2
Review of ground-state density-functional theory

Before we begin discussing the formal and practical aspects of time-dependent density-functional theory, it is an essential prerequisite to review the basic concepts of ground-state density-functional theory (DFT). Today, DFT is widely accepted as a universal approach to electronic-structure calculations, and it is being used by tens of thousands of researchers worldwide, working in areas as diverse as drug design, metallurgy, nanotechnology, geology, and astrophysics, to name just a few. In this chapter we explain in a nutshell what DFT is, how it works, and why it is so popular.

There are several reasons why a solid grasp of the basics of DFT is necessary before one begins to study TDDFT:

- Many of the essential concepts and practical ingredients of TDDFT—such as the time-dependent Kohn–Sham equations, and approximations to the exchange–correlation potential—have their counterparts in static DFT and make use of similar ideas.
- A standard scenario in TDDFT is to begin with a system in its ground state, which is then acted upon by a time-dependent perturbation. The initial state needs to be calculated with DFT before one can start using TDDFT.
- There is a well-defined static limit in which TDDFT reduces to DFT, namely, when a system is initially in its ground state and there is no perturbation: the system is just sitting there and doing nothing. This case may sound trivial, but it is a very important check of the consistency of the time-dependent theory.

The purpose of the present chapter is to provide the "bare essentials" which are needed as part of the foundation upon which TDDFT is built. We will therefore discuss the basic formal framework of DFT, go over the various approximations to the exchange–correlation (xc) functionals that are in use today, and review some of the most important exact properties, practical aspects, and applications of the theory.

Needless to say, it is impossible to squeeze the vast amount of knowledge in the area of DFT accumulated over the years into a single introductory chapter. There exists a large body of literature on the subject of DFT; some of the most popular review articles and textbooks are listed in Appendix N. The interested reader is encouraged to consult these to gain a more comprehensive overview of the history and the current state of the art of this vibrant and diverse field of research.

2.1 The formal framework of DFT

2.1.1 The electronic many-body problem

DFT is a formally exact approach to the static electronic many-body problem. What do we mean by this? In this section, we define the electronic many-body problem to consist in finding the solutions of the static Schrödinger equation for a system of N interacting nonrelativistic electrons,

$$\hat{H}\Psi_j(\mathbf{x}_1,\ldots,\mathbf{x}_N) = E_j\Psi_j(\mathbf{x}_1,\ldots,\mathbf{x}_N)\,. \tag{2.1}$$

Here, the antisymmetric N-electron wave function $\Psi_j(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ is the jth eigenstate of the Hamiltonian $\hat{H}$, with associated energy eigenvalue E_j. We use $\mathbf{x}_j \equiv (\mathbf{r}_j,\sigma_j)$ as a shorthand notation for the space and spin coordinates of the jth electron. In the following, we shall not explicitly indicate the arguments $\mathbf{x}_1,\ldots,\mathbf{x}_N$ of an N-electron wave function, unless needed.

The total Hamiltonian of the N-electron system is given by

$$\hat{H} = \hat{T} + \hat{V} + \hat{W}\,, \tag{2.2}$$

where the kinetic-energy operator is

$$\hat{T} = \sum_{j=1}^{N} -\frac{\nabla_j^2}{2} \tag{2.3}$$

(∇_j denotes the gradient operator with respect to $\mathbf{r}_j$, the position vector of the jth electron), the potential operator is

$$\hat{V} = \sum_{j=1}^{N} v(\mathbf{r}_j)\,, \tag{2.4}$$

and the electron–electron interaction is in general

$$\hat{W} = \frac{1}{2}\sum_{\substack{j,k\\ j\neq k}}^{N} w(|\mathbf{r}_j-\mathbf{r}_k|)\,. \tag{2.5}$$

The usual choice is of course the Coulomb interaction $w(|\mathbf{r}_j-\mathbf{r}_k|) = 1/|\mathbf{r}_j-\mathbf{r}_k|$, but different forms of the interaction, including zero interaction, are also allowed.

Notice that eqn (2.1) is the Schrödinger equation for the electronic degrees of freedom only. A more general formulation of the structure of matter would also include the nuclear degrees of freedom on an equal footing; here, we treat all nuclei as fixed and producing a given Coulomb potential, which contributes to the total external potential $v(\mathbf{r})$. This is known as the Born–Oppenheimer approximation. In Chapter 17 we will give a precise definition of the Born–Oppenheimer approximation, and we will also discuss approaches to the coupling of electronic and nuclear dynamics.

If the system is in the jth many-body eigenstate, the associated value of a physical observable $\hat{O}$ is obtained from the expectation value $O_j = \langle \Psi_j | \hat{O} | \Psi_j \rangle$. For instance, the ground-state energy is given by

$$E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \,, \tag{2.6}$$

where Ψ_0 is the ground-state wave function. In other words, the wave function tells us all we could ever wish to know about the quantum state of the system and the outcomes of possible measurements described by the operator $\hat{O}$.[1] The situation was most concisely summarized by Dirac (1929) in the following famous quote:

> The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.

As of today, solving the full many-body Schrödinger equation (2.1) has remained an intractable numerical problem, except for special cases such as two-electron systems and systems with high symmetry, and in reduced dimensionality. Later, when we deal with the time-dependent case, we will often use such special cases as numerically exact benchmarks.

Over the years, many ingenious schemes have been devised to find approximate solutions of eqn (2.1). Among these methods are Hartree–Fock (HF) theory and configuration interaction (CI) expansions (see Appendix D), diagrammatic Green's function techniques, and quantum Monte Carlo approaches. In the remainder of this book, we shall often make reference to these schemes.

For now, however, let us ask a different and somewhat provocative question: is the search for the full many-body wave function Ψ of an N-electron system a reasonable goal? We give another quote, due to Kohn (1999):

> In general the many-electron wave function $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ for a system of N electrons is not a legitimate scientific concept, when $N \gtrapprox 10^3$.

How could that be? The explanation for this statement is that wave functions encounter an "exponential wall," in the sense that the number of parameters required to construct a (reasonably accurate) approximation to Ψ grows exponentially with the number of electrons and rapidly becomes overwhelming. Even for moderate particle numbers (on the order of 10), it becomes impossible[2] (or at least impractical) to even *store* such a many-body wave function!

To put it in simple words: the many-body wave function contains vastly more information than one would ever care to know about an N-electron system. In practice, we are interested in properties of the system that can be characterized by a single number such as its energy, dipole moment, or polarizability, or by functions of a single or a few variables such as the single-particle probability density, the pair density, or few-body correlation functions (see Appendix C for definitions). Calculating the full Ψ to obtain these quantities seems like a terribly wasteful approach if N is large. The

[1] This statement is true if the system is in a pure quantum state; otherwise, one needs to work with the statistical density matrix.

[2] Some simple estimates are given in Exercise 2.1 at the end of this chapter.

essence of DFT is to relieve us of this duty: one can in principle obtain all properties of a many-body system exactly, *without* having to solve the many-body Schrödinger equation! In the following, we will deliver a proof of this bold statement.

2.1.2 The Hohenberg–Kohn theorem

The origin of DFT dates back to the publication of a landmark paper by Hohenberg and Kohn (1964), in which they provided the basic existence proof that became later known as the Hohenberg–Kohn theorem.

We consider a system of N interacting electrons in a finite region of space, governed by the many-body Hamiltonian (2.2). The single-particle probability density of the electronic ground state is given by

$$n_0(\mathbf{r}) = N \sum_{\sigma} \int d\mathbf{x}_2 \ldots \int d\mathbf{x}_N |\Psi_0(\mathbf{r}, \sigma, \mathbf{x}_2, \ldots, \mathbf{x}_N)|^2 \,, \tag{2.7}$$

where we use the shorthand notation $\int d\mathbf{x}_l = \sum_{\sigma_l} \int d^3 r_l$ to denote integration over the lth spatial coordinate and summation over the lth spin index. We consider here the original, basic version of the Hohenberg–Kohn theorem, which assumes that the ground state of the system is not degenerate.[3]

Since $n_0(\mathbf{r})$ is obtained from the wave function, which in turn obeys Schrödinger's equation, it is immediately clear and obvious that the ground-state density is a *functional* of the external potential $v(\mathbf{r})$, i.e., we can write it as $n_0[v](\mathbf{r})$ (in Appendix B, we explain what functionals are, and how to compute their derivatives).

It turns out that the converse is also true, which forms the content of the Hohenberg–Kohn theorem:

> **Hohenberg–Kohn theorem.** In a finite, interacting N-electron system with a given particle–particle interaction there exists a one-to-one correspondence between the external potential $v(\mathbf{r})$ and the ground-state density $n_0(\mathbf{r})$. In other words, the external potential is a unique functional of the ground-state density, $v[n_0](\mathbf{r})$, up to an arbitrary additive constant.

The proof turns out to be surprisingly simple, and proceeds in two steps via *reductio ad absurdum* (i.e., we assume the opposite of what we want to prove, and show that this leads to a logical contradiction). In the following, two potentials $v(\mathbf{r})$ and $v'(\mathbf{r})$ are considered different if they are not related by a constant shift: $v'(\mathbf{r}) \neq v(\mathbf{r}) + c$.

