}$ is the density distribution that solves the variational problem, Eq. (61).

2 Eq. (61) is usually referred to as the *second Hohenberg-Kohn theorem*. Its differ-

3 ential form usually includes a Lagrange multiplier, μ:

$$\frac{\delta F}{\delta\rho(\mathbf{r})} + V(\mathbf{r}) = \mu, \tag{63}$$

4 which would account for the constraint on the total number of electrons: $\int \rho(\mathbf{r})d\mathbf{r} =$

5 N. The value of μ can be incorporated in the definition of V if the variation of ρ

6 is restricted to the set of N-representable densities, which has to be done in any

7 case (otherwise, nor the HK, nor even the more general LL functional, would be

8 defined).

3.3 The Kohn-Sham equations

10 The theory of Hohenberg and Kohn, as encompassed in the two *theorems* of sec-

11 tion 3.2, would be almost trivial if all the arguments based on the use of Legendre

12 transforms could be made rigorous. Unfortunately this is not the case and many

13 important properties of the energy functional, including its very domain of def-

14 inition, are poorly known, not to mention our inability to compute it from first

15 principles, other than solving the Schrödinger equation (which is exactly what

16 DFT was designed to avoid in the first place). DFT would thus be kind of void if

17 workable and sufficiently accurate approximations to the density functional were

18 not available. The fundamental paper of Kohn and Sham (KS) [12] provides a

19 path to an entire class of such approximations.

20 The central idea of Kohn and Sham is to subtract from the HK (or LL) density

21 functional any sensible contribution to it that is *i)* physically motivated and *ii)*

22 easily calculable. By doing so one would confine all of our ignorance into the

23 difference, in the hope that it is small and easy to approximate accurately. The

24 HK density functional is the sum of a kinetic plus an *e-e* interaction term. When

25 applied to a system of non interacting electrons, the *first HK theorem* states that the

26 kinetic energy of its ground state, T_0, is a well defined functional of its ground-

27 state density: $\rho \mapsto V \mapsto \Psi \mapsto T_0[\rho]$. $T_0[\rho]$ is in fact the HK (or LL, if one

28 prefers) density functional for a system of non interacting electrons. Of course,

29 T_0 is *not* the kinetic energy of any system of interacting electrons, but it can be

30 used in the KS construction as a first approximation to it. Analogously, we can

31 separate out of the *e-e* interaction energy a classical contribution, corresponding

32 to the electrostatic self-interaction of a classical charge-density distribution, $\rho(\mathbf{r})$.

33 For reasons that will become apparent shortly, this classical contribution is usually

34 referred to as the *Hartree* energy:

$$E_H[\rho] = \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{64}$$

19

Following KS, the following ansatz can be made for the HK functional:

$$F[\rho] = T_0[\rho] + E_H[\rho] + E_{xc}[\rho], \tag{65}$$

where the exchange-correlation (XC) energy, E_{xc}, is the sum of the differences between the system's kinetic and *e-e* interaction energies and T_0 and E_H, respectively. Strictly speaking, Eq. (65) is nothing but a *definition* of the XC energy which was in fact named *stupidity energy* by R.P. Feynman [13], because it is the difference between what we can compute and what we can't! The actual use of the very concept of XC energy will depend on our ability to find approximations to it which are accurate enough and easy to compute in practice. Before dwelling further upon the definition, computation, and use of the XC functional, let us make the final step of the Kohn-Sham's path which leads to a practical scheme for using DFT in actual calculations. By using the ansatz of Eq. (65), the differential form of the HK variational principle, Eq. (63), would read:

$$\frac{\delta T_0}{\delta \rho(\mathbf{r})} + V(\mathbf{r}) + V_H(\mathbf{r}) + \mu_{xc}(\mathbf{r}) = \mu, \tag{66}$$

where the Hartree potential, $V_H(\mathbf{r})$, is the functional derivative of the Hartree energy functional, Eq. (64):

$$V_H(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \tag{67}$$

and $\mu_{xc}(\mathbf{r})$ is the (unknown) functional derivative of the (unknown) XC functional:

$$\mu_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})}. \tag{68}$$

Eq. (66) is formally equivalent to the HK variational principle for a system of *non interacting* electrons, subject to an effective potential:

$$V_{KS}(\mathbf{r}) = V(\mathbf{r}) + V_H(\mathbf{r}) + \mu_{xc}(\mathbf{r}). \tag{69}$$

