A bird's-eye view of density-functional theory

Klaus Capelle

Departamento de Química e Física Molecular
Instituto de Química de São Carlos
Universidade de São Paulo
Caixa Postal 780, São Carlos, 13560-970 SP, Brazil

(final version 20 February 2003)

Contents

1 Preface

This paper is the outgrowth of lectures the author gave at the Chemistry Institute of the University of São Paulo at São Carlos, Brazil, and at the VIII'th Summer School on Electronic Structure of the Brazilian Physical Society. It is an attempt to introduce density-functional theory (DFT) in a language accessible for students entering the field or researchers from other fields. It is not meant to be a scholarly review of DFT, but rather an informal guide to its conceptual basis and some recent developments and advances. The Hohenberg-Kohn theorem and the Kohn-Sham equations are discussed in some detail, including comparisons with the equations of Thomas-Fermi, Hartree-Fock, and Dyson. Approximate density functionals, selected aspects of applications of DFT, and a variety of extensions of standard DFT are also discussed, albeit in less detail. Throughout it is attempted to provide a balanced treatment of aspects that are relevant for chemistry and aspects relevant for physics, but with a strong bias towards conceptual foundations. The paper is intended to be read before (or in parallel with) one of the many excellent more technical reviews available in the literature. The author apologizes to all researchers whose work has not received proper consideration. The limits of the author's knowledge, as well as the limits of the available space and the nature of the intended audience, have from the outset prohibited any attempt at comprehensiveness.

2 What is density-functional theory?

Density-functional theory is one of the most popular and successful quantum mechanical approaches to matter. It is nowadays routinely applied for calculating, e.g., the binding energy of molecules in chemistry and the band structure of solids in physics. First applications relevant for fields traditionally considered more distant from quantum mechanics, such as biology and mineralogy are beginning to appear. Superconductivity, atoms in the focus of strong laser pulses, relativistic effects in heavy elements and in atomic nuclei, classical liquids, and magnetic properties of alloys have all been studied with DFT.

DFT owes this versatility to the generality of its fundamental concepts and the flexibility one has in implementing them. In spite of this flexibility and generality, DFT is based on quite a rigid conceptual framework. This and the next section introduce some aspects of this framework in general terms. The following two sections, 4 and 5, then deal in detail with two core elements of DFT, the Hohenberg-Kohn theorem and the Kohn-Sham equations. The final two sections, 6 and 7, contain a (necessarily less detailed) description of approximations typically made in practical DFT calculations, and of some extensions and generalizations of DFT.

To get a first idea of what density-functional theory is about, it is useful to take a step back and recall some elementary quantum mechanics. In quantum mechanics we learn that all information we can possibly have about a given system is contained in the system's wave function, Ψ. Here we will exclusively be concerned with the electronic structure of atoms, molecules, and solids. The nuclear degrees of freedom (e.g., the crystal

lattice in a solid) appear only in the form of a potential $v(\mathbf{r})$ acting on the electrons, so that the wave function depends only on the electronic coordinates.[1] Nonrelativistically, this wave function is calculated from Schrödinger's equation, which for a single electron moving in a potential $v(\mathbf{r})$ reads

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + v(\mathbf{r}) \right] \Psi(\mathbf{r}) = \epsilon \Psi(\mathbf{r}). \tag{1}$$

If there is more than one electron (i.e., one has a many-body problem) Schrödinger's equation becomes

$$\left[\sum_i^N \left(-\frac{\hbar^2 \nabla_i^2}{2m} + v(\mathbf{r}_i) \right) + \sum_{i<j} U(\mathbf{r}_i, \mathbf{r}_j) \right] \Psi(\mathbf{r}_1, \mathbf{r}_2 \dots, \mathbf{r}_N) = E \Psi(\mathbf{r}_1, \mathbf{r}_2 \dots, \mathbf{r}_N), \tag{2}$$

where N is the number of electrons and $U(\mathbf{r}_i, \mathbf{r}_j)$ is the electron-electron interaction. For a Coulomb system (the only type of system we consider here) one has

$$\hat{U} = \sum_{i<j} U(\mathbf{r}_i, \mathbf{r}_j) = \sum_{i<j} \frac{q^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{3}$$

Note that this is the same operator for any system of particles interacting via the Coulomb interaction, just as the kinetic energy operator

$$\hat{T} = -\frac{\hbar^2}{2m} \sum_i \nabla_i^2 \tag{4}$$

is the same for any nonrelativistic system. Whether our system is an atom, a molecule, or a solid thus depends only on the potential $v(\mathbf{r}_i)$. For an atom, e.g.,

$$\hat{V} = \sum_i v(\mathbf{r}_i) = -\sum_i \frac{Ze^2}{|\mathbf{r}_i - \mathbf{R}|}, \tag{5}$$

where Ze is the nuclear charge ($e > 0$) and $\mathbf{R}$ the nuclear position. For a molecule or a solid one has

$$\hat{V} = \sum_i v(\mathbf{r}_i) = -\sum_{ik} \frac{Z_k e^2}{|\mathbf{r}_i - \mathbf{R}_k|}, \tag{6}$$

where the sum on k extends over all nuclei in the system, each with charge $Z_k e$ and position $\mathbf{R}_k$. It is only the spatial arrangement of the $\mathbf{R}_k$ (together with the corresponding boundary conditions) that distinguishes, fundamentally, a molecule from a solid.[2] Similarly, it is only through the term $\hat{U}$ that the (essentially simple) single-body quantum mechanics of Eq. (1) differs from the extremely complex many-body problem posed by Eq. (2). These properties are built into DFT in a very fundamental way.

[1] This is the so-called Born-Oppenheimer approximation. We follow the usual custom of calling $v(\mathbf{r})$ a 'potential' although it is, strictly speaking, a potential energy.

[2] One sometimes says that $\hat{T}$ and $\hat{U}$ are 'universal', while $\hat{V}$ is system-dependent, or 'nonuniversal'. We will come back to this terminology.

The usual quantum-mechanical approach to Schrödinger's equation (SE) can be summarized by the following sequence

$$v(\mathbf{r}) \overset{SE}{\Longrightarrow} \Psi(\mathbf{r}_1, \mathbf{r}_2 \dots, \mathbf{r}_N) \overset{\langle\Psi|\dots|\Psi\rangle}{\Longrightarrow} \text{observables}, \tag{7}$$

i.e., one specifies the system by choosing $v(\mathbf{r})$, plugs it into Schrödinger's equation, solves that equation for the wave function Ψ, and then calculates expectation values of observables with this wave function. One among the observables that are calculated in this way is the particle density

$$n(\mathbf{r}) = N \int d^3 r_2 \int d^3 r_3 \dots \int d^3 r_N \Psi^*(\mathbf{r}, \mathbf{r}_2 \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2 \dots, \mathbf{r}_N). \tag{8}$$

Many powerful methods for solving Schrödinger's equation have been developed during decades of struggling with the many-body problem. In physics, for example, one has diagramatic perturbation theory (based on Feynman diagrams and Green's functions), while in chemistry one often uses configuration interaction (CI) methods, which are based on systematic expansion in Slater determinants. A host of more special techniques also exists. The problem with these methods is the great demand they place on one's computational resources: it is simply impossible to apply them efficiently to large and complex systems. Nobody has ever calculated the chemical properties of a 100-atom molecule with full CI, or the electronic structure of a real semiconductor using nothing but Green's functions.

It is here where DFT provides a viable alternative, less accurate perhaps, but much more versatile. DFT explicitly recognizes that nonrelativistic Coulomb systems differ only by their potential $v(\mathbf{r})$, and supplies a prescription for dealing with the universal operators $\hat{T}$ and $\hat{U}$ once and for all.[3] Furthermore, DFT provides a way to systematically map the many-body problem, with $\hat{U}$, onto a single-body problem, without $\hat{U}$. All this is done by promoting the particle density $n(\mathbf{r})$ from just one among many observables to the status of key variable, on which the calculation of all other observables can be based. This approach forms the basis of the large majority of electronic-structure calculations in physics and chemistry. Much of what we know about the electrical, magnetic, and structural properties of materials has been calculated using DFT, and the extent to which DFT has contributed to the science of molecules is reflected by the 1998 Nobel Prize in Chemistry, which was awarded to Walter Kohn [1], the founding father of DFT, and John Pople [2], who was instrumental in implementing DFT in computational chemistry.

The density-functional approach can be summarized by the sequence

$$n(\mathbf{r}) \Longrightarrow \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \Longrightarrow v(\mathbf{r}), \tag{9}$$

i.e., knowledge of $n(\mathbf{r})$ implies knowledge of the wave function and the potential, and hence of all other observables. Although this sequence describes the conceptual structure of DFT, it does not really represent what is done in actual applications of it, which typically proceed along rather different lines. In the following sections I try to explain both the conceptual structure and some of the many possible shapes and disguises under which this structure appears in applications.

---

[3] We will see that in practice this prescription can be implemented only approximately. Still, these approximations retain a high degree of universality in the sense that they often work well for more than one type of system.

The literature on DFT is large, and rich in excellent reviews and overviews. Some representative examples of full reviews and systematic collections of research papers are Refs. [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. The present overview of DFT is much less detailed and advanced than these treatments. Introductions to DFT that are more similar in spirit to the present one (but differ in emphasis and selection of topics) are the contribution of Levy in Ref. [7], the one of Kurth and Perdew in Ref. [12], and Ref. [16] by Makov and Argaman. My aim in the present text is to give a bird's-eye view of DFT in a language that should be accessible to an advanced undergraduate student who has completed a first course in quantum mechanics, in either chemistry or physics. Many interesting details, proofs of theorems, illustrative applications, and exciting developments had to be left out, just as any discussion of issues that are specific to only certain subfields of either physics or chemistry. All of this, and much more, can be found in the references cited above, to which the present little text may perhaps serve as a prelude.

3 From wave functions to density functionals (via Green's functions and density matrices)

This section can be skipped on a first reading. It is not necessary for understanding the following sections (except perhaps for its last paragraph, on functionals), but presents a different point of view on what one does when one practices density-functional theory.