In the first step of the proof, we note that two different potentials cannot produce the same ground-state wave function, i.e., Ψ_0 and Ψ_0' must differ by more than a trivial phase factor e^{-ic}. To convince yourself of this, assume that Ψ_0 and Ψ_0' are the same, and subtract the many-body Schrödinger equations. This leads to $\hat{V} - \hat{V}' = E_0 - E_0'$, in contradiction with the requirement that $v'(\mathbf{r}) \neq v(\mathbf{r}) + c$. The relationship between potentials and wave functions is therefore *unique*.[4]

[3] Extending the Hohenberg–Kohn theorem to degenerate ground states is straightforward (Dreizler and Gross, 1990).

[4] In generalizations of DFT to multicomponent systems (e.g., spin-DFT and current-DFT), a unique relationship between potentials and ground-state wave functions is no longer guaranteed, as

In the second step of the proof, we want to show that two different ground-state wave functions produce different ground-state densities, i.e.,

$$\begin{matrix}\Psi_0 \\ \\ \Psi_0'\end{matrix} \begin{matrix}\searrow \\ \\ \nearrow\end{matrix} \quad n_0(\mathbf{r}) \qquad \text{cannot happen,} \tag{2.8}$$

where $\Psi_0' \neq e^{-ic}\Psi_0$. To prove this, assume the contrary, namely, that both Ψ_0 and Ψ_0' (which come from Schrödinger equations with different potentials v and v') produce the same density $n_0(\mathbf{r})$. The ground-state energy associated with $v'(\mathbf{r})$ is given by

$$E_0' = \langle \Psi_0' | \hat{H}' | \Psi_0' \rangle \, . \tag{2.9}$$

Using the Rayleigh–Ritz variational principle and the fact that the two wave functions Ψ_0 and Ψ_0' are different, we have

$$E_0' < \langle \Psi_0 | \hat{H}' | \Psi_0 \rangle = \langle \Psi_0 | \hat{H} + \hat{V}' - \hat{V} | \Psi_0 \rangle = E_0 + \int d^3r [v'(\mathbf{r}) - v(\mathbf{r})] n_0(\mathbf{r}) \, . \tag{2.10}$$

It is essential that in this expression E_0' is strictly less than the right-hand side. Now we simply interchange primed and unprimed quantities:

$$E_0 < \langle \Psi_0' | \hat{H} | \Psi_0' \rangle = \langle \Psi_0' | \hat{H}' + \hat{V} - \hat{V}' | \Psi_0' \rangle = E_0' + \int d^3r [v(\mathbf{r}) - v'(\mathbf{r})] n_0(\mathbf{r}) \, . \tag{2.11}$$

The last equality on the right-hand side of eqns (2.10) and (2.11) reflects the assumption that Ψ_0 and Ψ_0' have the same density $n_0(\mathbf{r})$. Next, we simply add eqns (2.10) and (2.11), which yields

$$E_0 + E_0' < E_0 + E_0' \, , \tag{2.12}$$

an obvious contradiction. We have thus shown that Ψ_0 and Ψ_0' give different densities n_0 and n_0'; but in the first step we showed that Ψ_0 and Ψ_0' also come from different potentials v and v'. Therefore, a unique one-to-one correspondence exists between potentials and ground-state densities, which can be formally expressed by writing $v[n_0](\mathbf{r})$, and thus $\hat{V}[n_0]$.

Since the kinetic and electron-interaction operators $\hat{T}$ and $\hat{W}$ are fixed, this immediately implies that the Hamiltonian (2.2) is a functional of the ground-state density, $\hat{H}[n_0]$. Via the Schrödinger equation (2.1), this means that all eigenstates of the system become density functionals as well, $\Psi_j[n_0]$, and not just the ground state Ψ_0.

We thus arrive at the surprising conclusion that all ground- and excited-state properties of a given N-electron system are, in principle, determined entirely by its ground-state particle density! To appreciate the meaning of this statement, let us consider an example. Figure 2.1 shows a schematic illustration of the ground-state density of a diatomic molecule. At the positions of the nuclei, the density has a sharp cusp but is

shown by Capelle and Vignale (2001, 2002). Fortunately, this nonuniqueness is harmless in practice, since the second step of the Hohenberg–Kohn proof still goes through (Kohn *et al.*, 2004).

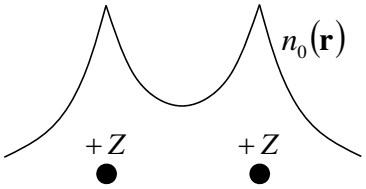

Fig. 2.1 In a molecule, the positions of the nuclei are indicated by cusps in the electronic ground-state density $n_0(\mathbf{r})$. The slope of the density at each cusp determines the nuclear charge (Kato, 1957).

otherwise smooth; this rigorous property of any many-body system subject to a superposition of Coulomb potentials is known as Kato's theorem (Kato, 1957). Specifically, this theorem says that

$$Z_i = -\frac{1}{2\bar{n}(\mathbf{r})} \left.\frac{\partial \bar{n}(\mathbf{r})}{\partial r}\right|_{\mathbf{r}=\mathbf{R}_i} , \tag{2.13}$$

where Z_i is the charge of the ith nucleus located at position $\mathbf{R}_i$, and $\bar{n}$ is the spherical average of the charge density. This is all we need:

- the positions of the cusps uniquely determine the locations of the nuclei;
- the magnitude and slope of the density at the cusps dictate the nuclear charges;
- the integral over the density gives the total number of electrons.

Thus, from the density given in Fig. 2.1, and similarly for all other molecular ground-state densities, it is straightforward to reconstruct the complete molecular Hamiltonian, which in turn determines all properties of the system. This example has been restricted to external potentials that come from a collection of positively charged nuclei, but the Hohenberg–Kohn theorem says that, in principle, *any* external potential can be uniquely reconstructed from the ground-state density.

In addition to establishing the one-to-one correspondence between potentials and ground-state densities, the original work by Hohenberg and Kohn (1964) contains two important corollaries related to the total-energy functional

$$E_{v_0}[n] = \langle \Psi[n] | \hat{T} + \hat{V}_0 + \hat{W} | \Psi[n] \rangle , \tag{2.14}$$

which is associated with a given external potential $v_0(\mathbf{r})$. Here, $n(\mathbf{r})$ is a density of an N-particle system, and $\Psi[n]$ is the unique ground-state wave function which produces this density. As a consequence of the Rayleigh–Ritz principle, $E_{v_0}[n]$ has the following variational property:

$$\begin{aligned} E_{v_0}[n] &> E_0 \qquad \text{for} \quad n(\mathbf{r}) \neq n_0(\mathbf{r}) , \\ E_{v_0}[n] &= E_0 \qquad \text{for} \quad n(\mathbf{r}) = n_0(\mathbf{r}) , \end{aligned} \tag{2.15}$$

where $n_0(\mathbf{r})$ is the ground-state density belonging to $v_0(\mathbf{r})$. This means that the exact ground-state density $n_0(\mathbf{r})$ of an interacting N-electron system can be found from the Euler equation

$$\frac{\delta}{\delta n(\mathbf{r})}\left[E_{v_0}[n] - \mu \int d^3r'\, n(\mathbf{r}')\right] = 0\,. \tag{2.16}$$

Here, μ is a Lagrange multiplier which ensures the correct total number of electrons. The significance of eqn (2.16) is that it allows one, in principle, to find the ground-state density of any system without solving the Schrödinger equation (2.1).

What can we say about the functional $E_{v_0}[n]$? Let us rewrite eqn (2.14) as

$$E_{v_0}[n] = F[n] + \int d^3r\, n(\mathbf{r}) v_0(\mathbf{r})\,. \tag{2.17}$$

The dependence of the total-energy functional on the external potential is very simple. The remaining part of the energy,

$$F[n] = \langle \Psi[n] | \hat{T} + \hat{W} | \Psi[n] \rangle = T[n] + W[n]\,, \tag{2.18}$$

is *universal* in the sense that it is the same for any N-electron system with the same electron–electron interaction, no matter what external potential is acting on it. Equation (2.16) then becomes

$$\frac{\delta F[n]}{\delta n(\mathbf{r})} + v_0(\mathbf{r}) = \mu\,. \tag{2.19}$$

The Hohenberg–Kohn theorem of DFT thus represents nothing less than a complete paradigm change of the electronic many-body problem: the wave function Ψ (a function of $3N$ variables) is replaced by the ground-state density n_0 (a function of three variables) as the fundamental quantity to be calculated. Of course, the catch is that the universal functional $F[n]$ is unknown, and much work has been carried out over the years to understand its properties and to find suitable approximations.

2.1.3 Constrained search

Aside from practical considerations, there are also some formal issues with the original Hohenberg–Kohn version of DFT. By construction, $E_{v_0}[n]$ is defined only for those functions $n(\mathbf{r})$ that are actual ground-state densities belonging to some external potential—such functions $n(\mathbf{r})$ are called v-representable. This is a prerequisite for the functional derivatives in eqns (2.16) and (2.19) to exist. The v-representability problem has remained elusive to this day, but at least it has been established that every density function on a lattice (finite or infinite) is associated with the ground eigenspace of some potential (Chayes *et al.*, 1985; Ullrich and Kohn, 2002).