The cleverness of this observation stems from the fact that we do know how to calculate the ground-state density—as well as any other property which can be expressed as the expectation value of a local one-electron operator, $\hat{O} = \sum_i O(\mathbf{r}_i)$— of non-interacting electron systems:

$$\langle \Psi | \hat{O} | \Psi \rangle = \int O(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}, \tag{70}$$

$$\rho(\mathbf{r}) = \sum_n |\phi_n(\mathbf{r})|^2, \tag{71}$$

1 where the sum extends over the first N low-lying eigenfunctions of the one-body
2 Schrdinger equation:

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial \mathbf{r}^2} + V_{KS}(\mathbf{r})\right)\phi_n(\mathbf{r}) = \epsilon_n\phi_n(\mathbf{r}). \tag{72}$$

3 Once this equation has been solved, the ground-state energy can in principle be
4 calculated using (62), (65), and (71), and the obvious prescription for the non-
5 interacting kinetic energy:

$$T_0[\rho] = -\frac{1}{2}\sum_n \int \phi_n^\star(\mathbf{r})\frac{\partial^2\phi_n(\mathbf{r})}{\partial \mathbf{r}^2}d\mathbf{r}. \tag{73}$$

6 The result reads:

$$E^\circ[V] = -\frac{1}{2}\sum_n \int \phi_n^\star(\mathbf{r})\frac{\partial^2\phi_n(\mathbf{r})}{\partial \mathbf{r}^2}d\mathbf{r} + \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_H[\rho] + E_{xc}[\rho]. \tag{74}$$

Eq. (74) can be rewritten in a different form which is often used in applications.
Let us consider the ground-state energy of the auxiliary Kohn-Sham system of
non-interacting electrons, which equals the sum of the N lowest-lying eigenvalues
of Eq. (72):

$$\sum_n \epsilon_n = -\frac{1}{2}\sum_n \int \phi_n^\star(\mathbf{r})\frac{\partial^2\phi_n(\mathbf{r})}{\partial \mathbf{r}^2}d\mathbf{r} + \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}+$$
$$\int V_H(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + \int \mu_{xc}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}. \tag{75}$$

7 By comparing Eq. (75) with Eq. (74), we arrive at the expression:

$$E^\circ[V] = \sum_n \epsilon_n - \int \mu_{xc}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} - \frac{1}{2}E_H[\rho] + E_{xc}[\rho]. \tag{76}$$

8 All in all, the Kohn-Sham ansatz allows one to replace the very complicated
9 many-body Schrödinger equation, Eq. (38), for the ground state of a system of
10 interacting electrons with the much handier Schrödinger equation for a system of
11 non-interacting electrons, Eq. (72), where, however, the effective potential, V_{KS},
12 depends on its own solutions through Eqs. (69) and (71). For this reason, this
13 Schrödinger equation is said to be *self-consistent* or *non-linear*. Eq. (72) has a
14 very similar structure to the Hartree equation [14], where the Hartree potential
15 generated by the classical charge-density distribution, $\rho(\mathbf{r})$, is supplemented with
16 an appropriate XC term, $\mu_{xc}(\mathbf{r})$, Eq. (68). Of course, our ability to even state Eq.
17 (72) depends on our knowledge (or lack thereof) of the XC energy, Eq. (65), and
18 potential, Eq. (68).