It is a fundamental postulate of quantum mechanics that the wave function contains all possible information about the system. Normally, this is much more information that one can handle: for a system with $N = 100$ particles the many-body wave function is an extremely complicated function of 300 spatial and 100 spin[4] variables that would be impossible to manipulate algebraically or to extract any information from, even if it were possible to calculate it in the first place. For this reason one searches for less complicated objects to formulate one's theory. Such objects should contain the experimentally relevant information, such as energies, densities, etc, but do not need to contain explicit information about the coordinates of every single particle. One such object is called the *Green's function*. In mathematics one usually defines the Green's function of a linear operator $\mathcal{L}$ via $[z - \mathcal{L}(z, r)]G(r, r; z) = \delta(r - r')$, where $\delta(r, r')$ is Dirac's delta function. For a single free quantum particle one has, for example,

$$\left[E + \frac{\hbar^2 \nabla^2}{2m} \right] G^{(0)}(\mathbf{r}, \mathbf{r}'; E) = \delta(\mathbf{r} - \mathbf{r}'). \tag{10}$$

Many applications of such single-particle Green's functions are discussed in Ref. [17]. In many-body physics one finds it necessary to introduce more complicated Green's functions. In an interacting system the single-particle Green's function is modified by the presence of the interaction between the particles. In time-independent situations it now satisfies

---

[4]To keep our notation simple we have suppressed spin indices in the preceeding section. We will continue to do so here, and only put them back when discussing spin-density-functional theory, in Sec. 7.

the equation

$$\left[E + \frac{\hbar^2 \nabla^2}{2m} \right] G(\mathbf{r}, \mathbf{r}'; E) = \delta(\mathbf{r} - \mathbf{r}') + R[U(\mathbf{r}, \mathbf{r}')], \tag{11}$$

where $R[U]$ is a complicated term that depends explicitly on the interaction $\hat{U}$ [17, 18]. From this Green's function one can calculate the expectation value of any single-body operator (such as the potential, the kinetic energy, the particle density, etc), but also that of certain two-particle operators, such as the Hamiltonian in the presence of particle-particle interactions.[5]

One way to obtain this Green's function is via solution of what is called Dyson's equation,

$$G(\mathbf{r}, \mathbf{r}'; t, t') = G^{(0)}(\mathbf{r}, \mathbf{r}'; t, t')$$
$$+ \int d^3x \int d^3x' \int d^3\tau \int d^3\tau' \, G^{(0)}(\mathbf{r}, \mathbf{x}; t, \tau) \Sigma(\mathbf{x}, \tau, \mathbf{x}', \tau') G(\mathbf{x}', \mathbf{r}'; \tau', t'), \tag{12}$$

where Σ is known as the irreducible self energy [17, 18] and $G^{(0)}$ is the Green's function in the absence of any interaction.[6] This equation (which we will not attempt to solve here) has a characteristic property that we will meet again when we study the (much simpler) Kohn-Sham and Hartree-Fock equations, in Sec. 5: the operator on the right-hand side, which determines G on the left-hand side, depends on G itself. The mathematical problem posed by this equation is thus *nonlinear*. We will return to such nonlinearity when we discuss self-consistent solution of the Kohn-Sham equation. The quantity Σ will appear again in Sec. 5.3 when we discuss the meaning of the eigenvalues of the Kohn-Sham equation.

A proper discussion of Σ, G, and R requires a formalism known as second quantization [17, 18], and is beyond the scope of the present overview. One important limiting case, however, can be discussed easily. This is the single-particle density matrix, which can be defined by

$$\gamma(\mathbf{r}, \mathbf{r}') = -i\hbar \lim_{t' \to t} G(\mathbf{r}, \mathbf{r}', t - t'). \tag{13}$$

An equivalent, and more immediately useful, definition is

$$\gamma(\mathbf{r}, \mathbf{r}') = N \int d^3r_2 d^3r_3 \dots \Psi^*(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, \dots \mathbf{r}_N) \Psi(\mathbf{r}', \mathbf{r}_2, \mathbf{r}_3, \dots \mathbf{r}_N). \tag{14}$$

The structure of this expression is quite simple: all coordinates that γ does not depend upon are set equal in Ψ and Ψ^*, and integrated over. The usefulness of this definition becomes apparent when we consider how one would calculate the expectation values of a multiplicative single-particle operator $\hat{A} = \sum_i^N a(\mathbf{r}_i)$ (such as the potential $\hat{V} = \sum_i^N v(\mathbf{r}_i)$),

$$\langle A \rangle = \int d^3r_1 \dots \int d^3r_N \, \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N) \left[\sum_i^N a(\mathbf{r}_i) \right] \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N) \tag{15}$$

$$= N \int d^3r_1 \dots \int d^3r_N \, \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N) a(\mathbf{r}_1) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots \mathbf{r}_N) \tag{16}$$

$$= \int d^3r \, a(\mathbf{r}) \gamma(\mathbf{r}, \mathbf{r}). \tag{17}$$

---

[5]Note that expressions like 'two-particle operator' refer to the number of particles involved in the definition of the operator (two in the case of an interaction, one for a potential energy, etc), not to the total number of particles present in the system.

[6]By Fourier transformation with respect to $t - t'$ one passes from $G(\mathbf{r}, \mathbf{r}'; t, t')$ to $G(\mathbf{r}, \mathbf{r}'; E)$.

The last equation implies that if one knows $\gamma(\mathbf{r}, \mathbf{r})$ one can calculate the expectation value of any multiplicative single-particle operator in terms of it, regardless of the number of particles present in the system. The simplification is enormous and density matrices are very popular in, e.g., computational chemistry for precisely this reason. More details are given in, e.g., Ref. [4]. Here we just point out that for nonmultiplicative operators (such as the kinetic energy, which contains a derivative) one requires the full single-particle matrix $\gamma(\mathbf{r}, \mathbf{r}')$ and not only $\gamma(\mathbf{r}, \mathbf{r})$.

On the other hand, one does not know how to calculate expectation values of two-particle operators, such as the interaction itself, or the full Hamiltonian (i.e., the total energy), in terms of the density matrix $\gamma(\mathbf{r}, \mathbf{r}')$. For this purpose one requires a more complicated function, called the two-particle density matrix. This situation is to be contrasted with that of the single-particle Green's function, for which one knows how to calculate the expectation values of $\hat{U}$ and $\hat{H}$. Obviously some information has gotten lost in passing from G to γ. This can also be seen very clearly from Eq. (13) or (14), which show that information that is contained in G and Ψ, respectively, is erased ('integrated out') in the definition of $\gamma(\mathbf{r}, \mathbf{r}')$.

Apparently even less information is contained in the particle density $n(\mathbf{r})$, which is obtained from the diagonal element of $\gamma(\mathbf{r}, \mathbf{r}')$

$$n(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r}')|_{\mathbf{r}' \to \mathbf{r}}. \tag{18}$$

(This follows either from substituting the density operator $\hat{n}(\mathbf{r}) = \sum_i^N \delta(\mathbf{r} - \mathbf{r}_i)$ into Eq. (17) or from comparing (8) with (14).) The particle density is an even simpler function than $\gamma(\mathbf{r}, \mathbf{r}')$: it depends on one coordinate only, it can easily be visualized as a three-dimensional charge distribution, and it is directly accessible in experiments. These advantages, however, seem to be more than compensated by the fact that one has integrated out an enormous amount of specific information about the system in going from wave functions to Green's functions, and on to density matrices and finally the density itself.

The great surprise of density-functional theory is that in fact no information has been lost at all, at least as long as one considers the system only in its ground state: one can prove that the ground-state density $n_0(\mathbf{r})$ completely determines the ground-state wave function $\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots \mathbf{r}_N)$. Hence, in the ground state, a function of one variable is equivalent to a function of N variables! This property, which constitutes the celebrated Hohenberg-Kohn theorem and is discussed in more detail in the next section, shows that we have only integrated out *explicit* information on our way from wave functions via Green's functions and density matrices to densities. *Implicitly* all the information that was contained in the ground-state wave function is still contained in the ground-state density. Part of the art of practical DFT is how to get this implicit information out, once one has obtained the density!

Before we discuss these issues more carefully, let us introduce a useful mathematical tool. Since the wave function is determined by the density, we can write it as $\Psi_0 = \Psi[n_0](\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$, which indicates that Ψ_0 is a function of its N spatial variables, but a *functional* of $n_0(\mathbf{r})$. More generally, a functional $F[n]$ can be defined (in an admittedly mathematically sloppy way) as a rule for going from a function to a number, just as a function $y = f(x)$ is a rule (f) for going from a number (x) to a number (y). A simple

example of a functional is the particle number,

$$N = \int d^3r \, n(\mathbf{r}) = N[n], \tag{19}$$

which is a rule for obtaining the number N, given the function $n(\mathbf{r})$. Note that the name given to the argument of n is completely irrelevant, since the functional depends on the *function* itself, not on its variable. Hence we do not need to distinuish $F[n(\mathbf{r})]$ from, e.g., $F[n(\mathbf{r}')]$. Another important case is that in which the functional depends on a parameter, such as in

$$v_H[n](\mathbf{r}) = q^2 \int d^3r' \, \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{20}$$

which is a rule that for any value of the parameter $\mathbf{r}$ associates a value $v_H[n](\mathbf{r})$ with the function $n(\mathbf{r}')$. This term is the so-called Hartree potential, which we will repeatedly encounter below.

4 DFT as a many-body theory: The Hohenberg-Kohn theorem

At the heart of DFT is the Hohenberg-Kohn (HK) theorem. This theorem states that for ground states Eq. (8) can be inverted: given a *ground-state* density $n_0(\mathbf{r})$ it is possible, in principle, to calculate the corresponding *ground-state* wave function $\Psi_0(\mathbf{r}_1, \mathbf{r}_2 \ldots, \mathbf{r}_N)$. This means that Ψ_0 is a functional of n_0. (Consequently, all ground-state observables are functionals of n_0, too.) If Ψ_0 can be calculated from n_0 and vice versa, both functions are equivalent and contain exactly the same information. At first sight this seems impossible: how can a function of one (vectorial) variable $\mathbf{r}$ be equivalent to a function of N (vectorial) variables $\mathbf{r}_1 \ldots \mathbf{r}_N$? How can one arbitrary variable contain the same information as N arbitrary variables?

The crucial fact which makes this possible is that knowledge of $n_0(\mathbf{r})$ implies implicit knowledge of much more than that of an arbitrary function $n(\mathbf{r})$. Knowledge that a given function is a *ground-state* density implies knowledge of an extremely detailed subsidiary condition: this function represents the spatial distribution of probability for the *lowest energy solution* to an N-particle second-order differential equation: Schrödinger's equation. It is not yet obvious, of course, that this particular auxiliar condition is enough to determine $\Psi_0(\mathbf{r}, \mathbf{r}_2 \ldots, \mathbf{r}_N)$ uniquely from $n_0(\mathbf{r})$, but it is clear where the additional information that makes the construction possible in principle comes from: it is hidden in the subscript 0, implying that this n_0 is not any arbitrary function but a very special one.

To prove mathematically that this is indeed true is not very complicated. In fact, it is a bit surprising that it took 38 years from Schrödinger's first paper on quantum mechanics [19] to Hohenberg and Kohn's paper containing their famous theorem [20]. The mathematical details of the proof can be found in the original paper [20] or in standard textbooks [3, 4] and need not be reproduced here. However, it may be useful to point out that there is more than one proof available in the literature. First, there is the original HK proof by contradiction. Second, there is the very elegant and powerful proof by

constrained search, given independently by M. Levy [21] and E. Lieb [22]. These proofs can be found in any book on the subject, and we refrain from repetition of the treatments given there.