To some extent, most of the difficulties related to the domain of the functional $E_{v_0}[n]$ have been overcome in an elegant way with the so-called *constrained search* formalism (Levy, 1979; Lieb, 1983). Other subtle issues such as the differentiability of the universal functional $F[n]$ have been reviewed by van Leeuwen (2003).

The key idea of the constrained-search formalism is quite simple. The Rayleigh–Ritz variational principle of quantum mechanics says that the ground-state energy E_0 which belongs to the Hamiltonian $\hat{H} = \hat{T} + \hat{V}_0 + \hat{W}$ for a given particle number N can be expressed mathematically as follows:

$$E_0 = \min_{\Psi} \langle \Psi | \hat{T} + \hat{V}_0 + \hat{W} | \Psi \rangle\,. \tag{2.20}$$

This means that we search over all antisymmetric N-particle wave functions until the lowest value of $\langle\Psi|\hat{H}|\Psi\rangle$ is found. But this search can also be carried out in two stages:

$$E_0 = \min_n \left\{ \min_{\Psi\to n} \langle\Psi|\hat{T} + \hat{V}_0 + \hat{W}|\Psi\rangle \right\} . \tag{2.21}$$

In other words, we first search over all Ψ's that produce a given density $n(\mathbf{r})$, and then determine that density which gives the overall lowest energy.[5] This density is then identified with the ground-state density $n_0(\mathbf{r})$.

It is straightforward to identify the object in parentheses as the functional

$$E_{v_0}[n] = \min_{\Psi\to n} \langle\Psi|\hat{T} + \hat{V}_0 + \hat{W}|\Psi\rangle , \tag{2.22}$$

where the minimization is constrained to those wave functions Ψ which yield the density n. This leads to the following definition of the universal functional $F[n]$:

$$F[n] = \min_{\Psi\to n} \langle\Psi|\hat{T} + \hat{W}|\Psi\rangle . \tag{2.23}$$

The advantage of the constrained-search formalism is that it gives an operational definition of the universal functional $F[n]$ in the form of a constructive procedure. This procedure, although not very practical (it would involve a search over an infinite number of wave functions), plays a very important formal and conceptual role in DFT.

The question of whether similar ideas can be applied in TDDFT has remained open to this day, following some early attempts (Kohl and Dreizler, 1986); we will therefore not elaborate further on the constrained-search formalism in this book.

2.1.4 The Kohn–Sham equations

Let us consider a noninteracting system, i.e., a system where the part $\hat{W}$ of the total Hamiltonian (2.2) is absent. We denote the Hamiltonian of this system by $\hat{H}_s$,

$$\hat{H}_s = \hat{T} + \hat{V}_s = \sum_{j=1}^{N} \left(-\frac{\nabla_j^2}{2} + v_s(\mathbf{r}_j) \right) . \tag{2.24}$$

The Hohenberg–Kohn theorem applies to this special case, too, and ensures a one-to-one correspondence between $v_s(\mathbf{r})$ and $n_{0s}(\mathbf{r})$, the ground-state density of the noninteracting system. We can write the associated total energy functional as

$$E_{v_s}[n] = T_s[n] + \int d^3r\, n(\mathbf{r}) v_s(\mathbf{r}). \tag{2.25}$$

In this case, the universal density functional $F[n]$ (which still depends on the interaction between the particles) reduces to the noninteracting kinetic-energy functional $T_s[n]$, and the Euler equation (2.16) becomes

[5] There is a nice analogy to doing the search in this way (Parr and Yang, 1989). Imagine you want to find the tallest student in a school. One way would be to have all students line up in the schoolyard and select the tallest one. This corresponds to eqn (2.20). The other method would be to ask the tallest student in each classroom to step outside, and then to pick the tallest student in the entire school from this smaller, preselected group. This corresponds to eqn (2.21).

$$\frac{\delta E_{v_s}[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + v_s(\mathbf{r}) = \mu \,. \tag{2.26}$$

Solving eqn (2.26) gives the exact ground-state density $n_{0s}(\mathbf{r})$ of the noninteracting system. Unfortunately, in practice $T_s[n]$ is only approximately known, and any solution of eqn (2.26) inevitably suffers from the deficiencies associated with any such approximation.[6]

There is, however, an easy way out of this dilemma. For noninteracting systems, the many-body ground-state wave function reduces to a single Slater determinant,[7]

$$\Psi_s(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \varphi_2(\mathbf{r}_1) & \ldots & \varphi_N(\mathbf{r}_1) \\ \varphi_1(\mathbf{r}_2) & \varphi_2(\mathbf{r}_2) & \ldots & \varphi_N(\mathbf{r}_2) \\ \vdots & \vdots & & \vdots \\ \varphi_1(\mathbf{r}_N) & \varphi_2(\mathbf{r}_N) & \ldots & \varphi_N(\mathbf{r}_N) \end{vmatrix}, \tag{2.27}$$

where the single-particle orbitals $\varphi_j(\mathbf{r})$ satisfy the Schrödinger equation

$$\left(-\frac{\nabla^2}{2} + v_s(\mathbf{r})\right) \varphi_j(\mathbf{r}) = \varepsilon_j \varphi_j(\mathbf{r}) \,, \tag{2.28}$$

and the ground-state density is obtained from the N lowest single-particle orbitals as

$$n_s(\mathbf{r}) = \sum_{j=1}^{N} |\varphi_j(\mathbf{r})|^2 \,. \tag{2.29}$$

Equation (2.25) and eqns (2.28) and (2.29) are formally equivalent and exact ways of obtaining the ground-state density of a noninteracting system. But the second method, which utilizes the single-particle orbitals, is much more convenient and accurate, and is therefore preferred for practical applications.

It was the key insight of Kohn and Sham (1965) that one could take advantage of an effective single-particle picture to transform DFT into the practical scheme that is nowadays the basis of most applications of DFT. All one needs to do is to rewrite the total-energy functional (2.17) in a clever way by addition and subtraction:

$$\begin{aligned} E_{v_0}[n] &= T[n] + W[n] + \int d^3r \, n(\mathbf{r}) v_0(\mathbf{r}) \\ &= T_s[n] + \int d^3r \, n(\mathbf{r}) v_0(\mathbf{r}) + \frac{1}{2} \int d^3r \int d^3r' \, \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{\rm xc}[n] \,, \end{aligned} \tag{2.30}$$

[6] The simplest approximation to $T_s[n]$ is the Thomas–Fermi expression $T_s^{\rm TF}[n] = \int d^3r \tau_s^h(\mathbf{r})$, which is based on the kinetic energy per unit volume of the noninteracting homogeneous electron liquid, $\tau_s^h(n) = (3/10)(3\pi^2)^{2/3} n^{5/3}$, evaluated at the local density $n(\mathbf{r})$. It turns out that the Thomas–Fermi model does not reproduce atomic shell structures, and does not lead to molecular binding. For a review of the Thomas–Fermi and related models in DFT, see Dreizler and Gross (1990).

[7] More generally, in the presence of degeneracies, the noninteracting ground state can be a linear combination of Slater determinants; in some situations, an ensemble of determinants is required. However, most real ground-state densities can be represented as noninteracting systems with a single Slater determinant, and it is desirable to do so for practical reasons, because common approximations are based on single-determinant descriptions. The situation can become quite subtle if the real system has certain symmetry requirements associated with a multiplet. In that case, a properly symmetrized Kohn–Sham theory can be constructed via constrained search (Görling, 1993).

where the classical Coulomb energy (or Hartree energy) is given by

$$E_{\rm H}[n] = \frac{1}{2}\int d^3r \int d^3r' \, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \,, \tag{2.31}$$

and the remaining part is called the exchange–correlation energy:

$$E_{\rm xc}[n] = T[n] - T_s[n] + W[n] - \frac{1}{2}\int d^3r \int d^3r' \, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \,. \tag{2.32}$$

Plugging the total-energy functional (2.30) into the Euler equation (2.16), we obtain

$$\frac{\delta T_s[n]}{\delta n(\mathbf{r})} + v(\mathbf{r}) + \int d^3r' \, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})} = \mu \,. \tag{2.33}$$

Now compare this with the noninteracting Euler equation (2.26). It is immediately clear that the two equations are formally identical if we identify

$$v_s[n](\mathbf{r}) = v(\mathbf{r}) + \int d^3r' \, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + v_{\rm xc}[n](\mathbf{r}) \,, \tag{2.34}$$

where the xc potential is defined as

$$v_{\rm xc}[n](\mathbf{r}) = \frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})} \,. \tag{2.35}$$

This means that the ground-state density of the *interacting* system can be found by solving the single-particle Schrödinger equation

$$\left(-\frac{\nabla^2}{2} + v_s[n](\mathbf{r})\right)\varphi_j(\mathbf{r}) = \varepsilon_j \varphi_j(\mathbf{r}) \tag{2.36}$$

and summing the squares of the lowest N occupied orbitals,

$$n_0(\mathbf{r}) = \sum_{j=1}^{N} |\varphi_j(\mathbf{r})|^2 \,. \tag{2.37}$$

Equations (2.34)–(2.37) are called the *Kohn–Sham equations*. We see that they emerge as a consequence of rewriting the Hohenberg–Kohn variational principle in a clever way. Via eqn (2.36), the Kohn–Sham orbitals are functionals of the density, $\varphi_j[n](\mathbf{r})$.