3.4 The local-density approximation

In a homogeneous system ($V(\mathbf{r}) = $ cnst: the *homogeneous electron gas*), the density is constant, $\rho(\mathbf{r}) = \frac{N}{\Omega} \equiv n$, and any functional of $\rho(\mathbf{r})$ reduces to a function of n: $E[\rho] = N\epsilon(n)$, where Ω and N are the total volume and number of electrons, respectively, and $\epsilon(n)$ the energy per particle in the homogeneous electron gas as a function of the density. A common approximation to the density dependence of energy functionals assumes that each small volume of the system, ω_i, contributes to the energy the same amount that a same volume would in a homogeneous electron gas with a density equal to the local density of the inhomogeneous system (*local density approximation*, LDA):

$$
\begin{aligned}
E[\rho] &\approx \sum_i \rho_i \omega_i \epsilon(\rho_i) \\
&\approx \int \epsilon(\rho(\mathbf{r})) \rho(\mathbf{r}) d\mathbf{r}.
\end{aligned} \tag{77}
$$

When applied to the independent-electron kinetic energy functional, $T_0[\rho]$, Eq. (77) gives rise to the Thomas-Fermi approximation [15, 16]:

$$
T_{TF}[\rho] = \frac{3\hbar^2}{10m}(3\pi)^{\frac{2}{3}} \int \rho(\mathbf{r})^{\frac{5}{3}} d\mathbf{r}. \tag{78}
$$

Eq. (78) results in a rather poor approximation for the kinetic energy, which is however capable to provide qualitatively meaningful results for atoms. The Hartree energy, Eq. (64), cannot be approximated by any sensible local-density form because of the long-range Coulomb interaction between distant charge fluctuations. It turns out, however, that when applied to the exchange-correlation energy functional, the LDA gives surprisingly good results. In the LDA, the XC energy functionals reads:

$$
E_{xc}^{LDA}[\rho] = \int \varepsilon_{xc}(\rho(\mathbf{r})) \rho(\mathbf{r}) d\mathbf{r}, \tag{79}
$$

where $\varepsilon_{xc}(n)$ is the difference between the energy per particle of a homogeneous electron gas at density n, which can be obtained from very accurate Quantum Monte Carlo calculations [17], and the corresponding independent-electron result, $\varepsilon_0(n) = \frac{3\hbar^2}{10m}(3\pi)^{\frac{2}{3}} n^{\frac{2}{3}}$, *cf.* Eq. (78). The functional derivative of Eq. (79) gives the LDA XC potential:

$$
\mu_{xc}^{LDA} = \varepsilon_{xc}(\rho(\mathbf{r})) + \rho(\mathbf{r})\varepsilon'_{xc}(\rho(\mathbf{r})), \tag{80}
$$

where $\varepsilon'_{xc}(n)$ is the derivative of $\varepsilon_{xc}(n)$.

¹ The LDA was demonstrated to be exact in the two limits of slowly varying
² densities,

$$\frac{|\nabla\rho(\mathbf{r})|}{\rho(\mathbf{r})} \ll k_F(\mathbf{r}) = \left(3\pi^2\rho(\mathbf{r})\right)^{\frac{1}{3}}, \tag{81}$$

³ where $k_F(\mathbf{r})$ is the *local Fermi momentum*, and in that of high density ($\rho(\mathbf{r})a_0^3 \gg$
⁴ 1, a_0 being the Bohr radius) [12].

⁵ 3.5 Pair correlations, correlation holes, and energy functionals

⁶ Real materials hardly satisfy the conditions originally set for the validity of the
⁷ LDA. The success met by over 40 years of DFT practice demonstrates though that
⁸ the predictive power of LDA is much superior than could be expected from the
⁹ original statements by Kohn and Sham. Some of this success can be explained on
¹⁰ the basis of the properties of the *exchange-correlation hole* [18] that we are now
¹¹ going to introduce and describe.

¹²   We have seen in Sec. 3.2, Eq. (57), that the HK density functional is the sum
¹³ of the ground-state expectation values of the kinetic-energy and *e-e* interaction
¹⁴ operators, which are one-body and two-body operators, respectively. Let us see
¹⁵ how one- and two-body expectation values can be expressed in general in terms
¹⁶ of reduced density matrices.