Furthermore, there are two arguments, perhaps less well known than the two classical proofs just mentioned, which lend additional credibility to the theorem. One of these is based on a clever perturbation theoretic argument, due to M. Gillan. It can be found in section 10.10 of Ref. [23]. The other is valid only for Coulomb potentials. It is based on Kato's theorem, which states [24, 25] that for such potentials the electron density has a cusp at the position of the nuclei, where it satisfies

$$Z_k = -\frac{a_0}{2n(\mathbf{r})} \frac{dn(\mathbf{r})}{dr}\bigg|_{\mathbf{r} \to \mathbf{R}_k}. \tag{21}$$

Here $\mathbf{R}_k$ denotes the positions of the nuclei, Z_k their atomic number, and a_0 is the Bohr radius. For a Coulomb system one can thus, in principle, read off all information necessary for completely specifying the Hamiltonian directly from examining the density distribution: the integral over $n(\mathbf{r})$ yields N, the total particle number; the position of the cusps of $n(\mathbf{r})$ are the positions of the nuclei, $\mathbf{R}_k$; and the derivative of $n(\mathbf{r})$ at these positions yields Z_k by means of Eq. (21). This is all one needs to specify the complete Hamiltonian of Eq. (2) (and thus implicitly all its eigenstates). In practice one almost never knows the density distribution sufficiently well to implement the search for the cusps and calculate the local derivatives. Still, Kato's theorem provides a vivid illustration of how the density can indeed contain sufficient information to completely specify a nontrivial Hamiltonian.[7]

For future reference we now provide a commented summary of the content of the HK theorem. This summary consists of four statements:

(1) The ground-state (GS) wave function is a unique functional of the GS density: $\Psi_0 = \Psi[n_0]$. This is the essence of the HK theorem. As a consequence, the GS expectation value of any observable $\hat{O}$ is a functional of $n_0(\mathbf{r})$, too:

$$O_0 = O[n_0] = \langle \Psi[n_0] | \hat{O} | \Psi[n_0] \rangle. \tag{22}$$

(2) Perhaps the most important observable is the GS energy. This energy

$$E_{v,0} = E_v[n_0] = \langle \Psi[n_0] | \hat{H} | \Psi[n_0] \rangle, \tag{23}$$

where $\hat{H} = \hat{T} + \hat{U} + \hat{V}$, has a useful variational property:

$$E_v[n_0] \leq E_v[n']. \tag{24}$$

where n_0 is GS density in potential $\hat{V}$ and n' is some other density. This is very similar to the usual variational principle for wave functions. If I calculate the expectation value of a Hamiltonian with a trial wave function Ψ' that is not its GS wave function Ψ_0 I can never obtain an energy below the true GS energy,

$$E_{v,0} = E_v[\Psi_0] = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \leq \langle \Psi' | \hat{H} | \Psi' \rangle = E_v[\Psi']. \tag{25}$$

---

[7]Note that, unlike the full Hohenberg-Kohn theorem, Kato's theorem does apply only to superpositions of Coulomb potentials, and can therefore not be applied directly to the effective Kohn-Sham potential.

Similarly, in DFT, if I calculate the GS energy of a Hamiltonian using a density that is not its GS density I can never find a result below the true GS energy. This is what Eq. (24) says, and it is so important for practical applications of DFT that it is sometimes called the *second Hohenberg-Kohn theorem* (Eq. (22) is the first one, then).

(3) Recalling that the kinetic and interaction energies of a nonrelativistic Coulomb system are described by universal operators, we can also write E_v as

$$E_v[n] = T[n] + U[n] + V[n], \tag{26}$$

where $T[n]$ and $U[n]$ are *universal functionals* [defined as expectation values of the type (22) of $\hat{T}$ and $\hat{U}$], independent of $v(\mathbf{r})$. On the other hand,

$$V[n] = \int d^3r \, n(\mathbf{r})v(\mathbf{r}) \tag{27}$$

is obviously nonuniversal (it depends on $v(\mathbf{r})$, i.e., on the system under study), but extremely simple: once the system is specified, i.e., $v(\mathbf{r})$ is known, the functional $V[n]$ is known explicitly.

(4) There is a fourth substatement to the HK theorem. This statement is that the GS density does determine not only the GS wave function Ψ_0, but also the potential $v(\mathbf{r})$:

$$v(\mathbf{r}) = v[n_0](\mathbf{r}). \tag{28}$$

As a consequence the explicit reference to the potential in the previous equations is not necessary, and one can rewrite, e.g. Eq. (23) as

$$E_0 = E[n_0] = \langle \Psi[n_0] | \hat{T} + \hat{U} + \hat{V}[n_0] | \Psi[n_0] \rangle. \tag{29}$$

Another consequence is that n_0 now does determine not only the GS wave function but the complete Hamiltonian ($\hat{T}$ and $\hat{U}$ are fixed), and thus all excited states, too:

$$\Psi_k(\mathbf{r}_1, \mathbf{r}_2 \ldots, \mathbf{r}_N) = \Psi_k[n_0], \tag{30}$$

where k labels the entire spectrum of the many-body Hamiltonian $\hat{H}$.

Originally this fourth statement was considered to be as sound as the other three. However, it has become clear very recently, as a consequence of work of the author with G. Vignale [26, 27] and, independently, of H. Eschrig and W. Pickett [28], that there are significant exceptions to it. In fact, the fourth substatement holds only when one formulates DFT exclusively in terms of the charge density, as we have done up to this point. It does not hold when one works with spin densities (spin-DFT) or current densities (current-DFT).[8] In these (and some other) cases the densities still determine the wave function, but they do not uniquely determine the corresponding potentials. This so-called *nonuniqueness problem* has been discovered only recently, and its consequences are now beginning to be explored [26, 27, 28, 29]. It is clear, however, that the fourth substatement is, from a practical point of view, the least important of the four, and most applications of DFT do not have to be reconsidered as a consequence of its eventual failure. (But some do: see Refs. [26, 27] for examples.)

---

[8]In Section 7 we will briefly discuss these formulations of DFT.

Another conceptual problem with the HK theorem, much better known and more studied than nonuniqueness, is representability. To understand what representability is about, consider the following two questions: (i) How do I know, given an arbitrary function $n(\mathbf{r})$, that this function can be represented in the form (8), i.e., that it is a density arising from an antisymmetric N-body wave function $\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N)$? (ii) How do I know, given a function that can be written in the form (8), that this density is a *ground-state* density of a local potential $v(\mathbf{r})$? The first of these questions is known as the N-representability problem, the second is called v-representability. Note that these are quite important questions: if one should find for example, in a numerical calculation, a minimum of $E_v[n]$ that is not N-representable, then this mimimum is not the physically acceptable solution to the problem at hand. Luckily, the N-representability problem of the single-particle density has been solved: any square-integrable nonnegative function can be written in the form (8). The v-representability problem does not have a similarly simple solution, but one can use the constrained search algorithm of Levy and Lieb to show that it is irrelevant for the proof of the HK theorem [3, 4].

After these abstract considerations let us now consider one way in which one can make practical use of DFT. Assume we have specified our system (i.e., $v(\mathbf{r})$ is known). Assume further that we have reliable approximations for $U[n]$ and $T[n]$. All one has to do then is to minimize

$$E_v[n] = T[n] + U[n] + \int d^3r\, n(\mathbf{r})v(\mathbf{r}) \tag{31}$$

with respect to $n(\mathbf{r})$. The minimizing function $n_0(\mathbf{r})$ is the system's GS charge density and the value $E_{v,0} = E_v[n_0]$ is the GS energy. Assume now that $v(\mathbf{r})$ depends on a parameter a. This can be, for example, the lattice constant in a solid or the angle between two atoms in a molecule. Calculation of $E_{v,0}$ for many values of a allows one to plot the curve $E_{v,0}(a)$ and to find the value of a that mimimizes it. This value, a_0, is the GS lattice constant or angle. In this way one can calculate quantities like molecular geometries and sizes, lattice constants, unit cell volumes, charge distributions, total energies, etc. By looking at the change of $E_{v,0}(a)$ with a one can, moreover, calculate compressibilities and bulk moduli (in solids) and vibrational frequencies (in molecules). By comparing the total energy of a composite system (e.g., a molecule) with that of its constituent systems (e.g., individual atoms) one obtains dissociation energies. By calculating the total energy for systems with one electron more or less one obtains electron affinities and ionization energies. All this follows from DFT without having to solve the many-body Schrödinger equation and without having to make a single-body approximation.

In theory it should be possible to calculate *all* observables, since the HK theorem guarantees that they are all functionals of $n_0(\mathbf{r})$. In practice, one does not know how to do this explicitly. Another problem is that the minimization of $E_v[n]$ is, in general, a tough numerical problem on its own. And, moreover, one needs reliable approximations for $T[n]$ and $U[n]$ to begin with. In the next section, on the Kohn-Sham equations, we will see one widely used method for solving these problems. Before looking at that, however, it is worthwhile to recall an older, but still occasionally useful, alternative: the Thomas-Fermi approximation.

In this approximation one sets

$$U[n] \approx U_H[n] = \frac{q^2}{2} \int d^3r \int d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{32}$$

i.e., approximates the full interaction energy by the Hartree energy, the classical electrostatic interaction energy. One further sets, initially,

$$T[n] \approx T^{LDA}[n] = \int d^3r\, t^{hom}(n(\mathbf{r})), \tag{33}$$

where $t^{hom}(n)$ is the kinetic energy density of a homogeneous system with (constant) density n. This formula is a first example of a local-density approximation (LDA). In this type of approximation one imagines the real inhomogeneous system (with density $n(\mathbf{r})$ in potential $v(\mathbf{r})$) to be decomposed in small cells in each of which $n(\mathbf{r})$ and $v(\mathbf{r})$ are approximately constant. In each cell (i.e., locally) one can then use the expression of a homogeneous system to approximate the contribution of the cell to the real inhomogeneous one. Making the cells infinitesimally small and summing over all of them yields Eq. (33). For a noninteracting system (specified by subscript s, for 'single-particle') the function $t_s^{hom}(n)$ is known explicitly, $t_s^{hom}(n) = 3\hbar^2(3\pi^2)^{2/3}n^{5/3}/(10m)$. This is exploited to further approximate

$$T[n] \approx T^{LDA}[n] \approx T_s^{LDA}[n] = \int d^3r\, t_s^{hom}(n(\mathbf{r})). \tag{34}$$

The Thomas-Fermi approximation then consists in setting

$$E[n] = T[n] + U[n] + V[n] \approx E^{TF}[n] = T_s^{LDA}[n] + U_H[n] + V[n]. \tag{35}$$

This expression is the starting point for a large body of literature in chemistry and physics [10, 25]. More recent approximations for $T[n]$ can be found, e.g., in Refs. [30, 31].

5 DFT as an effective single-body theory: The Kohn-Sham equations

Density-functional theory can be implemented in many ways. The minimization of an explicit energy functional, discussed up to this point, is not normally the most efficient among them. Much more widely used is the Kohn-Sham approach. Interestingly, this approach owes its success and popularity partly to the fact that it does not exclusively work in terms of the particle (or charge) density, but brings a special kind of wave functions (single-particle orbitals) back into the game. We will now see in some detail how this is done.

5.1 Exchange-correlation energy

The Thomas-Fermi approximation (34) for $T[n]$ is not very good. A more accurate scheme for treating the kinetic energy functional is based on decomposing $T[n]$ into one part

that represents the kinetic energy of noninteracting particles, i.e., $T_s[n]$, and one that represents the remainder, denoted $T_c[n]$ (the subscripts s and c stand for 'single-particle' and 'correlation', respectively[9]),

$$T[n] = T_s[n] + T_c[n]. \tag{36}$$

$T_s[n]$ is not known exactly as a functional of n [and using the LDA to approximate it leads one back to the Thomas-Fermi approximation (34)], but it is easily expressed in terms of the single-particle orbitals $\phi_i(\mathbf{r})$ of a noninteracting system with density n, as

$$T_s[n] = -\frac{\hbar^2}{2m} \sum_i^N \int d^3r \, \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}), \tag{37}$$

because for noninteracting particles the total kinetic energy is just the sum of the individual kinetic energies. Since all $\phi_i(\mathbf{r})$ are functionals of n, this expression for T_s is an explicit orbital functional but an implicit density functional, $T_s[n] = T_s[\{\phi_i[n]\}]$, where the notation indicates that T_s depends on the full set of occupied orbitals ϕ_i, each of which is a functional of n. Other such orbital functionals will be discussed in Sec. 6.