Let us now consider the noninteracting kinetic-energy functional $T_s[n]$, and assume that we have solved the Kohn–Sham equations. Then

$$T_s[n_0] = \sum_{j=1}^{N} \int d^3r \, \varphi_j^*(\mathbf{r}) \left(-\frac{\nabla^2}{2}\right) \varphi_j(\mathbf{r}) \tag{2.38}$$

$$= \sum_{j=1}^{N} \varepsilon_j - \int d^3r \, n_0(\mathbf{r}) v_s[n_0](\mathbf{r}) \,. \tag{2.39}$$

According to eqn (2.38), T_s is an *explicit* functional of the Kohn–Sham orbitals, but an *implicit* functional of the density: $T_s[n] = T_s\big[\{\varphi_j[n]\}\big]$. We use here the notation

Table 2.1 Total ground-state energy E_{v_0} (in a.u.) of rare-gas atoms, calculated with the LDA, and various components of the total energy according to eqn (2.30). We define $E_{\rm ext} = \int d^3r\, n_0(\mathbf{r}) v_0(\mathbf{r})$.

Atom	E_{v_0}	T_s	$E_{\rm ext}$	$E_{\rm H}$	$E_{\rm xc}$
He	−2.83	2.77	−6.63	2.00	−0.97
Ne	−128.23	127.74	−309.99	65.73	−11.71
Ar	−525.95	524.97	−1253.13	231.46	−29.24
Kr	−2750.15	2747.81	−6577.87	1171.72	−91.82
Xe	−7228.86	7225.10	−17159.16	2880.92	−175.71

$[\{\varphi_j\}]$ to indicate functional dependence on the set of occupied orbitals. Later, we will also encounter approximate xc functionals which depend explicitly on the orbitals but only implicitly on the density.

Plugging eqn (2.39) into eqn (2.30) gives an alternative, convenient expression for the exact ground-state energy of the interacting system:

$$E_{v_0}[n_0] = \sum_{j=1}^{N} \varepsilon_j - \frac{1}{2}\int d^3r \int d^3r'\, \frac{n_0(\mathbf{r}) n_0(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \int d^3r\, n_0(\mathbf{r}) v_{\rm xc}[n_0](\mathbf{r}) + E_{\rm xc}[n_0]\,. \tag{2.40}$$

Table 2.1 shows total ground-state energies of rare-gas atoms, obtained from a Kohn–Sham calculation using the local-density approximation (LDA) for $E_{\rm xc}$ (we will explain the LDA and other approximations in Section 2.3). Also shown are the individual contributions to E_{v_0} [see eqn (2.30)], where T_s and $E_{\rm ext} = \int d^3r\, n_0(\mathbf{r}) v_0(\mathbf{r})$ are clearly dominant. The main advantage of the Kohn–Sham approach to DFT now becomes clear: as seen in eqn (2.38), the noninteracting kinetic-energy functional $T_s[n]$ is treated *exactly* in the Kohn–Sham scheme; this is made possible by introducing the effective single-particle orbitals.

But there's more! We have not yet commented on those parts of the total energy that result from electron–electron interaction. Again, the Kohn–Sham theory manages to treat the largest portion of that energy exactly, namely the Hartree energy $E_{\rm H}$. The remaining part, $E_{\rm xc}$, is the only unknown in DFT, but, as seen from Table 2.1, it is also by far the smallest part of the total energy. It turns out, however, that $E_{\rm xc}$ contributes significantly (up to 100%) to the binding energy of matter.[8] Finding good approximations to $E_{\rm xc}$ is therefore extremely important.

In practice, the Kohn–Sham equations are almost always implemented in their spin-resolved form. For an N-system with $N = N_\uparrow + N_\downarrow$, the total ground-state density becomes a sum of spin-up and spin-down densities,

[8] A drastic example is given by van der Waals interactions (see Chapter 14), which are caused purely by correlation effects.

$$n_0(\mathbf{r}) = n_{0\uparrow}(\mathbf{r}) + n_{0\downarrow}(\mathbf{r}) = \sum_{\sigma=\uparrow,\downarrow} \sum_{j=1}^{N_\sigma} |\varphi_{j\sigma}(\mathbf{r})|^2 \,, \tag{2.41}$$

and the Kohn–Sham spin orbitals follow from

$$\left(-\frac{\nabla^2}{2} + v_{s\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) \right) \varphi_{j\sigma}(\mathbf{r}) = \varepsilon_{j\sigma} \varphi_{j\sigma}(\mathbf{r}) \,. \tag{2.42}$$

The Kohn–Sham effective potential is given by

$$v_{s\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) = v_\sigma(\mathbf{r}) + \int d^3r' \, \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + v_{\mathrm{xc}\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) \,, \tag{2.43}$$

where the xc energy and potential are defined as functionals of the individual spin densities:

$$v_{\mathrm{xc}\sigma}[n_\uparrow, n_\downarrow](\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]}{\delta n_\sigma(\mathbf{r})} \,. \tag{2.44}$$

The spin-resolved Kohn–Sham formalism (2.41)–(2.44) is more general than the spin-independent version (2.34)–(2.37) since it includes spin-dependent external potentials $v_\sigma(\mathbf{r})$. In most applications, however, the external potential is simply the Coulomb potential of the nuclei, and under these circumstances the spin-independent and the spin-dependent Kohn–Sham formalism are in principle equivalent. In practice, the functional dependence of E_{xc} on the spin-up and spin-down densities provides more flexibility and is better suited for the construction of approximations.

2.2 Exact properties

The key quantities in Kohn–Sham DFT are the xc energy functional E_{xc} and the resulting xc potential v_{xc}; neither of them is exactly known. Approximations are therefore unavoidable; these should be as accurate as possible yet sufficiently simple to be of practical use.

On occasion, DFT has been portrayed critically as a theory which is in principle exact but in practice lacks strategies that allow one to construct approximations which can be systematically improved. This is put in contrast with wave-function-based schemes from quantum chemistry such as configuration interaction methods. Such criticism misses the point, for two reasons.

First of all, one of the main advantages of DFT is precisely that it does *not* necessarily follow a systematic strategy for treating xc many-body effects: this allows clever shortcuts which are capable of delivering unexpected degrees of accuracy with relatively low computational effort. This would be much more expensive to obtain from more traditional wave-function-based methods.

But this does not mean that DFT is unsystematic. On the contrary, xc functionals *can* be systematically constructed on a formal level by using many-body perturbation theory and proceeding order by order (however, with increasing difficulty). We will discuss such approaches for the time-dependent case in Chapter 13.

In practice, many successful strategies for constructing approximate xc functionals are based on the idea that, in general, satisfying more and more exact properties and

constraints will lead to better functionals. In this section, we will summarize some of the known exact properties of xc functionals.

2.2.1 Orbitals, eigenvalues, and asymptotics

In the Kohn–Sham scheme, the exact ground-state density (2.41) is constructed from the N lowest occupied single-particle orbitals following the conventional Aufbau principle. It is important to stress that the Kohn–Sham orbitals $\varphi_j(\mathbf{r})$ themselves, strictly speaking, serve no purpose other than to produce the ground-state density; in particular, the Kohn–Sham ground-state Slater determinant is not meant to reproduce the fully interacting ground-state wave function:

$$\Psi_{s0}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \neq \Psi_0(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)\,. \tag{2.45}$$

Although they both have the same one-particle density, Ψ_{s0} and Ψ_0 are in general very different functions, and other quantities derived from the wave functions (for instance, one- or two-particle density matrices) can not be expected to be close. This plays a particularly important role in the calculation of certain observables which are hard to express as explicit functionals of the density, but easy to write down as functionals of the Kohn–Sham orbitals. We will discuss such observables in detail for the time-dependent case in Chapter 5.

In general, the Kohn–Sham energy eigenvalues ε_j do not have a rigorous physical meaning either. However, this statement calls for a bit more discussion. First of all, there is an important exception which concerns the highest occupied eigenvalue. For a finite N-electron system, ε_N equals minus the ionization energy of the system[9] (the energy it costs to remove an electron to infinity) (Almbladh and von Barth, 1985):

$$\varepsilon_N(N) = E(N) - E(N-1) \equiv -I(N)\,, \tag{2.46}$$

where $E(N)$ and $E(N-1)$ are the total ground-state energies of the N- and $(N-1)$-electron systems, respectively. This can be regarded as the DFT version of the Koopmans theorem.[10] Equation (2.46) assumes that the Kohn–Sham effective potential vanishes at infinity. In a similar way, the highest Kohn–Sham eigenvalue of the $(N+1)$-particle system is related to the electron affinity (the energy gained by bringing in a particle from infinity):

$$\varepsilon_{N+1}(N+1) = E(N+1) - E(N) \equiv -A(N)\,. \tag{2.47}$$

It is tempting to identify differences of Kohn–Sham energy eigenvalues, $\varepsilon_a - \varepsilon_i$, with excitation energies of the physical system (here, a and i are labels for unoccupied and occupied orbitals, respectively). We will investigate this point in detail later, in

[9] In infinite metallic systems, the highest occupied Kohn–Sham eigenvalue equals the Fermi energy (Giuliani and Vignale, 2005). However, the shape of the Kohn–Sham Fermi surface differs in general from the true Fermi surface (Mearns, 1988).