¹⁷   Let $\hat{O} = \sum_i \hat{o}(i)$ be a (generally non-local) one-body operator. Its ground-
¹⁸ state expectation value reads:

$$\langle\Psi|\hat{O}|\Psi\rangle = \sum_i \int \Psi^\star(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_i, \cdots \mathbf{r}_N) o(\mathbf{r}_i, \mathbf{r}_i') \times$$
$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_i', \cdots \mathbf{r}_N) \, d\mathbf{r}_1 d\mathbf{r}_2 \cdots d\mathbf{r}_i d\mathbf{r}_i' \cdots d\mathbf{r}_N \tag{82}$$
$$= \int o(\mathbf{r}, \mathbf{r}')\gamma(\mathbf{r}, \mathbf{r}')d\mathbf{r}d\mathbf{r}', \tag{83}$$

¹⁹ where $o(\mathbf{r}, rr') = \langle\mathbf{r}|\hat{o}|\mathbf{r}'\rangle$ is the kernel of the $\hat{o}$ operator, and

$$\gamma(\mathbf{r}, \mathbf{r}') = N \int \Psi^\star(\mathbf{r}, \mathbf{r}_2, \cdots \mathbf{r}_N)\Psi(\mathbf{r}', \mathbf{r}_2, \cdots \mathbf{r}_N) \, d\mathbf{r}_2 d\mathbf{r}_3 \cdots d\mathbf{r}_N \tag{84}$$

²⁰ is the one-electron density matrix. Expectation values of local one-electron oper-
²¹ ators can be expressed in terms of the diagonal of the one-electron density matrix:
²²

$$\rho(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r}), \tag{85}$$

²³ which is of course the electron charge-density distribution, Eqs. (39) and (40).
²⁴ The evaluation of expectation values of differential operators, such as *e.g.*, the

₁ kinetic-energy operator $\hat{K}$ requires the knowledge of off-diagonal elements of the
₂ density matrix:

$$\langle \hat{K} \rangle = -\frac{\hbar^2}{2m} \int \left[\frac{\partial^2}{\partial \mathbf{r}'^2} \gamma(\mathbf{r}, \mathbf{r}') \right]_{\mathbf{r}=\mathbf{r}'} d\mathbf{r}. \tag{86}$$

₃ Let us now consider the case of a local[1] two-electron operator,
₄ $\hat{W} = \frac{1}{2} \sum_{i \neq j} w(\mathbf{r}_i, \mathbf{r}_j)$. Without loss of generality we can assume that $\hat{W}$ is the
₅ *e-e* interaction operator, so that $w(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}$. Its ground-state expectation
₆ value is:

$$\langle \Psi | \hat{W} | \Psi \rangle = \frac{1}{2} \int \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \, d\mathbf{r}_1 d\mathbf{r}_2, \tag{87}$$

₇ where:

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) P^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \tag{88}$$
$$\equiv \Gamma(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2), \tag{89}$$

$P^{(2)}$ being the pair probability density and Γ the two-electron density matrix:

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}'_1, \mathbf{r}'_2) = N(N-1) \int \Psi^\star(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \cdots \mathbf{r}_N) \times$$
$$\Psi(\mathbf{r}'_1, \mathbf{r}'_2, \mathbf{r}_3, \cdots \mathbf{r}_N) \, d\mathbf{r}_3 \cdots d\mathbf{r}_N. \tag{90}$$

₈ The pair probability density can be written as:

$$P^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = P^{(1)}(\mathbf{r}_1) \pi(\mathbf{r}_2 | \mathbf{r}_1), \tag{91}$$

₉ where $\pi(\mathbf{r}_2, \mathbf{r}_1)$ is the *conditional* probability density for an electron sitting at
₁₀ $\mathbf{r}_2$, given that another electron sits at $\mathbf{r}_1$. If the electrons were uncorrelated, the
₁₁ conditional probability would be equal to the one-electron reduced probability:
₁₂ $\pi(\mathbf{r}_2 | \mathbf{r}_1) = P^{(1)}(\mathbf{r}_2) \equiv \rho(\mathbf{r}_2)/N$. In the presence of electronic correlations, the
₁₃ pair density, Eq. (88), reads:

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_1)(N-1)\pi(\mathbf{r}_2 | \mathbf{r}_1) \tag{92}$$
$$= \rho(\mathbf{r}_1) \big[\rho(\mathbf{r}_2) - (N-1)\pi(\mathbf{r}_2 | \mathbf{r}_1) - \rho(\mathbf{r}_2) \big] \tag{93}$$
$$= \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) + \rho(\mathbf{r}_1) h_{xc}(\mathbf{r}_1, \mathbf{r}_2), \tag{94}$$

₁₄ where the *exchange-correlation hole*, h_{xc}, is defined as:

$$h_{xc}(\mathbf{r}_1, \mathbf{r}_2) = (N-1)\pi(\mathbf{r}_2 | \mathbf{r}_1) - \rho(\mathbf{r}_2). \tag{95}$$

---

[1]We restrict ourselves to local two-electron operators because the *e-e* interaction, which is the only relevant such operator occurring in nature, is local. In the case of one-electron operators, instead there are a few important ones, such as the kinetic energy, momentum, current density, etc. which are non-local

1 The XC hole as a function of $\mathbf{r}_2$ is normalized to -1 because it is the difference
2 between two functions normalized to $N - 1$ and N.

3 Eq. (94) allows one to write the ground-state expectation value of the *e-e*
4 interaction in terms of the ground-state electron density distribution and XC hole:
5

$$\langle \hat{W} \rangle = E_H[\rho] + \int \rho(\mathbf{r}_1)\varepsilon(\mathbf{r}_1)d\mathbf{r}_1, \tag{96}$$

6 where:

$$\varepsilon(\mathbf{r}_1) = \frac{1}{2} \int \frac{h_{xc}(\mathbf{r}_2|\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_2. \tag{97}$$

7 One would be tempted to identify the expression of Eq. (97) with the XC contribu-
8 tion to the HK functional, Eq. (65). Actually, this cannot be done because the XC
9 energy, as defined in Eq. (65), also contains a kinetic contribution, $\langle \hat{K} \rangle - T_0[\rho]$. In
10 order to express the sum of the kinetic plus potential energy contributions to the
11 XC functional in terms of the latter alone, it is convenient to introduce and use the
12 concept of *adiabatic connection* [18].

13 Let $\hat{H}_\lambda$ be the Hamiltonian:

$$\hat{H}_\lambda = \hat{K} + \lambda\hat{W} + \hat{V}_\lambda, \tag{98}$$

14 where $0 \leq \lambda \leq 1$ is a parameter, and the external potential $V_\lambda(\mathbf{r})$ is determined
15 by the condition that the ground-state charge density determined by $\hat{H}_\lambda$ is inde-
16 pendent of λ, $\rho_\lambda(\mathbf{r}) = \rho(\mathbf{r})$. This condition shows that for $\lambda = 1$ V_λ equals the
17 physical external potential acting on the systems, whereas for $\lambda = 0$ it equals the
18 Kohn-Sham potential:

$$V_{\lambda=1}(\mathbf{r}) = V(\mathbf{r}) \tag{99}$$
$$V_{\lambda=0}(\mathbf{r}) = V_{KS}(\mathbf{r}). \tag{100}$$

19 The Hellmann-Feynman theorem gives the derivative of the ground-state energy
20 with respect to λ:

$$\frac{dE(\lambda)}{d\lambda} = \langle \hat{W} \rangle_\lambda + \int \rho(\mathbf{r})V'_\lambda(\mathbf{r})d\mathbf{r}, \tag{101}$$

21 where $V'_\lambda(\mathbf{r}) = \partial V_\lambda(\mathbf{r})/\partial\lambda$. Note that the first term on the rhs of Eq. (101) is
22 the expectation of the λ-independent operator $\hat{W}$ over a two-electron distribution
23 which depends on λ, whereas in the second term the one-electron distribution ρ
24 does not depend on λ by hypothesis, while $V'_\lambda(\mathbf{r})$ does. Also note that, although
25 we do not know what V'_λ is, its integral from $\lambda = 0$ to 1 is equal to the difference
26 between the physical external potential and it KS counterpart, see Eqs. (99) and

(100):

$$
\begin{aligned}
\int V'_\lambda(\mathbf{r})d\lambda &= V(\mathbf{r}) - V_{KS}(\mathbf{r}) \\
&= -V_H(\mathbf{r}) - \mu_{xc}(\mathbf{r}).
\end{aligned}
\tag{102}
$$