We now rewrite the exact energy functional as

$$E[n] = T[n] + U[n] + V[n] = T_s[\{\phi_i[n]\}] + U_H[n] + E_{xc}[n] + V[n], \tag{38}$$

where by definition E_{xc} contains the differences $T - T_s$ (i.e. T_c) and $U - U_H$. Eq. (38) is formally exact, but of course E_{xc} is unknown — although the HK theorem guarantees that it is a density functional. This functional, $E_{xc}[n]$, is called the *exchange-correlation* (xc) energy. It is often decomposed as $E_{xc} = E_x + E_c$, where E_x is due to the Pauli principle (exchange energy) and E_c is due to correlations. (T_c is then a part of E_c.) The exchange energy can be written explicitly in terms of the single-particle orbitals as[10]

$$E_x[\{\phi_i[n]\}] = -\frac{q^2}{2} \sum_{jk} \int d^3r \int d^3r' \, \frac{\phi_j^*(\mathbf{r})\phi_k^*(\mathbf{r}')\phi_j(\mathbf{r}')\phi_k(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}, \tag{39}$$

but no exact expression in terms of the density is known.

A simple way to understand the origin of correlation is to consider the operator for the Coulomb interaction

$$\hat{U} = \frac{q^2}{2} \int d^3r \int d^3r' \, \frac{\hat{n}(\mathbf{r})\hat{n}(\mathbf{r}') - \hat{n}(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{40}$$

where $\hat{n}$ is the density operator and the term with the delta function subtracts out the interaction of a charge with itself. The expectation value of this operator, $U = \langle \Psi | \hat{U} | \Psi \rangle$, involves the expectation value of a product of density operators. In the Hartree term (32), on the other hand, this expectation value of a product is replaced by a product of expectation values of the form $n = \langle \Psi | \hat{n} | \Psi \rangle$. This replacement neglects quantum

---

[9]One can prove that there is no exchange contribution to T, i.e., $T_x \equiv 0$.

[10]This differs from the exchange energy used in Hartree-Fock theory only in the substitution of Hartree-Fock orbitals $\phi_i^{HF}(\mathbf{r})$ by Kohn-Sham orbitals $\phi_i(\mathbf{r})$. Some consequences of this substitution are explored in Ref. [32].

fluctuations[11] about the expectation values: by writing $\hat{n} = n + \delta\hat{n}_{fluc}$ and substituting in the previous equation we see that the difference between the true operator and the Hartree energy is entirely due to the fluctuations $\delta\hat{n}_{fluc}$ and the self-interaction correction. The exchange-correlation functional introduced in Eq. (38) thus corrects the Hartree term by including the self-interaction correction and the above quantum fluctuations, and the noninteracting kinetic energy by including T_c. The fluctuations are not entirely free, since the Pauli principle prohibits that two particles with the same spin come arbitrarily close. The influence of this constraint on the total energy is described by the exchange energy E_x, while the remainder is called the correlation energy E_c.

Clearly E_c is an enormously complex object, and DFT would be of little use if one had to know it exactly for making calculations. The practical advantage of writing $E[n]$ in the form Eq. (38) is that the unknown functional $E_{xc}[n]$ is typically much smaller than the known terms T_s, U_H and V. One can thus hope that reasonably simple approximations for $E_{xc}[n]$ provide useful results for $E[n]$. Some successful approximations are discussed in Sec. 6. Two properties of the exchange, $E_x[n]$, and correlation, $E_c[n]$, contributions to this functional are the scaling conditions first obtained by Levy and Perdew [34]

$$
\begin{align}
E_x[n_\lambda] &= \lambda E_x[n] \tag{41}\\
E_c[n_\lambda] &> \lambda E_c[n] \qquad \text{for} \lambda > 1 \tag{42}\\
E_c[n_\lambda] &< \lambda E_c[n] \qquad \text{for} \lambda < 1, \tag{43}
\end{align}
$$

where $n_\lambda(\mathbf{r}) = \lambda^3 n(\lambda\mathbf{r})$ is a scaled density. These properties serve as constraints in the construction of approximations for the functionals $E_x[n]$ and $E_c[n]$. Many other similar properties are known. A useful overview is the contribution of M. Levy in Ref. [15].

5.2 Kohn-Sham equations

Since T_s is now written as an orbital functional one cannot directly minimize Eq. (38) with respect to n. Instead one commonly employs a scheme suggested by Kohn and Sham [35], for performing the minimization indirectly. This scheme starts by writing the minimization as

$$0 = \frac{\delta E}{\delta n} = \frac{\delta T_s}{\delta n} + \frac{\delta V}{\delta n} + \frac{\delta U_H}{\delta n} + \frac{\delta E_{xc}}{\delta n} = \frac{\delta T_s}{\delta n} + v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}). \tag{44}$$

As a consequence of Eq. (27), $\delta V/\delta n = v(\mathbf{r})$, the 'external' potential the electrons move in. The term $\delta U_H/\delta n$ simply yields the Hartree potential, introduced in Eq. (20). For the term $\delta E_{xc}/\delta n$, which can only be calculated explicitly once an approximation for E_{xc} has been chosen, one commonly writes v_{xc}.

Consider now a noninteracting system of particles moving in external potential $v_s(\mathbf{r})$. For this system the minimization condition is simply

$$0 = \frac{\delta E}{\delta n_s} = \frac{\delta T_s}{\delta n_s} + v_s(\mathbf{r}), \tag{45}$$

---

[11] At finite temperature there are also thermal fluctuations. To properly include these one must use a finite-temperature formulation of DFT [33]. See also the contribution of B. L. Gyorffy et al. in Ref. [15] for DFT treatment of fluctuations.

since there are no Hartree and xc terms in the absence of interactions. Comparing this with the previous equation we find that both minimizations have the same solution $n_s(\mathbf{r}) \equiv n(\mathbf{r})$ if v_s is chosen to be

$$v_s(\mathbf{r}) = v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}). \tag{46}$$

Consequently, one can calculate the density of the interacting (many-body) system in potential $v(\mathbf{r})$ by solving the equations of a noninteracting (single-body) system in potential $v_s(\mathbf{r})$.[12] In particular, the Schrödinger equation of this auxiliary system,

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + v_s(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \tag{47}$$

yields orbitals that reproduce the density $n(\mathbf{r})$ of the original system (these are the same orbitals employed in Eq. (37)),

$$n(\mathbf{r}) \equiv n_s(\mathbf{r}) = \sum_i^N |\phi_i(\mathbf{r})|^2. \tag{48}$$

Eqs. (46) to (48) are the celebrated Kohn-Sham (KS) equations. They replace the problem of minimizing $E[n]$ by that of solving a noninteracting Schrödinger equation. (Recall that the minimization of $E[n]$ originally replaced the problem of solving the many-body Schrödinger equation!)

Since both v_H and v_{xc} depend on n, which depends on the ϕ_i, which in turn depend on v_s, the problem of solving the KS equations is a nonlinear one, just as is the one of solving the (much more complicated) Dyson equation (12). The usual way of solving such problems is to start with an initial guess for $n(\mathbf{r})$, calculate the corresponding $v_s(\mathbf{r})$, and then solve the differential equation (47) for the ϕ_i. From these one calculates a new density, using (48), and starts again. The process is repeated until it converges. (Only rarely this requires more than a few dozen iterations.) The technical name for this procedure is 'self-consistency cycle'.

Once one has a converged solution n_0, one can calculate the total energy from Eq. (38) or, equivalently and more conveniently, from

$$E_0 = \sum_i^N \epsilon_i - \frac{q^2}{2} \int d^3r \int d^3r' \frac{n_0(\mathbf{r})n_0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - \int d^3r \, v_{xc}(\mathbf{r})n_0(\mathbf{r}) + E_{xc}[n_0]. \tag{49}$$

This equation shows that E_0 is not simply the sum of all ϵ_i. In fact, it should be clear from our derivation of Eq. (47) that the ϵ_i are completely artificial objects: they are the eigenvalues of an auxiliary single-body Schrödinger equation whose eigenfunctions (orbitals) yield the correct density. It is only this density that has strict physical meaning in the KS equations. The KS eigenvalues, on the other hand, bear only a semiquantitative resemblance with the true energy spectrum [36], but are not to be trusted quantitatively.

The only exception to this rule is the highest occupied KS eigenvalue. Denoting by $\epsilon_N(M)$ the N'th eigenvalue of a system with M electrons, one can show rigorously

---

[12]The question whether such a potential $v_s(\mathbf{r})$ always exists in the mathematical sense, is called the noninteracting v-representability problem. This problem is unsolved, but only very rarely relevant for practical applications.

that $\epsilon_N(N) = -I$, the negative of the first ionization energy of the N-body system, and $\epsilon_{N+1}(N+1) = -A$, the negative of the electron affinity of the same N-body system. These relations hold for the exact functional only. When calculated with an approximate functional of the LDA or GGA type, the highest eigenvalues usually do not provide good approximations to the experimental I and A. Better results for these observables are obtained by calculating them as total-energy differences, according to $I = E_0(N-1) - E_0(N)$ and $A = E_0(N) - E_0(N+1)$, where $E_0(N)$ is the ground-state energy of the N-body system. Alternatively, self-interaction corrections can be used to obtain improved ionization energies and electron affinities from Kohn-Sham eigenvalues [37].

Given the auxiliary nature of the other Kohn-Sham eigenvalues, it comes as a great (and welcome) surprise that in many situations (typically characterized by the presence of fermionic quasiparticles and absence of strong correlations) the Kohn-Sham eigenvalues ϵ_i do, empirically, provide a quite good approximation to the actual energy levels of extended systems [38]. Moreover, it was found recently [39] that the energy gap between the highest occupied KS eigenvalue and the lowest unoccupied one agrees quite well with experimental values for the true energy gap, even for difficult cases such as transition-metal oxides, provided the KS eigenvalues are calculated with the B3LYP hybrid functional (see Sec. 6.2 for more on this functional.) Similarly, it has recently been argued that the differences between exact KS eigenvalues (as opposed to approximate ones, obtained, e.g., in the LDA) provide surprisingly good approximations for the excitation energies of finite systems [40]. Much research remains to be done before it is clear to what extent such conclusions can be generalized, and how situations in which the KS eigenvalues approximate the true excitation spectrum are to be characterized microscopically.[13]

Most band-structure calculations in solid-state physics are actually calculations of the KS eigenvalues ϵ_i. This simplification has proved enormously successful, but when one uses it one must be aware of the fact that one is taking the auxiliary single-body equation (47) literally as an approximation to the many-body Schrödinger equation. DFT, practiced in this mode, is not a rigorous many-body theory anymore, but a mean-field theory (albeit one with a very sophisticated mean field $v_s(\mathbf{r})$).