[10] The original Koopmans theorem (Koopmans, 1934) states that the HF orbital energies correspond to electron removal energies; this ignores correlation effects, as well as effects due to possible reorganization of the remaining electrons (see also Appendix D). The DFT–Koopmans theorem, by contrast, is exact, but only holds for the highest-occupied-orbital eigenvalue.

Section 9.1. For now, we merely state that the Kohn–Sham excitation energies $\varepsilon_a - \varepsilon_i$ in general differ from the exact excitation energies of the interacting many-body system; however, they can often be taken as reasonable first approximations. The quality of the Kohn–Sham excitation energies depends crucially on the quality of the approximate xc potential, in particular when it comes to describing unoccupied levels.

Let us now consider the asymptotic behavior of the Kohn–Sham effective potential of a finite, overall neutral system composed of N electrons and N positive charges (for an atom, this is the nuclear charge; for a molecule, the sum of all nuclear charges). For $r \to \infty$, the external and Hartree potentials behave as follows:

$$v(\mathbf{r}) \longrightarrow -\frac{N}{r}\,, \qquad v_{\rm H}(\mathbf{r}) = \int d^3r'\,\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \longrightarrow \frac{N}{r}\,. \tag{2.48}$$

If an electron ventures far away into the outer regions of the system, leaving $N-1$ electrons behind, it should see the Coulomb potential of the remaining positive ion, and it is the job of the xc potential to take care of this. The xc potential of a finite system therefore has the following asymptotic behavior:[11]

$$v_{\rm xc}(\mathbf{r}) \longrightarrow -\frac{1}{r}\,, \qquad r \to \infty\,. \tag{2.49}$$

Similar results can be derived for the behavior of the xc potential outside the surface of an extended solid, metallic or insulating, where the asymptotic behavior is given by the classical image potential (Almbladh and von Barth, 1985).

2.2.2 Self-interaction

The electron–electron interaction part $\hat{W}$ of the many-body Hamiltonian (2.2) excludes all terms with $j = k$, which accounts for the fact that an electron interacts with all other electrons but not with itself. It is therefore an exact requirement of the Kohn–Sham formalism that it must be free of such self-interaction. This implies the following condition:

$$E_{\rm H}[n_{j\sigma}] + E_{\rm xc}[n_{j\sigma}, 0] = 0\,, \tag{2.50}$$

where $n_{j\sigma}(\mathbf{r}) = |\varphi_{j\sigma}(\mathbf{r})|^2$ is the density of the Kohn–Sham spin orbital $\varphi_{j\sigma}(\mathbf{r})$. In other words, the xc energy of each occupied Kohn–Sham orbital must fully cancel its self-Hartree (direct Coulomb) energy.

The exact exchange energy in DFT is defined as

$$E_{\rm x}^{\rm exact} = -\frac{1}{2}\sum_\sigma \sum_{i,j=1}^{N_\sigma} \int d^3r \int d^3r'\,\frac{\varphi^*_{i\sigma}(\mathbf{r}')\varphi_{j\sigma}(\mathbf{r}')\varphi_{i\sigma}(\mathbf{r})\varphi^*_{j\sigma}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\,. \tag{2.51}$$

This is the same expression as for the exchange energy in HF theory, but evaluated with Kohn–Sham orbitals (coming from a local potential) rather than HF orbitals (coming from a nonlocal potential). We will discuss these subtleties at greater length

[11] The asymptotic behavior of the xc potential is dominated by the exchange part, $v_{\rm x}(\mathbf{r}) \to -1/r$. The correlation potential drops to zero much more quickly: $v_{\rm c}(\mathbf{r}) \to -\mathbf{r}\,\alpha\,\mathbf{r}/2r^6$, to within an arbitrary constant, where α is the polarizability tensor of the $(N-1)$-particle system.

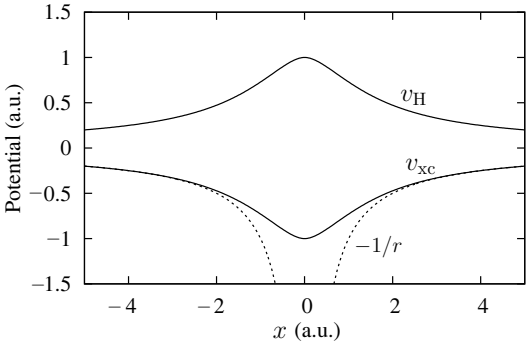

Fig. 2.2 Exact Hartree and xc potentials for the hydrogen atom ($N = 1$).

in Chapter 11 when we talk about orbital functionals and the (TD)OEP method. From the exchange energy functional (2.51), it follows immediately that

$$E_{\rm H}[n_{j\sigma}] + E_{\rm x}[n_{j\sigma}, 0] = 0\,, \tag{2.52}$$

i.e., the self-Hartree energy is compensated by the self-exchange energy for each orbital. The self-correlation energy, on the other hand, vanishes by itself:

$$E_{\rm c}[n_{j\sigma}, 0] = 0\,. \tag{2.53}$$

The freedom from self-interaction in $E_{\rm xc}$ automatically guarantees that the xc potential satisfies the asymptotic behavior (2.49), because in this way it can cancel the asymptotic self-Hartree potential. We will soon see that many of the most widely used approximate xc functionals do not satisfy this condition, and we will show in Section 2.3.4 how it can be approximately restored using so-called self-interaction corrections.

The simplest case, which exhibits the self-interaction condition in a particularly impressive manner, is for systems with $N = 1$ electron, such as the hydrogen atom. The Kohn–Sham equation for the H atom reads

$$\left(-\frac{\nabla^2}{2} - \frac{1}{r}\right)\varphi_j(\mathbf{r}) = \varepsilon_j\varphi_j(\mathbf{r})\,, \tag{2.54}$$

which immediately shows that $v_{\rm xc}(\mathbf{r}) = -v_{\rm H}(\mathbf{r}) = -\int d^3r\, n(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'|$. The xc potential (more precisely, the exchange potential—there is no correlation in a one-electron system) has the job of compensating the self-Hartree potential of the single electron. This is illustrated in Fig. 2.2.

2.2.3 The band gap in solids and derivative discontinuities

Different types and states of matter can often be characterized by the presence or absence of *gaps*. In general, a gap means that there is a finite energetic difference between two states of a system. There exist various types of gaps in physics, such as the (quasi)particle gap, the optical gap, and the superconducting gap. These gaps are distinguished by the degrees of freedom that are involved in changing the state of the

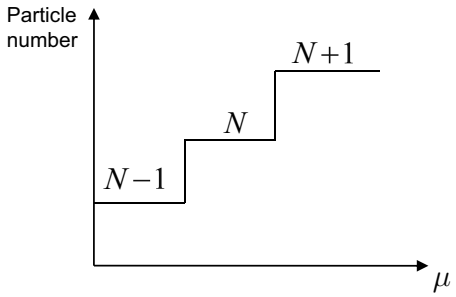

Fig. 2.3 For a system with a gap coupled to a particle reservoir, the chemical potential μ is a piecewise constant function of the total particle number. When the particle number passes through an integer, μ jumps discontinuously.

system. There are different experimental probes for each type of gap. For instance, the excitation gap (also known as the optical gap) can be probed by means of optical spectroscopy, and it refers to the energy difference between the electronic ground state and the first excited state of a system for a *fixed* particle number N. As we shall see later, optical gaps can be successfully calculated with TDDFT.

The fundamental band gap in an insulating periodic solid is an example of a particle gap,[12] involving ground-state energies of systems with *different* particle numbers. Experimentally, particle gaps play an important role in transport phenomena and in electron-transfer reactions. The fundamental particle gap E_g is defined as

$$E_g(N) = I(N) - A(N), \tag{2.55}$$

where the ionization energy $I(N)$ and electron affinity $A(N)$ are defined in eqns (2.46) and (2.47). From this, the fundamental particle gap can be expressed as a difference of Kohn–Sham eigenvalues,

$$E_g(N) = \varepsilon_{N+1}(N+1) - \varepsilon_N(N)\,. \tag{2.56}$$

It is important to note that these are the highest occupied Kohn–Sham energy eigenvalues of two different systems, namely with N and with $N+1$ electrons. In a macroscopic solid with 10^{23} electrons, it is clearly practically impossible to calculate the band gap by following this definition.