The physical ($\lambda = 1$) ground-state energy can be calculated from the fictitious KS ($\lambda = 0$) value by integrating its derivative, Eq. (101), with respect to the strength of the coupling constant, λ:

$$
\begin{aligned}
E[V] &= E_{\lambda=0} + \int_0^1 \frac{dE(\lambda)}{d\lambda}d\lambda \\
&= \sum_n \epsilon_n + \int_0^1 \langle \hat{W} \rangle_\lambda d\lambda + \int \rho(\mathbf{r})\big(V(\mathbf{r}) - V_{KS}(\mathbf{r})\big)d\mathbf{r} \quad (103) \\
&= \sum_n \epsilon_n + \int_0^1 \langle \hat{W} \rangle_\lambda - 2E_H[\rho] - \int \mu_{xc}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}. \quad (104)
\end{aligned}
$$

By comparing Eq. (104) with Eq. (76) and using Eqs. (96) and (97) for the expectation value of $\hat{W}$, we find:

$$
E_{xc}[\rho] = \frac{1}{2} \int \rho(\mathbf{r}_1) \frac{\bar{h}_{xc}(\mathbf{r}_2|\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2,
\tag{105}
$$

where $\bar{h}_{xc}$ is an effective XC hole defined as the average over the coupling constant, λ, of the XC corresponding to the $\hat{H}_\lambda$ Hamiltonian:

$$
\bar{h}_{xc}(\mathbf{r}_2|\mathbf{r}_1) = \int_0^1 h^\lambda_{xc}(\mathbf{r}_2|\mathbf{r}_1)d\lambda.
\tag{106}
$$

Eq. (105) shows that the XC energy does not depend on many details of the XC hole. In particular, the XC energy density $\varepsilon(\mathbf{r}_1)$ of Eq. (97) can be seen as the classical electrostatic potential generated at $\mathbf{r}_1$ by the charge-density distribution $\frac{1}{2}h_{xc}(\mathbf{r}_2|\mathbf{r}_1)$. This quantity depends only the spherical component of this charge distribution, as a function of $\mathbf{r}_2 - \mathbf{r}_1$, because multipolar potentials generated by a regular charge-density ditribution vanish at the origin. As a consequence, *any* XC hole possessing the correct spherical component would give the exact XC energy, no matter how wrong the non-spherical components are. This argument was used by Gunnarsson and Lundqvist [18] to justify the unexpected good quality of the results provided by the LDA. The latter can in fact be reformulated by the assumption that the XC hole in real materials is well approximated by the XC hole of the homogeneous electron gas. In fact, as poor as this approximation may be in general, it fulfills the sum rul that the XC hole is correctly normalized to -1,

26

and furthermore the quality of the XC energy given by this approximation does not depend on its inability to predict the non-spherical components of the XC hole (the XC of the homegeneous electron gas is in fact spherically symmetric). Last but not least, the XC hole was shown to be rather short-range. By dimensional arguments, its range can only be of the order of $\rho(\mathbf{r})^{-\frac{1}{3}}$, which is also what the LDA would predict. All of these considerations taken together are believed to justify, if not to demonstrate, the unexpected predictive power that the LDA has displayed all over 40 years of intense benchmarking [19].

This section has to be reviewed

3.6 The generalized gradient approximation(s)

Ever since the founding papers of Hohenberg, Kohn, and Sham [10, 12], it has been known that, as the LDA becomes exact in the limit of slowly varying densities, Eq. (81), in this regime corrections to it can be expressed in terms of an expansion in powers of the density gradients. This observation lead to an attempt to build energy functionals on the basis of such an expansion [20, 21]. It was soon to be realized that such an approach is bound to fail as the results provided by such *improved* functionals are in fact often worse than the original LDA results.