5.3   Hartree, Hartree-Fock, and Dyson equations

A partial justification for the interpretation of the KS eigenvalues as an approximation to quasi-particle energies can be given by comparing the KS equation with other self-consistent equations of many-body physics. Among the simplest such equations are the Hartree equation

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + v(\mathbf{r}) + v_H(\mathbf{r}) \right] \phi_i^H(\mathbf{r}) = \epsilon_i^H \phi_i^H(\mathbf{r}), \qquad (50)$$

and the Hartree-Fock (HF) equation

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + v(\mathbf{r}) + v_H(\mathbf{r}) \right] \phi_i^{HF}(\mathbf{r}) - q^2 \int d^3r' \frac{\gamma(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \phi_i^{HF}(\mathbf{r}') = \epsilon_i^{HF} \phi_i^{HF}(\mathbf{r}), \qquad (51)$$

---

[13]Several more rigorous approaches to excited states in DFT, which do not require the KS eigenvalues to have physical meaning, are mentioned in Sec. 7.

where $\gamma(\mathbf{r}, \mathbf{r}')$ is the density matrix of Eq. (14). It is a fact known as Koopman's theorem [41] that the HF eigenvalues ϵ_i^{HF} can be interpreted as unrelaxed electron-removal energies (i.e., ionization energies of the i'th electron, neglecting reorganization of the remaining electrons after removal). As mentioned above, in DFT only the highest occupied eigenvalue corresponds to an ionization energy, but unlike in HF this energy includes relaxation effects.

The KS equation (47) includes both exchange and correlation via the multiplicative operator v_{xc}. Both exchange and correlation are normally approximated in DFT,[14] whereas HF accounts for exchange exactly, through the integral operator containing $\gamma(\mathbf{r}, \mathbf{r}')$, but neglects correlation completely. In practise DFT results are typically at least as good as HF ones and often comparable to much more sophisticated correlated methods — and the KS equations are much easier to solve than the HF equations.[15]

All three single-particle equations, Hartree, Hartree-Fock, and Kohn-Sham can also be interpreted as approximations to Dyson's equation (12), which can be rewritten [18] as

$$\left(-\frac{\hbar^2 \nabla^2}{2m} + v(\mathbf{r})\right)\psi_k(\mathbf{r}) + \int d^3r' \, \Sigma(\mathbf{r}, \mathbf{r}', E_k)\psi_k(\mathbf{r}') = E_k \psi_k(\mathbf{r}), \qquad (52)$$

where Σ is the irreducible self energy introduced in Eq. (12). The E_k appearing in this equation are the true (quasi-)electron addition and removal energies of the many-body system. Needless to say, it is much more complicated to solve this equation than the HF or KS equations. It is also much harder to find useful approximations for Σ than for v_{xc}.[16] Obviously, the KS equation employs a local, energy-independent potential v_s in place of the nonlocal, energy-dependent operator Σ. Whenever this is a good approximation, the ϵ_i are also a good approximation to the E_k.

The interpretation of the KS equation (47) as an approximation to Eq. (52) is suggestive and useful, but certainly not necessary for DFT to work: if the KS equations are only used to obtain the density, and all other observables, such as total energies, are calculated from this density, then the KS equations in themselves are not an approximation at all, but simply a very useful mathematical tool.

5.4 Basis functions

In practice, numerical solution of the KS differential equation (47) typically proceeds by expanding the KS orbitals in a suitable set of basis functions and solving the resulting secular equation for the coefficients in this expansion and/or for the eigenvalues for which it has a solution.

In physics much is known about the construction of basis functions for solids due to decades of experience with band-structure calculations. This includes many calculations that predate the widespread use of DFT in physics. There is a fundamental dichotomy between methods that work with fixed basis functions that do not depend on energy, and methods that employ energy-dependent basis functions. Fixed basis functions are used

---

[14]A possibility to treat exchange exactly in DFT is offered by the OEP method discussed in Sec. 6.3.

[15]This is due to the integral operator in the HF equations.

[16]The so-called GW approximation [42, 43, 44] is one such approximation for Σ, but in actual implementations of it one usually takes DFT-KS results as an input.

e.g., in plane-wave expansions, tight-binding or LCAO (linear combination of atomic orbitals) approximations, or the OPW (orthogonalized plane wave) method. Examples for methods using energy-dependent functions are the APW (augmented plane wave) or KKR (Korringa-Kohn-Rostoker) approaches. This distinction became less clear-cut with the introduction of 'linear methods' [45], in which energy-dependent basis functions are linearized (Taylor expanded) around some fixed reference energy. The most widely used methods for solving the Kohn-Sham equation in solid-state physics, LMTO (linear muffin tin orbitals) and LAPW (linear augmented plane waves), are of this latter type.

The situation is quite similar in chemistry. Due to decades of experience with Hartree-Fock and CI calculations much is known about the construction of basis functions that are suitable for molecules. Almost all of this continues to hold in DFT — a fact that has greatly contributed to the recent popularity of DFT in chemistry. Chemical basis functions are classified with respect to their behaviour as a function of the radial coordinate into Slater type orbitals (STOs), which decay exponentially far from the origin, and Gaussian type orbitals (GTOs), which have a gaussian behaviour. STOs more closely resemble the true behaviour of atomic wave functions [in particular the cusp condition of Eq. (21)], but GTOs are easier to handle numerically. The so-called 'contracted basis functions', in which STO basis functions are reexpanded in a small number of GTOs, represent a compromise between the accuracy of STOs and the convenience of GTOs. The most common methods for solving the Kohn-Sham equations in quantum chemistry are of this type [2, 41]. More details on the development of suitable basis functions can be found, e.g., in Ref. [41].

6   Making DFT practical: Approximations

There are basically three distinct types of approximations involved in a DFT calculation. One is conceptual, and concerns the interpretation of KS eigenvalues and orbitals as physical energies and wave functions. This approximation is optional — if one does not want to make it one simply does not attach meaning to the eigenvalues of Eq. (47). The pros and cons of this procedure were discussed in Secs. 5.2 and 5.3. The second type of approximation is numerical, and concerns methods for actually solving the differential equation (47). A main aspect here is the selection of suitable basis functions, briefly discussed in Sec. 5.4. The third type of approximation involves constructing an expression for the unknown xc functional $E_{xc}[n]$, which contains all many-body aspects of the problem [cf. Eq. (38)]. It is with this type of approximation that we are concerned in the present section.

This section is intended to give the reader an idea of what types of functionals exist, and what their main features are. It does not deal with the actual construction of these functionals. For this the reader is referred to the reviews [3]-[15] or to the original papers cited below. Sticking to the bird's-eye philosophy of this overview I have also refrained from including numerical data on the performance of each functional — extensive comparisons of a wide variety of functionals can be found in Refs. [3]-[15] and in the original literature cited below.

6.1   Local functionals: LDA

Historically (and in many applications also practically) the most important type of approximation is the local-density approximation (LDA). To understand the concept of an LDA recall first how the noninteracting kinetic energy $T_s[n]$ is treated in the Thomas-Fermi approximation: In a homogeneous system one knowns that, per volume[17]

$$t_s^{hom}(n) = \frac{3\hbar^2}{10m}(3\pi^2)^{2/3}n^{5/3} \tag{53}$$

where $n = const$. In an inhomogeneous system, with $n = n(\mathbf{r})$, one approximates locally

$$t_s(\mathbf{r}) \approx t_s^{hom}(n(\mathbf{r})) = \frac{3\hbar^2}{10m}(3\pi^2)^{2/3}n(\mathbf{r})^{5/3} \tag{54}$$

and obtains the full kinetic energy by integration over all space

$$T_s^{LDA}[n] = \int d^3r\, t_s^{hom}(n(\mathbf{r})) = \frac{3\hbar^2}{10m}(3\pi^2)^{2/3}\int d^3r\, n(\mathbf{r})^{5/3}. \tag{55}$$

For the kinetic energy the approximation $T_s[n] \approx T_s^{LDA}[n]$ is much inferior to the exact treatment of T_s in terms of orbitals, offered by the Kohn-Sham equations, but the LDA concept turned out to be highly useful for another component of the total energy (38), the exchange-correlation energy $E_{xc}[n]$. For the exchange energy $E_x[n]$ the procedure is extremely simple, since one knows [3, 4]

$$e_x^{hom}(n) = -\frac{3q^2}{4}\left(\frac{3}{\pi}\right)^{1/3}n^{4/3}, \tag{56}$$

so that

$$E_x^{LDA}[n] = -\frac{3q^2}{4}\left(\frac{3}{\pi}\right)^{1/3}\int d^3r\, n(\mathbf{r})^{4/3}. \tag{57}$$

This is the LDA for E_x.[18]

   For the correlation energy $E_c[n]$ the situation is more complicated, since $e_c^{hom}(n)$ is not known exactly: the determination of the correlation energy of a homogeneous interacting electron system is already a difficult many-body problem on its own. Early approximate expressions for $e_c^{hom}(n)$ were based on applying perturbation theory (e.g. the random-phase approximation) to this problem [46, 47]. These approximations became outdated with the advent of highly precise Quantum Monte Carlo (QMC) calculations for the electron liquid, by Cerperley and Alder [48]. Modern expressions for $e_c^{hom}(n)$ [49, 50, 51] are parametrizations of these data. These expressions are implemented in most standard DFT program packages and in typical applications give almost identical results. On the other hand, the earlier parametrizations of the LDA, based on perturbation theory [46, 47], can deviate substantially from the QMC ones, and are better avoided.

---

[17]The change from a capital T to a lower-case t is commonly used to indicate quantities per volume.

[18]If one adds this term to the Thomas-Fermi expression (35) one obtains the so-called Thomas-Fermi-Dirac approximation to $E[n]$. It one multiplies it with an adjustable parameter α one obtains the so-called $X\alpha$ approximation to $E_{xc}[n]$. These approximations are not much used today.

Independently of the parametrization, the LDA for $E_{xc}[n]$ formally consists in

$$E_{xc}[n] \approx E_{xc}^{LDA}[n] = \int d^3r \left[e_x^{hom}(n(\mathbf{r})) + e_c^{hom}(n(\mathbf{r})) \right]. \tag{58}$$

This approximation for $E_{xc}[n]$ has proved amazingly successful, even when applied to systems that are quite different from the electron liquid that forms the reference system for the LDA. For many decades the LDA has been applied in, e.g., calculations of band structures and total energies in solid-state physics. In quantum chemistry it is much less popular, because it fails to provide results that are accurate enought to permit a quantitative discussion of the chemical bond in molecules (so-called 'chemical accuracy' requires calculations with an error of not more than about $1kcal/mole \approx 0.0434eV/particle$).

6.2 Semilocal functionals: GEA, GGA, and beyond

In the LDA one exploits knowledge of the density at point $\mathbf{r}$. Any real system is spatially inhomogeneous, i.e., it has a spatially varying density $n(\mathbf{r})$, and it would clearly be useful to also include information on the rate of this variation in the functional. A first attempt at doing this were the so-called 'gradient-expansion approximations' (GEA). In this class of approximation one tries to systematically calculate gradient-corrections of the form $|\nabla n(\mathbf{r})|$, $|\nabla n(\mathbf{r})|^2$, $\nabla^2 n(\mathbf{r})$, etc, to the LDA. A famous example is the lowest-order gradient correction to the Thomas-Fermi approximation for $T_s[n]$,

$$T_s[n] \approx T_s^W[n] = T_s^{LDA}[n] + \frac{\hbar^2}{8m} \int d^3r \frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})}. \tag{59}$$

This second term on the right-hand side is called the Weizsäcker term.[19] Similarly, in

$$E_x[n] \approx E_x^{GEA(2)}[n] = E_x^{LDA}[n] - \frac{10q^2}{432\pi(3\pi^2)^{1/3}} \int d^3r \frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})^{4/3}} \tag{60}$$

the second term on the right-hand side is the lowest-order gradient correction to $E_x^{LDA}[n]$. In practice, the inclusion of low-order gradient corrections almost never improves on the LDA, and often even worsens it. Higher-order corrections (e.g., $\propto |\nabla n(\mathbf{r})|^\alpha$ or $\propto \nabla^\beta n(\mathbf{r})$ with $\alpha, \beta > 2$), on the other hand, are exceedingly difficult to calculate, and little is known about them.