Let us now define the particle gap in a noninteracting system, also known as the Kohn–Sham gap:

$$E_{g,s}(N) = \varepsilon_{N+1}(N) - \varepsilon_N(N)\,. \tag{2.57}$$

In contrast with the interacting gap E_g, the Kohn–Sham gap $E_{g,s}$ is simply the difference between the highest occupied and lowest unoccupied single-particle levels in the *same* N-particle system. We can relate the two gaps by

$$E_g = E_{g,s} + \Delta_{\rm xc}, \tag{2.58}$$

[12]The discussion of the particle gap in this section mainly follows Capelle *et al.* (2010), who also discuss the spin gap in terms of derivative discontinuities and its connection to TDDFT.

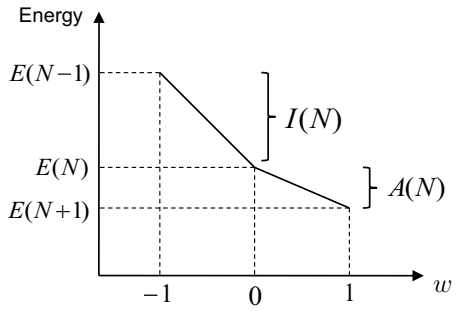

Fig. 2.4 The total energy $E(N+w)$ is a set of straight lines as a function of the particle number $N+w$. At integer values of N, there is a derivative discontinuity.

which defines $\Delta_{\rm xc}$ as the many-body correction to the Kohn–Sham gap. By making use of the previous relations, we can cast $\Delta_{\rm xc}$ as

$$\Delta_{\rm xc} = \varepsilon_{N+1}(N+1) - \varepsilon_{N+1}(N)\,. \tag{2.59}$$

It turns out that the many-body gap correction $\Delta_{\rm xc}$ can be related to a very fundamental property of density functionals, known as derivative discontinuities, as we will now explain.

Let us consider an open system with particle number N, connected to a particle reservoir with fixed chemical potential μ. Assume that our system has a gap, and that μ lies between the highest occupied and the lowest unoccupied energy level of the system. But the position of these levels relative to μ is not fixed: we can rigidly shift the level spectrum by changing the adjustable constant in the potential.

As long as the potential changes are small enough that μ remains in the gap, nothing interesting happens. But as soon as, say, the lowest unoccupied level moves below μ, an additional particle will be allowed to come in and the total number changes abruptly from N to $N+1$. To reflect the new situation, the chemical potential will take on a new value, compatible with $N+1$. Thus, μ is a piecewise constant function of the particle number, as illustrated in Fig. 2.3.

Let us now consider the Euler equations (2.19) and (2.26) for the interacting and the noninteracting system. The key observation is that the right-hand sides of both equations contain μ and therefore change discontinuously with particle number—but then, so must the left-hand sides! This means that the functional derivatives of $F[n]$ and $T_s[n]$ change discontinuously for variations $\delta n(\mathbf{r})$ such that N passes through an integer, and are not precisely defined at that integer.

In this picture, the particle number is treated as a continuous variable which is allowed to take nonintegral values. Formally, a particle number $N+w$, where $0<w<1$, can be achieved with an ensemble which mixes systems with N and $(N+1)$ particles via a density operator $\hat\rho = (1-w)|\Psi_0^N\rangle\langle\Psi_0^N| + w|\Psi_0^{N+1}\rangle\langle\Psi_0^{N+1}|$ (Perdew *et al.*, 1982). An important outcome of this ensemble formalism is that the total ground-state energy is given by $E(N+w) = (1-w)E(N) + wE(N+1)$; in other words, the energy is a set of straight lines connecting values at integer particle numbers (see Fig. 2.4).

The ionization potential and the electron affinity can thus be related to the derivatives of the energy to the left and to the right of the discontinuity at N:

$$-A(N) = E(N+1) - E(N) = \left.\frac{\partial E}{\partial N}\right|_{N^+} = \left.\frac{\delta E}{\delta n(\mathbf{r})}\right|_{N^+} \tag{2.60}$$

and

$$-I(N) = E(N) - E(N-1) = \left.\frac{\partial E}{\partial N}\right|_{N^-} = \left.\frac{\delta E}{\delta n(\mathbf{r})}\right|_{N^-}, \tag{2.61}$$

where $N^\pm = \lim_{\delta\to 0} N \pm \delta$. The last equalities in these equations arise from the definition of the chemical potential, $\mu = \partial E/\partial N$, and the Euler equation of DFT, eqn (2.19). The fundamental gap is obtained from eqn (2.55) as the derivative discontinuity of the energy at integer N:

$$E_g(N) = \left.\frac{\delta E[n]}{\delta n(\mathbf{r})}\right|_{N^+} - \left.\frac{\delta E[n]}{\delta n(\mathbf{r})}\right|_{N^-}. \tag{2.62}$$

In the total-energy expression (2.30), only the noninteracting kinetic energy and the xc energy are discontinuous, so

$$E_g(N) = \left.\frac{\delta T_s[n]}{\delta n(\mathbf{r})}\right|_{N^+} + \left.\frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}\right|_{N^+} - \left.\frac{\delta T_s[n]}{\delta n(\mathbf{r})}\right|_{N^-} - \left.\frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}\right|_{N^-}. \tag{2.63}$$

The entire argument up to this point can be repeated for a noninteracting system, and one obtains

$$E_{g,s} = \left.\frac{\delta T_s[n]}{\delta n(\mathbf{r})}\right|_{N^+} - \left.\frac{\delta T_s[n]}{\delta n(\mathbf{r})}\right|_{N^-}. \tag{2.64}$$

Comparison with eqn (2.58) gives

$$\Delta_{\rm xc} = \left.\frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}\right|_{N^+} - \left.\frac{\delta E_{\rm xc}[n]}{\delta n(\mathbf{r})}\right|_{N^-}. \tag{2.65}$$

This identifies the many-body correction to the Kohn–Sham gap as the derivative discontinuity of the xc energy arising from the nonuniqueness of the potentials with respect to an additive constant.

Besides giving a contribution to the band gap in solids, the discontinuity $\Delta_{\rm xc}$ of the Kohn–Sham potential also plays an important role in many other physical situations. For instance, it ensures that heteroatomic molecules dissociate into neutral fragments: approximate xc potentials (such as the LDA and GGA—see below) which do not reproduce the discontinuity upon change of particle number lead to unphysical fractional charges in the end products of molecular dissociation (Perdew, 1985; Ruzsinszky *et al.*, 2006).

As we will see later, discontinuities of the xc potential also play a role in TDDFT, for instance in field-induced ionization processes and in transport through nanoscale junctions.

2.2.4 Uniform limit

The homogeneous electron liquid[13] is an extremely important model system, which has been widely studied in theoretical physics [for a comprehensive overview see Pines and Nozières (1966) and Giuliani and Vignale (2005)]. The system is uniformly extended and consists of an infinite number of interacting electrons characterized by a single parameter, the uniform particle density n. For later reference, let us list a few relevant quantities in two and three dimensions (2D and 3D): the Wigner–Seitz radius

$$r_s = \left(\frac{3}{4\pi n}\right)^{1/3} \qquad \text{(3D)}, \tag{2.66}$$

$$r_s = \left(\frac{1}{\pi n}\right)^{1/2} \qquad \text{(2D)}, \tag{2.67}$$

the Fermi wave vector

$$k_F = \left(3\pi^2 n\right)^{1/3} \qquad \text{(3D)} \tag{2.68}$$

$$k_F = \left(2\pi n\right)^{1/2} \qquad \text{(2D)}, \tag{2.69}$$

and the Fermi energy $E_F = k_F^2/2$ in both 2D and 3D.

An extension of this model is the spin-polarized homogeneous electron liquid, characterized by two parameters, the spin-up and spin-down particle densities $n_\uparrow$ and $n_\downarrow$. Alternatively, one often uses the total density n and the (dimensionless) spin polarization ζ, defined as

$$n = n_\uparrow + n_\downarrow\,, \qquad \zeta = \frac{n_\uparrow - n_\downarrow}{n}\,. \tag{2.70}$$

The quantity of interest is the xc energy density $e^h_{\rm xc}(n)$ [or $e^h_{\rm xc}(n_\uparrow, n_\downarrow)$ in the spin-polarized case] of the homogeneous electron liquid. Instead of the xc energy per unit volume, $e^h_{\rm xc}(n)$, one can also work with the energy per particle, $\widetilde{e}^h_{\rm xc}(n)$, defined as $e^h_{\rm xc} = n\widetilde{e}^h_{\rm xc}$. Since the density here is just a number, $e^h_{\rm xc}(n)$ and $\widetilde{e}^h_{\rm xc}(n)$ are simply *functions* of n.