The situation changed dramatically starting from the mid eighties with the proposal to modify the form of gradient dependence of the density functional from a power expansion in the density gradient (which is exact in the slowly varying regime, but inaccurate in practice), to a general functional of the form:

$$E_{xc}^{GGA}[\rho] = \int \epsilon_{xc}^{GGA}\big(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})\big)\rho(\mathbf{r})d\mathbf{r}. \tag{107}$$

This section has to be completed

3.7 Orbital energies and excited states

In a system of non-interacting electrons, the eigenvalues of the one-electron Schrödinger equation describe the energy necessary to remove one electron from the system (occupied orbitals) or to add one to it (empty orbitals). Orbital energies are thus differences between the energy levels of systems with $N \pm 1$ electrons and the N-electron ground state, and they are often referred to as *quasi-particle* (QP) energies. This property also holds in the Hartree-Fock approximation [22]. As the mathematical structure of the Kohn-Sham theory is that of a *mean-field approximation* (even if it is in principle an exact theory), such as Hartree-Fock, it is tempting to identify KS orbital energies with electron addition or removal energies. Although it can be demonstrated that in any finite system the highest occupied KS eigenvalue equals the first ionization potential [23], this identification

27

¹ is not possible in general, and KS energy eigenvalues do not have in general any
² physical meaning. In spite of this, and in the lack of equally simple theoretical
³ tools, KS energy levels are often assumed to be fair approximations for the QP
⁴ energies. One of the most striking demonstrations of the limits of this approach is
⁵ the well know inability of DFT to predict optical gaps in insulators [24].

In a paper of 1965 [25], Sham and Kohn provided a theoretical justification
for the practice of identifying KS eigenvalues with QP energies. QP energies are
poles of the one-particle propagator (Green's function) which satisfies the Dyson's
equation:

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial \mathbf{r}^2} + V(\mathbf{r}) + V_H(\mathbf{r}) - \epsilon\right)G(\mathbf{r}, \mathbf{r}'; \epsilon) +$$
$$\int \Sigma(\mathbf{r}, \mathbf{r}''; \epsilon)G(\mathbf{r}'', \mathbf{r}'; \epsilon)d\mathbf{r}'' = -\delta(\mathbf{r} - \mathbf{r}'), \quad (108)$$

⁶ where Σ is the so-called self-energy operator. If the self-energy were Hermitean
⁷ and energy-independent, Eq. (108) would be equivalent to the Schrödinger-like
⁸ equation:

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial \mathbf{r}^2} + V(\mathbf{r}) + V_H(\mathbf{r})\right)\phi_n(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}')\phi_n(\mathbf{r}')d\mathbf{r}' = \epsilon_n\phi_n(\mathbf{r}), \quad (109)$$

⁹ whose eigenvalues, ϵ_n, would then be the poles of the propagator, *i.e.* QP energies.
¹⁰ In the many-body literature, the eigenfunctions of Eq. (109) (or rather of its full-
¹¹ fledged form where the self-energy is energy-dependent) are called *quasi-particle*
¹² *amplitudes*, or *Feynman-Dyson amplitudes*. Note that the Hartree-Fock equation
¹³ has exactly the form of Eq. (109), with the self-energy approximated by the ex-
¹⁴ change operator: $\Sigma(\mathbf{r}, \mathbf{r}') = \gamma(\mathbf{r}, \mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$, γ being the one-electron density
¹⁵ matrix of Eq. (84).

¹⁶ Sham and Kohn showed that under the assumption that charge inhomogeneities
¹⁷ are both small ($|\Omega\rho(\mathbf{r})/N - 1| << 1$) and slowly varying ($|\nabla\rho(\mathbf{r})|/\rho(\mathbf{r})^{\frac{4}{3}} << 1$),
¹⁸ the self-energy operator can be approximated by a local potential which has the
¹⁹ same form as the XC potential of the ground-state KS theory! Of course, charge
²⁰ inhomogeneities in real materials are not small, nor are they slowly varying, and
²¹ the justification (or lack thereof) for using KS eigenvalues as estimates of QP
²² energies can only be rooted into extended practice and in the evaluation of its
²³ successes and failures. The success of such a practice is rather controversial, but
²⁴ there seems to be a fair consensus that in weakly correlated materials the occu-
²⁵ pied KS eigenvalues and eigenfunctions are reasonable approximations to the QP
²⁶ energies and amplitudes, whereas for empty states substantial improvements with
²⁷ respect to the plain KS model for one-electron excitations are needed. Much work

is presently being devoted to the development of reliable, accurate, and still numerically workable schemes to the quantum many-body problem. Schemes based on the partial summation of entire classes of terms in the perturbative expansion of the one- and two-electron propagators, such as the *GW* and the *Bethe-Salpether* approaches, seem rather accurate in the weak-correlation regime, albeit the extent to which they can be used for large-scale materials simulations is still to be demonstrated.