In this situation it was a major breakthrough when it was realized, in the early eigthies, that instead of power-series-like systematic gradient expansions one could experiment with more general functions of $n(\mathbf{r})$ and $\nabla n(\mathbf{r})$, which need not proceed order by order. Such functionals, of the general form

$$E_{xc}^{GGA}[n] = \int d^3r \, e(n(\mathbf{r}), \nabla n(\mathbf{r})), \tag{61}$$

have become known as generalized-gradient approximations (GGAs) [52].

---

[19]If one adds this term to the Thomas-Fermi expression (35) one obtains the so-called Thomas-Fermi-Weizsäcker approximation to $E[n]$. In a systematic gradient expansion the 8 in the denominator is replaced by a 72 [3, 4].

Different GGAs differ in the choice of the function $e(n(\mathbf{r}), \nabla n(\mathbf{r}))$. Note that this makes different GGAs much more different from each other than the different parametrizations of the LDA: essentially there is only one correct expression for $e_{xc}^{hom}(n)$, and the various parametrizations of the LDA [46, 47, 49, 50, 51] are merely different ways of writing it. On the other hand, depending on the method of construction employed for obtaining $e(n(\mathbf{r}), \nabla n(\mathbf{r}))$ one can obtain very different GGAs. In particular, GGAs used in chemistry typically proceed by fitting parameters to test sets of selected molecules. On the other hand, GGAs used in physics tend to emphasize exact constraints. Nowadays the most popular (and most reliable) GGAs are PBE (denoting the functional proposed in 1996 by Perdew, Burke, and Ernzerhof [53]) in physics, and BLYP (denoting the combination of Becke's 1988 exchange functional [54] with the 1988 correlation functional of Lee, Yang, and Parr [55]) in chemistry. Many other GGA-type functionals are also available, and new ones continue to appear.

Quite generally, current GGAs seem to give reliable results for all main types of chemical bond (covalent, ionic and metallic). For van der Waals interactions, however, most GGAs fail (PBE may be an exception [56]). To describe these very weak interactions several more specialized approaches have been developed within DFT [57, 58, 59, 60]. Both in physics and in chemistry the widespread use of GGAs has lead to major improvements as compared to LDA. 'Chemical accuracy', as defined above, has not yet been attained, but is not too far away either. A useful collection of explicit expressions for some GGAs can be found in the appendix of Ref. [61], and more detailed discussion of some selected GGAs and their performance is given in Ref. [62] and in the chapter of Kurth and Perdew in Ref. [12].

In spite of these advances, the quest for more accurate functionals goes ever on, and both in chemistry and physics various beyond-GGA functionals have appeared. Perhaps the most popular functional in chemistry[20] is B3LYP. This is a combination of the LYP GGA for correlation [55] with Becke's three-parameter hybrid functional B3 for exchange [63]. Common hybrid functionals, such as B3, mix a fraction of Hartree-Fock exchange into the DFT exchange functional (other mixtures are also possible). This mixing involves a certain amount of empiricism (the weight factors given to the HF and DFT exchange terms are adjustable) and optimization for selected classes of molecules. Formally, this might be considered a drawback, but in practice B3 has proven to be the most successful exchange functional for chemical applications, in particular when combined with the LYP functional for E_c. More extreme examples of this semiempirical mode of construction of functionals are Becke's 1997 hybrid functional [64], which contains 10 adjustable parameters, and the functionals of Refs. [65] and [66], each of which contains 21 parameters.

In common applications of such hybrid functionals self-consistency is enforced with respect to the single-particle orbitals (as in the Hartree-Fock Kohn-Sham method [4]), and not with respect to the density itself. Self-consistency with respect to the density would require to treat the Fock term in the hybrid functional as an implicit density functional. One way to do this is the OEP scheme discussed in the next section.

Another recent beyond-GGA development is the emergence of so-called Meta-GGAs, which depend, in addition to the density and its derivatives, also on the Kohn-Sham

---

[20]This is written in early 2002, but things can change rapidly in this field.

kinetic-energy density $\tau(\mathbf{r})$ [67, 68]

$$\tau(\mathbf{r}) = \frac{\hbar^2}{2m} \sum_i |\nabla \phi_i(\mathbf{r})|^2, \qquad (62)$$

so that E_{xc} can be written as $E_{xc}[n(\mathbf{r}), \nabla n(\mathbf{r}), \tau(\mathbf{r})]$. In several recent tests [69, 70, 71] Meta-GGAs have given favorable results, even when compared to the best GGAs, but the full potential of this type of approximation has not yet been explored systematically.

6.3   Nonlocal and orbital functionals

As we have seen in the case of T_s, it can be much easier to represent a functional in terms of single-particle orbitals than directly in terms of the density. Such functionals are known as orbital functionals, and Eq. (37) constitutes a simple example. Another important orbital-dependent functional is the exchange energy (Fock term) of Eq. (39). The Meta-GGAs and hybrid functionals mentioned in the previous section are also orbital functionals, because they depend on the kinetic energy density (62), and on a combination of the orbital functional (39) with ordinary GGAs, respectively.

Since E_x (and all other orbital functionals) depend on the density only implicitly, via the orbitals $\phi_i[n]$, it is not possible to directly calculate the functional derivative $v_{xc} = \delta E_{xc}/\delta n$. Instead one must use indirect approaches to minimize $E[n]$ and obtain v_{xc}. In the case of the kinetic energy functional $T_s[\{\phi_i[n]\}]$ this indirect approach is the Kohn-Sham scheme, described in Sec. 5. In the case of orbital expressions for E_{xc} the corresponding indirect scheme is known as the optimized effective potential (OEP) [72] or, equivalently, the optimized-potential model (OPM) [73]. The minimization of the orbital functional with respect to the density is achieved by repeated application of the chain rule for functional differentiation,

$$v_{xc}[n](\mathbf{r}) = \frac{\delta E_{xc}^{orb}[\{\phi_i\}]}{\delta n(\mathbf{r})} = \int d^3r' \int d^3r'' \sum_i \left[\frac{\delta E_{xc}^{orb}[\{\phi_i\}]}{\delta \phi_i(\mathbf{r}')} \frac{\delta \phi_i(\mathbf{r}')}{\delta v_s(\mathbf{r}'')} \frac{\delta v_s(\mathbf{r}'')}{\delta n(\mathbf{r})} + c.c. \right], \quad (63)$$

where E_{xc}^{orb} is the orbital functional (e.g., the Fock term). Further evaluation of Eq. (63) gives rise to an integral equation that determines the $v_{xc}[n]$ belonging to the chosen orbital functional $E_{xc}[\{\phi_i[n]\}]$ [72, 74]. As an alternative to solving the full OEP equation, Krieger, Li and Yafrate (KLI) have proposed a simple but surprisingly accurate approximation that greatly facilitates implementation of the OEP [72].

The application of the OEP methodology, either with or without the KLI approximation, to the Fock term (39) is also known as the exact-exchange method (EXX). The OEP-EXX equations have been solved for atoms [72, 73, 75] and solids [76, 77], with very encouraging results. Other orbital-dependent functionals that have been treated within the OEP scheme are the self-interaction correction [37] and the Colle-Salvetti functional [75]. A detailed review of the OEP and its KLI approximation is Ref. [74].

The high accuracy attained in applications of the OEP, and the fact that it is easier to devise orbital functionals than explicit density functionals, makes the OEP concept attractive, but the computational cost of solving the OEP integral equation is a major drawback. However, this computational cost is significantly reduced by the KLI approximation. In the context of the EXX method (i.e., using the Fock term as one's orbital

functional) this is the common way to proceed. A further reduction of computational complexity is achieved by not evaluating the orbital functional self-consistently, using (63), but only once, following a converged self-consistent LDA or GGA calculation. This 'post-GGA' or 'post-LDA' strategy completely avoids the OEP and has been used both for hybrid functionals and Meta-GGAs [63, 64, 67, 68]. In the case of hybrid functionals, still another mode of implementation has become popular. This alternative, which also avoids solution of Eq. (63), is to calculate the derivative of the Fock contribution to the hybrid with respect to the single-particle orbitals, instead of the density as in (63), while the other terms in the hybrid are still treated as explicit density functionals. The differential equation one then has to solve is of Hartree-Fock form, but with a weight factor in front of the Fock term, and some explicit density functionals added to it. This scheme, in which self-consistency is obtained with respect to the single-particle orbitals, can be considered an evolution of the Hartree-Fock Kohn-Sham method [4], and is how hybrids are commonly implemented.

Still another type of orbital functional is the self-interaction correction (SIC) [50]. This correction can be applied on top of any other density functional, and ensures that the resulting corrected functional satisfies $E_c[n] \equiv 0$ and $E_x[n] \equiv -E_H[n]$ for a one-electron system. The LDA is exact for a completely uniform system, and thus is self-interaction free in this limit, but neither it not common GGAs or Meta GGAs satisfy the requirement of freedom from self-interaction in general [67]. This is particularly critical for localized states, such as the d states in transition-metal oxides. For such systems SIC has been shown to greatly improve the uncorrected LDA [78, 79]. Unfortunately the SIC approach does not lead to Kohn-Sham equations of the usual form. As a consequence, a number of specialized algorithms for minimizing the SIC energy functional have been developed. For more details on these algorithms and some interesting applications in solid-state physics see Refs. [78, 79, 80]. For finite systems the SIC functional has also been implemented by means of the OEP [37]. A detailed review of implementations and applications of SIC can be found in the contribution of Temmerman et al. in Ref. [13].

Apart from orbital functionals, which are implicit nonlocal density functionals because the orbitals depend on the density in a nonlocal way, there is also a class of explicit nonlocal density functionals. Such nonlocal density functionals take into account, at any point $\mathbf{r}$, not only the density at that point, $n(\mathbf{r})$, and its derivatives, $\nabla n(\mathbf{r})$ etc, but also the behaviour of the density at different points $\mathbf{r}' \neq \mathbf{r}$. A typical example is

$$E_{xc}^{ADA}[n] = \int d^3r \, e_{xc}^{hom}(\bar{n}(\mathbf{r})). \tag{64}$$

In the LDA one would have $\bar{n}(\mathbf{r}) \equiv n(\mathbf{r})$, but in the average-density approximation (ADA) one takes [81]

$$\bar{n}(\mathbf{r}) = \int d^3r' \, n(\mathbf{r}')w[n](|\mathbf{r} - \mathbf{r}'|), \tag{65}$$

where $w[n](|\mathbf{r} - \mathbf{r}'|)$ is a weighting function that samples the density not only semilocally, as do the GGAs, but over a volume determined by the range of w. Conceptually similar to the ADA, but slightly more complicated formally and computationally, is the weighted-density approximation (WDA) [81]. The dependence of the ADA functional on the integral over $n(\mathbf{r})$, instead of on derivatives as in the GGAs, is the reason why such functionals are called nonlocal. In practice, this integral turns the functionals computationally expensive,

and in spite of their great promise they are much less used than GGAs. However, recent comparisons of ADA and WDA with LDA and GGAs for low-dimensional systems [69, 82] and for bulk silicon [83] show that nonlocal density functionals can outperform local and semilocal approximations.