The xc energy density can be separated into an exchange and a correlation part, $e^h_{\rm xc} = e^h_{\rm x} + e^h_{\rm c}$. The exchange part is known exactly and has the form

$$e^h_{\rm x}(n,\zeta) = e^h_{\rm x}(n,0) + [e^h_{\rm x}(n,1) - e^h_{\rm x}(n,0)]f(\zeta)\,, \tag{2.71}$$

with the exchange energy densities of the unpolarized and the fully polarized electron liquids given by

$$e^h_{\rm x}(n,0) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} n^{4/3}\,, \quad e^h_{\rm x}(n,1) = -\frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3} n^{4/3} \qquad \text{(3D)}\,, \tag{2.72}$$

$$e^h_{\rm x}(n,0) = -\frac{4}{3}\left(\frac{2}{\pi}\right)^{1/2} n^{3/2}\,, \quad e^h_{\rm x}(n,1) = -\frac{8}{3}\left(\frac{1}{\pi}\right)^{1/2} n^{3/2} \qquad \text{(2D)}\,, \tag{2.73}$$

[13] Whether one speaks of the homogeneous electron *liquid* or the homogeneous electron *gas* is mostly a matter of taste; both are really the same thing. We prefer to use "liquid" because later, in the dynamical case, we will frequently invoke concepts of classical hydrodynamics.

Table 2.2 Fitting parameters for the 3D correlation energy per particle, eqns (2.75) and (2.76) (Perdew and Wang, 1992).

	$e_{\rm c}(r_s,0)$	$e_{\rm c}(r_s,1)$	$-\alpha_{\rm c}(r_s)$
A	0.031091	0.015545	0.016887
α_1	0.21370	0.20548	0.11125
β_1	7.5957	14.1189	10.357
β_2	3.5876	6.1977	3.6231
β_3	1.6382	3.3662	0.88026
β_4	0.49294	0.62517	0.49671

and the interpolation function for dimension $d = 2$ and 3,

$$f(\zeta) = \frac{(1+\zeta)^{\frac{d+1}{d}} + (1-\zeta)^{\frac{d+1}{d}} - 2}{2(2^{1/d}-1)} \,. \tag{2.74}$$

The correlation energy density, on the other hand, is not exactly known in analytic form, but very accurate numerical results exist thanks to quantum Monte Carlo calculations (Ceperley and Alder, 1980; Tanatar and Ceperley, 1989). Based on the available numerical data, high-precision analytical parametrizations have been developed in 3D (Vosko *et al.*, 1980; Perdew and Wang, 1992) and 2D (Tanatar and Ceperley, 1989; Attaccalite *et al.*, 2002).

The Perdew–Wang interpolation formula in 3D is given by

$$\widetilde{e}^h_{\rm c}(r_s,\zeta) = \widetilde{e}^h_{\rm c}(r_s,0) + \alpha_{\rm c}(r_s)\frac{f(\zeta)}{f''(0)}(1-\zeta^4) + [\widetilde{e}^h_{\rm c}(r_s,1) - \widetilde{e}^h_{\rm c}(r_s,0)]f(\zeta)\zeta^4 \,. \tag{2.75}$$

The unknown quantities $\widetilde{e}^h_{\rm c}(r_s,0)$, $\widetilde{e}^h_{\rm c}(r_s,1)$, and $\alpha_{\rm c}(r_s)$ (the latter is the spin stiffness) are all parametrized by the analytic form

$$G(r_s,A,\alpha_1,\beta_1,\beta_2,\beta_3,\beta_4) = -2A(1+\alpha_1 r_s)\ln\left[1 + \frac{1/2A}{\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_4 r_s^2}\right], \tag{2.76}$$

where the parameters $A, \alpha_1, \beta_1, \beta_2, \beta_3, \beta_4$ are given in Table 2.2. This expression is widely used in DFT (see Section 2.3); it is nowadays preferred over the older parametrization of Vosko *et al.* (1980).

The recommended interpolation formula in 2D is that given by Attaccalite *et al.* (2002):

$$\widetilde{e}^h_{\rm c}(r_s,\zeta) = \left(e^{-\beta r_s} - 1\right) e^{(6)}_{\rm x}(r_s,\zeta) + \alpha_0(r_s) + \alpha_1(r_s)\zeta^2 + \alpha_2(r_s)\zeta^4 \,, \tag{2.77}$$

where $\beta = 1.3386$ and

$$e^{(6)}_{\rm x}(r_s,\zeta) = \widetilde{e}^h_{\rm x}(r_s,\zeta) - \left(1 + \frac{3}{8}\zeta^2 + \frac{3}{128}\zeta^4\right)\widetilde{e}^h_{\rm x}(r_s,0) \tag{2.78}$$

Table 2.3 Fitting parameters for the 2D correlation energy per particle, eqns (2.77)–(2.79) (Attaccalite *et al.*, 2002).

	$i=0$	$i=1$	$i=2$
A_i	-0.1925	0.117331	0.0234188
B_i	0.0863136	-3.394×10^{-2}	-0.037093
C_i	0.0572384	-7.66765×10^{-3}	0.0163618
E_i	1.0022	0.4133	1.424301
F_i	-0.02069	0	0
G_i	0.33997	6.68467×10^{-2}	0
H_i	1.747×10^{-2}	7.799×10^{-4}	1.163099

is the Taylor expansion of $\widetilde{e}_{\rm x}(r_s,\zeta)$ [see above, eqns (2.71)–(2.74)] beyond fourth order in ζ. The functions $\alpha_i(r_s)$ have a form similar to the Perdew–Wang interpolation:

$$\alpha_i(r_s) = A_i + (B_i r_s + C_i r_s^2 + D_i r_s^3)\ln\left[1+\frac{1}{E_i r_s + F_i r_s^{3/2} + G_i r_s^2 + H_i r_s^3}\right] . \quad (2.79)$$

The fitting parameters for this formula are listed in Table 2.3 (except for the parameter D_i, which is given by $D_i = -A_i H_i$).

For very low densities ($r_s \gtrapprox 100$ in 3D and $r_s \gtrapprox 34$ in 2D), the electron liquid forms a new phase known as the Wigner crystal, which is not described by the interpolation formulas discussed above. The complete phase diagram of the homogeneous electron liquid is still a subject of active research (Giuliani and Vignale, 2005).

2.3 Approximate functionals

2.3.1 The local-density approximation

The oldest approximation of DFT, the local-density approximation (LDA), was originally proposed by Kohn and Sham (1965). It expresses the xc energy of an inhomogeneous system as the integral over the xc energy density of a homogeneous electron liquid, evaluated at the local density:

$$E_{\rm xc}^{\rm LDA}[n] = \int d^3r\; e_{\rm xc}^h(\bar{n})\big|_{\bar{n}=n(\mathbf{r})} . \quad (2.80)$$

The required input, $e_{\rm xc}^h(n)$, was discussed in Section 2.2.4: the exchange part $e_{\rm x}^h$ is exactly known, and there exist highly accurate parametrizations for the correlation part $e_{\rm c}^h$.[14] The resulting xc potential is

[14] In practice, different approximate parametrizations of $e_{\rm c}^h(n)$ may lead to slightly different results, and people tend to distinguish between different LDAs (e.g., the Perdew–Wang LDA and the Vosko–Wilk–Nusair LDA). However, keep in mind that formally there is one and only one LDA (for each dimensionality), defined by eqn (2.80), with the xc energy density of the uniform electron liquid as input. This is in contrast to the generalized gradient approximations (GGAs) and hybrids, which come in many different flavors.

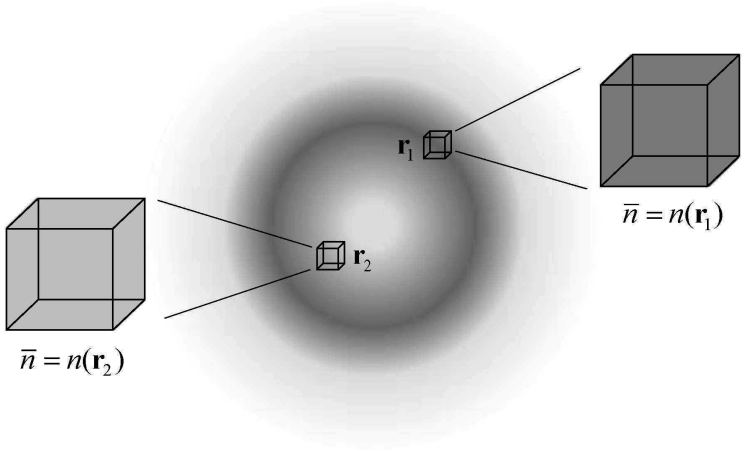

Fig. 2.5 The LDA assumes that the xc energy density in an infinitesimal volume element at position $\mathbf{r}$ is given by the xc energy density of a reference system with a uniform density $\bar{n}$. For $\bar{n}$, one takes the value of the local density $n(\mathbf{r})$ at this position.

$$v_{\rm xc}^{\rm LDA}(\mathbf{r}) = \left.\frac{de_{\rm xc}^h(\bar{n})}{d\bar{n}}\right|_{\bar{n}=n(\mathbf{r})} . \tag{2.81}$$

The physical meaning of the LDA is illustrated in Fig. 2.5: at each point $\mathbf{r}$ in space, the xc energy density $e_{\rm xc}(\mathbf{r})$ is approximated by that which one would obtain from a homogeneous electron liquid that has a density $n(\mathbf{r})$ everywhere.