In strongly correlated materials, such as transition-metal oxides, cuprates, or manganites to name a few, the ground- and low-lying excited states cannot be expressed as perturbations with respect to the eigenstates of any reference system of non interacting electrons. This fact has two important and somewhat related, although conceptually quite distinct, consequences. On the one hand, DFT, although able in principle to cope with these systems, as well as with any other, will be severely limited by the fact that the splitting of of the non-interacting kinetic energy out of the HK density functional would hardly alleviate our ignorance about it. As a consequence, the XC energy would be dominated in this case by our own substantial stupidity [13], more than it is tamed by the ingenuity of the Kohn-Sham construction [12]. On the other hand, the lack of any reference independent-electron systems, make electron propagators loose their *quasiparticle* singly peaked features, so that perturbative approaches, such as *GW* and its generalizations, loose much of their predictive power. The quest for methods which are robust in the strongly correlated regime, and yet simple enough to be implemented in practice, is a very active field of research, as the very existence of this summer school indicates. A good introduction to the state of the art in this field is a review paper recently written by Kotliar and coworkers [26].

This section has to be extensively reviewed

References

[1] E.H. Lieb, Int. J. Quant. Chem. **24**, 243 (1983).

[2] R.F. Nalewajski and R.G. Parr, J. Chem. Phys. **77** 399 (1982).

[3] R. Fukuda, T. Kotani, Y. Suzuki, and S. Yokojima, Progr. Theor. Phys. **92**, 833 (1994).

[4] N. Argaman and G. Makov, Am. J. Phys. **68**, 69 (2000).

[5] N. Argaman and G. Makov, Phys. Rev. B **66**, 52413 (2002).

[6] P.W. Ayers and W. Yang, J. Chem. Phys. **124**, 224108 (2006).

[7] H. Hellmann, *Einführung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), Sec. 54.

[8] R.P. Feynmann, Phys. Rev. **56**, 340 (1939).

[9] W. Kohn, Phys. Rev. Lett. **51**, 1596 (1983).

[10] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[11] M. Levy, Proc. Natl. Acad. Sci. USA **76**, 6062 (1979).

[12] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).

[13] R.P. Feynman, Statistical Mechanics (Benjamin, Reading, 1972, p. 249).

[14] D. R. Hartree, Proc. R. Soc. London **A113**, 621, (1928).

[15] L.H. Thomas, Proc. Cambridge Philos. Soc., **23**, 542 (1927).

[16] E. Fermi, Rend. Accad. Naz. Lincei **6**, 602 (1927); Z. Phys. **48**, 73 (1928).

[17] D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[18] O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[19] A review of DFT as applied to atoms and molecules, with emphasis on the LDA can be found in R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[20] F. Herman, J.P. Van Dyke, and I.P. Ortenburger, Phys. Rev. Lett. **22**, 807 (1969).

[21] P.S. Svendsen and U. von Barth, Phys. Rev. B **54**, 17402 (1996).

[22] T. Koopmans, Physica **1**, 104 (1933); see also, *e.g.*, F.L. Pilar, *Elementary quantum chemstry* (McGraw-Hill, New York, 1968) p.350.

[23] M. Levy, J.P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).

[24] R.W. Godby, M. Schlter and L.J. Sham, Phys. Rev. Lett. **56** 2415 (1986).

[25] L.J. Sham and W. Kohn, Phys. Rev. **145**, 561 (1965).

[26] G. Kotliar, S.Y. Savrasov, K. Haule, V.S. Oudovenko, O. Parcollet, and C.A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).