7 Extensions of DFT: New frontiers and old problems

Up to this point we have discussed DFT in terms of the charge (or particle) density $n(\mathbf{r})$ as fundamental variable. In order to reproduce the correct charge density of the interacting system in the noninteracting (Kohn-Sham) system, one must apply to the latter the effective KS potential $v_s = v + v_H + v_{xc}$, in which the last two terms simulate the effect of the electron-electron interaction on the charge density. This form of DFT, which is the one proposed originally [20], could also be called 'charge-only' DFT. It is not the most widely used DFT in practical applications. Much more common is a formulation that employs one density for each spin, $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$, i.e, works with two fundamental variables. In order to reproduce both of these in the noninteracting system one must now apply two effective potentials, $v_{s,\uparrow}(\mathbf{r})$ and $v_{s,\downarrow}(\mathbf{r})$.[21] This formulation of DFT is known as spin-DFT (SDFT) [46, 47]. Its fundamental variables $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$ can be used to calculate the charge density $n(\mathbf{r})$ and the spin-magnetization density $m(\mathbf{r})$ from

$$n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r}) \tag{66}$$

$$m(\mathbf{r}) = \mu_0(n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r})), \tag{67}$$

where $\mu_0 = q\hbar/2mc$ is the Bohr magneton. More generally, the Hohenberg-Kohn theorem of SDFT states that in the presence of a magnetic field $B(\mathbf{r})$ that couples only to the electron spin [via the familiar Zeeman term $\int d^3r\, m(\mathbf{r})B(\mathbf{r})$] the ground-state wave function and all ground-state observables are unique functionals of n and m or, equivalently, of $n_\uparrow$ and $n_\downarrow$.[22] Almost the entire further development of the HK theorem and the KS equations can be immediately rephrased for SDFT, just by attaching a suitable spin index to the densities. For this reason we could afford the luxury of exclusively discussing 'charge-only' DFT in the preceeding sections, without missing any essential aspects of SDFT.

There are, however, some exceptions to this simple rule. One is the fourth statement of the HK theorem, as discussed in Sec. 4. Another is the construction of functionals. For the exchange energy it is known, e.g., that [84]

$$E_x^{SDFT}[n_\uparrow, n_\downarrow] = \frac{1}{2}\left(E_x^{DFT}[2n_\uparrow] + E_x^{DFT}[2n_\downarrow]\right). \tag{68}$$

In analogy to the coordinate scaling of Eqs. (41) - (43), this property is often called 'spin-scaling', and it can be used to construct an SDFT exchange functional from a given

---

[21]More generally, one requires one effective potential for each density-like quantity to be reproduced in the KS system. Such potentials and corresponding densities are called conjugate variables.

[22]In the particular case $B = 0$ the SDFT HK theorem still holds and continues to be useful, e.g., for systems with spontaneous polarization. In principle one could also use 'charge-only' DFT to study such systems, but then $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$ become functionals of $n(\mathbf{r})$ and nobody knows how to determine these functionals.

DFT exchange functional. In the context of the LSDA, von Barth and Hedin [46] wrote the exchange functional in terms of an interpolation between the unpolarized and fully polarized electron gas which by construction satisfies Eq. (68). Alternative interpolation procedures can be found in Ref. [49]. GGA exchange functionals also satisfy Eq. (68) by construction. For the correlation energy no scaling relation of the type (68) holds, so that in practice correlation functionals are either directly constructed in terms of the spin densities or written by using, without formal justification, the same interpolation already used in the exchange functional. In the case of the LSDA this latter procedure was introduced in Ref. [46] and further analysed in Ref. [49].

The Kohn-Sham equations of SDFT are

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + v_{s\sigma}(\mathbf{r}) \right] \phi_{i\sigma}(\mathbf{r}) = \epsilon_{i\sigma} \phi_{i\sigma}(\mathbf{r}), \tag{69}$$

where $v_{s\sigma}(\mathbf{r}) = v_\sigma(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc,\sigma}(\mathbf{r})$. In a nonrelativistic calculation the Hartree term does not depend on the spin label,[23] but in the presence of an externally applied magnetic field $v_\sigma(\mathbf{r}) = v(\mathbf{r}) - \sigma \mu_0 B$ (where $\sigma = \pm 1$). Finally,

$$v_{xc,\sigma}(\mathbf{r}) = \frac{\delta E_{xc}^{SDFT}[n_\uparrow, n_\downarrow]}{\delta n_\sigma(\mathbf{r})}. \tag{70}$$

In the presence of an internal magnetic field B_{xc} (i.e., in spin-polarized systems) $v_{xc,\downarrow} - v_{xc,\uparrow} = \mu_0 B_{xc}$. This field is the origin of, e.g., ferromagnetism in transition metals. A more detailed discussion of SDFT can be found in Refs. [3, 4, 47], and a particularly clear exposition of the construction of xc functionals for SDFT is the contribution of Kurth and Perdew in Ref. [12].

If the direction of the spins is not uniform in space[24] one requires a formulation of SDFT in which the spin magnetization is not a scalar, as above, but a three-component vector $\mathbf{m}(\mathbf{r})$. Different proposals for extending SDFT to this situation are available [85, 86, 87]. One mechanism that can give rise to noncollinear magnetism is spin-orbit coupling. This is another relativistic effect [23], and as such it is not consistently treated in either DFT or SDFT. A generalization of DFT that does account for spin-orbit coupling and other relativistic effects is relativistic DFT (RDFT) [88, 89]. Here the fundamental variable is the relativistic four current j^μ. RDFT requires a more drastic reformulation of DFT than does SDFT. In particular, the KS equation of RDFT is now of the form of the single-particle Dirac equation, instead of the Schrödinger equation. There are also many subtle questions involving renormalizability and the use of the variational principle in the presence of negative energy states. For details on these problems and their eventual solution the reader is referred to the chapters by Engel et al. in Refs. [8] and [15], and to the book by Eschrig [14]. A didactical exposition of RDFT, together with representative applications in atomic and condensed-matter physics, can be found in the book by Strange [23], and a recent numerical implementation is presented in Ref. [90].

[23]Spin-spin dipolar interactions are a relativistic effect of order $(1/c)^2$, as are current-current interactions.
[24]Such 'noncollinear magnetism' appears, e.g., as canted or helical spin configurations in rare-earth compounds, or as domain walls in ferromagnets.

To study the magnetic properties of matter one would often like to be able to obtain information on the currents in the system and their coupling to possible external magnetic fields. Important classes of experiments for which this information is relevant are nuclear magnetic resonance and the quantum Hall effects. SDFT does not provide explicit information on the currents. RDFT in principle does, but standard implementations of it are formulated in a spin-only version, which prohibits extraction of information on the currents. Furthermore, the formalism of RDFT is considerably more complicated than that of SDFT. In this situation the formulation of nonrelativistic current-DFT (CDFT), accomplished by Vignale and Rasolt [91, 92], was a major step forward. CDFT is formulated explicitly in terms of the (spin) density and the nonrelativistic paramagnetic current density vector $\mathbf{j}_p(\mathbf{r})$. Some recent applications of CDFT are Refs. [93, 94, 95, 96]. E. K. U. Gross and the author have shown that the existence of *spin currents* implies the existence of a link between the xc functionals of SDFT and those of CDFT [97]. Conceptually, this link is similar to the one of Eq. (68) between functionals of DFT and SDFT, but the details are quite different. Approximations for xc functionals of CDFT are discussed in Refs. [97, 98].

Apart from SDFT, RDFT, and CDFT, there exist many other generalizations of DFT that were designed for one or other special purpose. As examples we mention superconductivity [99, 100, 101, 102] and spin-density waves [87, 103], but there are many more [3]-[15]. For reasons of space we cannot discuss these extensions here. Instead, let us take a brief look at a problem that requires more radical departures from the framework of conventional DFT: excited states. DFT is formulated in terms of ground-state densities, and it is not immediately obvious how one could extract information on excited states from them (although at least in the case of 'charge-only' DFT the fourth substatement of the HK theorem guarantees that this must be possible).

Apart from the *ad hoc* identification of the KS eigenvalues with true excitation energies, there are at least three more sound approaches to excited states in DFT that have met with some success. The early suggestion of Gunnarsson and Lundqvist [47] to use a symmetry-dependent xc functional to calculate the lowest-energy excited state of each symmetry class has been implemented approximately by von Barth [104], but suffers from lack of knowledge on the symmetry dependence of the functional. More recent work on this dependence is Ref. [105]. An alternative approach to excited states, not restricted to the lowest energy state of a given symmetry, is ensemble DFT, developed by Theophilou [106] and Oliveira, Gross, and Kohn [107]. In this formalism the functional depends on the particular choice for the ensemble, and a simple approximation for this dependence is available [107]. Some applications of this method have been worked out by Nagy [108]. Other recent DFT approaches to excited states can be found in Refs. [109], [110] and [111].

Perhaps the most radical departure from the framework of conventional DFT is time-dependent DFT (TD-DFT). The time-dependent generalization of the HK theorem, the Runge-Gross theorem, cannot be proven along the lines of the original HK theorem, but requires a different approach [112, 113]. For recent reviews of TD-DFT see Ref. [114]. Excited states have first been extracted from TD-DFT in Ref. [115]. This approach is now implemented in standard quantum-chemical DFT program packages [116, 117] and is increasingly applied also in solid-state physics [44]. Another important application of TD-

DFT is to systems in external time-dependent fields, such as atoms in strong laser fields [118, 119]. First steps towards studying dynamical magnetic phenomena with TD-SDFT have been taken very recently [120].

A very different way of using DFT, which does not depend directly on approximate solution of Kohn-Sham equations, is the quantification and clarification of traditional chemical concepts, such as electronegativity [4], hardness, softness, Fukui functions, and other reactivity indices [4, 121], or aromaticity [122]. The true potential of DFT for this kind of investigation is only beginning to be explored.

All extensions of DFT face the same formal questions (e.g., noninteracting v-representability of the densities, nonuniqueness of the KS potentials, meaning of the KS eigenvalues) and practical problems (e.g., how to efficiently solve the KS equations, how to construct accurate approximations to E_{xc}, how to treat systems with very strong correlations) as do the more widely used 'charge-only' DFT and SDFT. These questions and problems, however, have never stopped DFT from advancing, and at present DFT emerges as the method of choice for solving a wide variety of quantum mechanical problems in chemistry and physics — and in many situations, such as large and inhomogeneous systems, it is the only applicable first-principles method at all.

The future of DFT is bright [1, 36, 123] — but to be able to contribute to it, the reader must now leave the present superficial overview behind, and turn to the more advanced treatments available in the literature [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

References

[1] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).

[2] J. A. Pople, Rev. Mod. Phys. **71**, 1267 (1999).

[3] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).

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

[5] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (John Wiley & Sons, New York, 2001).

[6] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[7] J. M. Seminario (Ed.), *Recent Developments and Applications of Modern DFT* (Elsevier, Amsterdam, 1996).

[8] R. F. Nalewajski (Ed.), *Density Functional Theory I -IV* (Springer, Topics in Current Chemistry Vols. 180-183, 1996).

[9] V. I. Anisimov (Ed.), *Strong Coulomb Correlations in Electronic Structure Calculations: Beyond the Local Density Approximation* (Gordon & Breach, 1999).

[10] N. H. March, *Electron Density Theory of Atoms and Molecules* (Academic Press, London, 1992).

[11] B. B. Laird, R. B. Ross, and T. Ziegler (Eds.), *Chemical Applications of Density Functional Theory*, (American Chemical Society, 1996).

[12] D. Joulbert (Ed.), *Density Functionals: Theory and Applications* (Springer, Lecture Notes in Physics Vol. 500, 1998)

[13] J. F. Dobson, G. Vignale, and M. P. Das (Eds.), *Density Functional Theory: Recent Progress and New Directions* (Plenum, New York, 1998).

[14] H. Eschrig, *The Fundamentals of Density Functional Theory* (Teubner, Leipzig, 1996).

[15] E. K. U. Gross and R. M. Dreizler (Eds.), *Density Functional Theory* (Plenum, New York, 1995).

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

[17] E. N. Economou, *Green's Functions in Quantum Physics* (Springer, New York, 1979).

[18] E. K. U. Gross, E. Runge, O. Heinonen, *Many Particle Theory* (Adam Hilger, 1992).

[19] E. Schrödinger, Ann. Physik **79**, 489 (1926).

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

[21] M. Levy, Phys. Rev. A **26**, 1200 (1982).

[22] E. H. Lieb in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providencia, (Plenum, New York,1985).

[23] P. Strange, *Relativistic Quantum Mechanics with Applications in Condensed Matter and Atomic Physics* (Cambridge University Press, Cambridge, 1998).

[24] T. Kato, Commun. Pure Appl. Math. **10** 151 (1957).

[25] N. H. March, *Self-consistent fields in atoms* (Pergamon Press, Oxford, 1975).

[26] K. Capelle and G. Vignale, Phys. Rev. Lett. **86**, 5546 (2001).

[27] K. Capelle and G. Vignale, Phys. Rev. B **65**, 113106 (2002).

[28] H. Eschrig and W. E. Pickett, Solid State Commun. **118**, 123 (2001).

[29] N. Argaman and G. Makov, Phys. Rev. B **66**, 052413 (2002).

[30] L. W. Wang and M. P. Teter, Phys. Rev. B **45**, 13196 (1992).

[31] M. Foley and P. A. Madden, Phys. Rev. B **53**, 10589 (1996).

[32] E. K. U. Gross, M. Petersilka and T. Grabo in Ref. [11].

[33] N. D. Mermin, Phys. Rev. **137**, 1441 (1965).

[34] M. Levy and J. P. Perdew, Phys. Rev. A **32**, 2010 (1985).

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

[36] W. Kohn, A. D. Becke, and R. G. Parr, J. Phys. Chem. **100**, 12974 (1996).

[37] J. Chen, J. B. Krieger, Y. Li and G. J. Iafrate, Phys. Rev. A **54**, 3939 (1996).

[38] M. Lüders, A. Ernst, W. M. Temmerman, Z. Szotek, and P. J. Durham, J. Phys. Cond. Mat. **13**, 8587 (2001).

[39] J. Muskat, A. Wander, and N. M. Harrison, Chem. Phys. Lett. **342**, 397 (2001).

[40] A. Savin, C. J. Umrigar, and X. Gonze, Chem. Phys. Lett. **288**, 391 (1998).

[41] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (McGraw-Hill, New York, 1989).

[42] W. G. Aulbur, L. Jönsson, and J. W. Wilkins, Solid State Phys. **54**, 1 (1999).

[43] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).

[44] G. Onida, L. Reining, and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[45] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[46] U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).

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

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

[49] S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).

[50] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[51] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1993).

[52] J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).

[53] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996). *ibid* **78**, 1396(E) (1997).

[54] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[55] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[56] D. C. Patton and M. R. Pederson, Phys. Rev. A **56**, 2495 (1997).

[57] W. Kohn, Y. Meir, and D. E. Makarov, Phys. Rev. Lett. **80**, 4153 (1998).

[58] M. Lein, J. F. Dobson, and E. K. U. Gross, J. Comp. Chem. **20**, 12 (1999).

[59] J. F. Dobson and B. P. Dinte, Phys. Rev. Lett. **76**, 1780 (1996).

[60] Y. Andersson, D. C. Langreth, and B. I. Lundqvist, Phys. Rev. Lett. **76**, 102 (1996).

[61] C. Filippi, C. J. Umrigar, and M. Taut, J. Chem. Phys. **100**, 1290 (1994).

[62] P. Ziesche, S. Kurth, and J. P. Perdew, Comp. Mat. Sci. **11**, 122 (1998).

[63] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[64] A. D. Becke, J. Chem. Phys. **107**, 8554 (1997). See also A. D. Becke, J. Comp. Chem. **20**, 63 (1999).

[65] D. J. Tozer and N. C. Handy, J. Chem. Phys. **108**, 2545 (1998).

[66] T. van Voorhis and G. E. Scuseria, J. Chem. Phys. **109**, 400 (1998).

[67] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha, Phys. Rev. Lett. **82**, 2544 (1999).

[68] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).

[69] Y. H. Ki, I. H. Lee, S. Nagaraja, J. P. Leburton, R. Q. Hood, and R. M. Martin, Phys. Rev. B **61**, 5202 (2000).

[70] S. Kurth, J. P. Perdew, and P. Blaha, Int. J. Quantum Chem. **75**, 889 (1999).

[71] C. Adamo, M. Ernzerhof, and G. E. Scuseria, J. Chem. Phys. **112**, 2643 (2000).

[72] J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992); *ibid* **46**, 5453 (1992); *ibid* **47** 165 (1993).

[73] E. Engel and S. H. Vosko, Phys. Rev. A **47**, 2800 (1993).

[74] T. Grabo, T. Kreibich, S. Kurth, and E. K. U. Gross, in Ref. [9].

[75] T. Grabo and E. K. U. Gross, Int. J. Quantum Chem. **64**, 95 (1997). T. Grabo and E. K. U. Gross, Chem. Phys. Lett. **240**, 141 (1995).

[76] T. Kotani, Phys. Rev. Lett. **74**, 2989 (1995).

[77] M. Stadele, J. A. Majewski, P. Vogl, and A. Görling, Phys. Rev. Lett. **79**, 2089 (1997). M. Stadele, M. Moukara, J. A. Majewski, P. Vogl and A. Gorling, Phys. Rev. B **59**, 10031 (1999). Y.-H. Kim, M. Stadele and R. M. Martin, Phys. Rev. A **60**, 3633 (1999).

[78] A. Svane and O. Gunnarsson, Phys. Rev. Lett. **65**, 1148 (1990); *ibid* **72**, 1248 (1994).

[79] Z. Szotek, W. .M. Temmermann, and H. Winter, Phys. Rev. Lett. **72**, 1244 (1994).

[80] P. Strange, A. Svane, W. M. Temmermann, Z. Szotek, and H. Winter, Nature **399**, 756 (1999).

[81] O. Gunnarsson, M. Jonson, and B. I. Lundqvist, Phys. Rev. B **20**, 3136 (1979).

[82] A. Cancio, M. Y. Chou, and R. O. Hood, Phys. Rev. B **64**, 115112 (2002).

[83] P. Garcia-Gonzalez, Phys. Rev. B **62**, 2321 (2000).

[84] G. L. Oliver and J. P. Perdew, Phys. Rev. A **20**, 397 (1979).

[85] L. M. Sandratskii, Adv. Phys. **47**, 91 (1998).

[86] L. Nordström and D. Singh, Phys. Rev. Lett. **76**, 4420 (1996).

[87] K. Capelle and L. N. Oliveira, Phys. Rev. B **61**, 15228 (2000).

[88] A. K. Rajagopal and J. Callaway, Phys. Rev. B **7**, 1912 (1973).

[89] A. H. Mac Donald and S. H. Vosko, J. Phys. C **12**, 2977 (1979).

[90] E. Engel, T. Auth, and R. M. Dreizler, Phys. Rev. B **64**, 235126 (2001).

[91] G. Vignale and M. Rasolt, Phys. Rev. Lett. **59**, 2360 (1987).

[92] G. Vignale and M. Rasolt, Phys. Rev. B **37**, 10685 (1988).

[93] M. Ferconi and G. Vignale, Phys. Rev. B **50**, 14722 (1997).

[94] M. Battocletti, H. Ebert, and E. K. U. Gross, Europhys. Lett. **40**, 545 (1997).

[95] A. M. Lee, N. C. Handy, and S. M. Colwell, J. Chem. Phys. **103**, 10095 (1996).

[96] K. Capelle, Phys. Rev. B **65**, 100515 (2002).

[97] K. Capelle and E. K. U. Gross, Phys. Rev. Lett. **78**, 1872 (1997).

[98] P. Skudlarski and G. Vignale, Phys. Rev. B **48**, 8547 (1993).

[99] L. N. Oliveira, E. K. U. Gross, and W. Kohn, Phys. Rev. Lett. **60**, 2430 (1988).

[100] W. Kohn, E. K. U. Gross and L. N. Oliveira, J. de Physique **50**, 2601 (1989).

[101] W. M. Temmermann, Z. Szotek, B. L. Gyorffy, O. K. Andersen, and O. Jepsen, Phys. Rev. Lett. **76**, 307 (1996).

[102] S. Kurth, M. Marques, M. Lüders, and E. K. U. Gross, Phys. Rev. Lett. **83**, 2628 (1999).

[103] K. Capelle and L. N. Oliveira, Europhys. Lett. **49**, 376 (2000).

[104] U. von Barth, Phys. Rev. A **20**, 1693 (1979).

[105] A. Görling, Phys. Rev. Lett. **85**, 4229 (2000).

[106] A. K. Theophilou, J. Phys. C **12**, 5419 (1979).

[107] E. K. U. Gross, L. N. Oliveira, and W. Kohn, Phys. Rev. A **37**, 2805; *ibid* 2809; *ibid* 2821 (1988).

[108] A. Nagy, Phys. Rev. A **49**, 3074 (1994), *ibid* **42**, 4388 (1990).

[109] A. Görling, Phys. Rev A **59**, 3359 (1999).

[110] A. Nagy, Chem. Phys. Lett. **296**, 489 (1998). M. Levy and A. Nagy, Phys. Rev. Lett. **83**, 4361 (1999).

[111] K. Capelle, preprint [available as cond-mat/0209364].

[112] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).

[113] R. van Leeuwen, Phys. Rev. Lett. **82**, 3863 (1999).

[114] E. K. U. Gross, J. F. Dobson, and M. Petersilka in Ref. [8]. K. Burke and E. K. U. Gross in Ref. [12].

[115] M. Petersilka, U. J. Gossmann and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996). See also T. Grabo, M. Petersilka and E. K. U. Gross, J. Mol. Struc. (Theochem) **501**, 353 (2000).

[116] S. J. A. van Gisbergen, J. G. Snijders, G. te Velde, and E. J. Baerends, Comp. Phys. Comm. **118**, 119 (1999).

[117] R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, J. Chem. Phys. **109**, 8218 (1998).

[118] M. Lein, E. K. U. Gross and V. Engel, Phys. Rev. Lett. **85**, 4707 (2000).

[119] C. A. Ullrich and E. K. U. Gross, Comm. At. Mol. Phys. **33**, 211 (1997).

[120] K. Capelle, G. Vignale, and B. L. Gyorffy, Phys. Rev. Lett. **87**, 206403 (2001). K. Capelle and B. L. Gyorffy, Europhys. Lett. **61**, 354 (2003).

[121] H. Chermette, J. Comp. Chem. **20**, 129 (1999).

[122] F. De Proft and P. Geerlings, Chem. Rev. **101**, 1451 (2001).

[123] A. E. Mattsson, Science **298**, 759 (2002).