By construction, the LDA becomes exact in the uniform limit. For nonuniform systems it is reasonable to assume that the LDA works best if the density variations are slow; an appropriate length scale against which density variations can be measured is the inverse of the local Fermi wave vector $k_F(\mathbf{r})$, so that the condition for the validity of the LDA can be formulated as

$$\frac{|\nabla n(\mathbf{r})|}{n(\mathbf{r})} \ll k_F(\mathbf{r}) . \tag{2.82}$$

In practice, this condition is often drastically violated, for instance in the vicinity of nuclei. On the face of it, one would expect the LDA to work only for very special systems where the density is slowly varying throughout, for instance simple metals such as bulk sodium. But it turns out that the LDA works surprisingly well for many energetic and structural properties in a wide range of materials:

- Total atomic and molecular ground-state energies typically lie within 1–5% of the experimental values, with a systematic trend towards overbinding.
- Molecular equilibrium distances and geometries are reproduced within ∼3%.
- Fermi surfaces of bulk metals are reproduced within a few percent.
- Lattice constants of solids are typically reproduced within 2%.
- Vibrational frequencies and phonon energies are excellent (within a few percent).

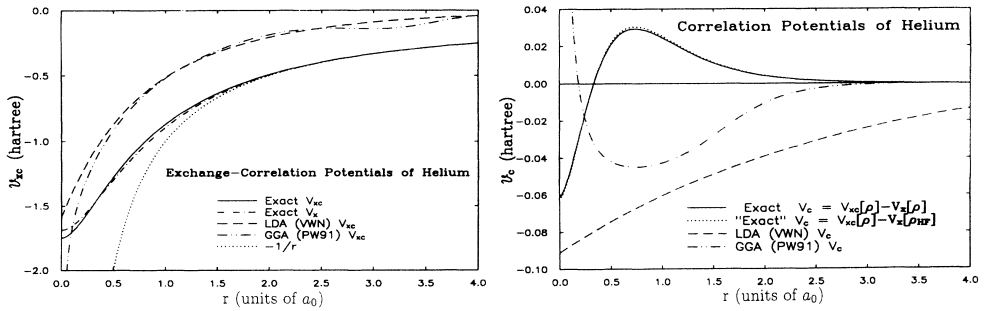

Fig. 2.6 Comparison of exact and approximate xc potentials for the helium atom. The LDA and GGA xc potentials have the wrong asymptotic behavior, approaching zero exponentially instead of as $-1/r$. [Reproduced with permission from APS from Umrigar and Gonze (1994), ©(1994)]

The LDA is almost always better than the HF approximation,[15] and often succeeds in predicting the right physical trends. In practice, the LDA is implemented in its more general, spin-polarized form (the LSDA), where

$$E_{\rm xc}^{\rm LSDA}[n_\uparrow, n_\downarrow] = \int d^3r\, e_{\rm xc}^h\Big(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})\Big)\,, \quad v_{{\rm xc}\sigma}^{\rm LSDA}(\mathbf{r}) = \left.\frac{de_{\rm xc}^h(\bar{n}_\uparrow, \bar{n}_\downarrow)}{d\bar{n}_\sigma}\right|_{\bar{n}_\sigma = n_\sigma(\mathbf{r})}\,. \tag{2.83}$$

The fundamental reason for the unexpected success of the LDA is the fact that it is based on a reference system (the homogeneous electron liquid) which is a real physical system and therefore satisfies a number of exact conditions such as sum rules and scaling properties (Perdew and Kurth, 2003).

One of the major shortcomings of the LDA is that it is not free of self-interaction, and therefore has the wrong asymptotic behavior. Rather than going like $-1/r$ for $r \to \infty$ (see Section 2.2.1), the LDA xc potential goes to zero exponentially fast:[16]

$$v_{\rm xc}^{\rm LDA} \longrightarrow -e^{-\alpha r}\,, \qquad r \to \infty\,. \tag{2.84}$$

This is illustrated in Fig. 2.6 for the helium atom, where the exact xc potential can be constructed numerically (see Appendix E) using a highly accurate ground-state density obtained with quantum Monte Carlo methods (Umrigar and Gonze, 1994). The exact $v_{\rm xc}$ is nicely seen to approach $-1/r$, whereas $v_{\rm xc}^{\rm LDA}$ and the gradient-corrected $v_{\rm xc}^{\rm GGA}$ (see next section) drop to zero too fast.

For two-electron singlet ground states such as those in helium, the exact exchange potential is given by

$$v_{\rm x}(\mathbf{r}) = -\frac{1}{2}\int d^3r'\,\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.85}$$

(this will be explained later in Chapter 11). Figure 2.6 shows that $v_{\rm x}$ makes by far the dominating contribution to the total xc potential, and $v_{\rm x} \longrightarrow -1/r$. The remaining

[15] With some notable exceptions such as molecular dissociation.

[16] This is easily seen from the fact that the density in a finite system drops off exponentially.

correlation potential is shown in the right-hand part of Fig. 2.6, and we see that it approaches zero much faster. Interestingly, the approximate LDA and GGA correlation potentials bear no resemblance at all to the exact correlation potential!

This points to another "deep reason" for the unexpected success of the LDA (as well as that of the gradient-corrected functionals), known as *error cancellation*. LDA exchange energies are typically overestimated, whereas correlation energies are underestimated, and the errors tend to compensate. This is fortunate, but not entirely accidental: the fact that the LDA satisfies exact sum rules mandates that integrated errors in the exchange and correlation cancel out to some extent.

While the wrong asymptotics of $v_{\rm xc}^{\rm LDA}$ has relatively mild consequences for integrated quantities such as the total energy, it is still a rather severe problem. Let us conclude this section with a list of the major shortcomings of the LDA:

- Owing to the wrong asymptotics (caused by spurious self-interaction), the Kohn–Sham energy eigenvalues come out too low in magnitude. In particular, the LDA eigenvalue of the highest occupied orbital differs by ∼30–50% from the ionization energy [see eqn (2.46)].
- The LDA does not produce stable negative ions, since the potential is too shallow to bind an extra electron.
- $v_{\rm xc}^{\rm LDA}$ does not have a derivative discontinuity. LDA band gaps in solids are typically off by a considerable amount (anywhere from around 50% for weakly correlated semiconductors to wrongly predicting metallic behavior in some transition metal oxides that are Mott insulators). For the same reason, the LDA predicts the wrong dissociation limit of molecules, leading to fractionally charged fragments.
- Finally, the LDA, while generally capturing the right physical trends, is not accurate enough for many chemical applications, nor to predict the correct ground states in many materials such as transition metal oxides.

We will soon discover that many of the successes and failures of the LDA carry over into the dynamical regime. As a case in point, the errors in the LDA energy eigenvalues will play an important role in the calculation of excitation energies (see Chapter 9).

2.3.2 Generalized gradient approximations

The idea of improving the LDA by constructing xc functionals which depend not only on the local density itself but also on its gradients has a long history, going back to the original work by Hohenberg and Kohn (1964). The resulting gradient expansion approximation (GEA) is based on the notion that the condition (2.82) holds, at least approximately, and one can therefore use a reduced density gradient as a small parameter for constructing expansions of density functionals in powers (orders) of gradients. One commonly used definition of the reduced density gradient is

$$s(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|}{2n(\mathbf{r})k_F(\mathbf{r})}\,, \tag{2.86}$$

but other forms also appear in the literature, for instance $x = |\nabla n|/n^{4/3} = 2(3\pi^2)^{1/3}s$. The GEA for the xc energy is given by

$$E_{\rm xc}^{\rm GEA}[n] = \int d^3r \left(e_{\rm xc}^h\Big(n(\mathbf{r})\Big) + C_{\rm xc}^{(2)}(n)s^2 + \ldots \right) . \tag{2.87}$$

A similar expansion can be written down for the noninteracting kinetic-energy functional $T_s[n]$.[17]

Over the years, a large amount of work has been invested in deriving terms of the gradient expansion series (up to second order for correlation, fourth order for exchange, and sixth order for the kinetic energy); see reviews by Dreizler and Gross (1990) and Perdew and Kurth (2003). But, eventually, the sobering realization set in that for $E_{\rm xc}$ this strategy does not lead to systematic improvements over the LDA—in many cases, the results even turned worse.[18] The main problem with the GEA, if only the first few terms are kept, is that a very important property of the LDA gets lost, namely, compliance with sum rules.

Generalized gradient approximations are a class of xc functionals which have the following general structure:

$$E_{\rm xc}^{\rm GGA}[n_\uparrow, n_\downarrow] = \int d^3r\, e_{\rm xc}^{\rm GGA}\Big(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}), \nabla n_\uparrow(\mathbf{r}), \nabla n_\downarrow(\mathbf{r})\Big) . \tag{2.88}$$

GGAs do not arise from systematic order-by-order expansions in terms of the density gradients. Instead, the key idea is to construct explicit mathematical expressions for $e_{\rm xc}^{\rm GGA}(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow)$ which satisfy as many of the known exact properties of $e_{\rm xc}[n]$ as possible—in a way, this can be viewed as attempting to find heuristic resummations of the gradient expansion series. There is no unique prescription for doing so: sometimes this requires the help of empirical parameters, and sometimes this is done without any empirical input. Thus, while the GEA is uniquely defined,[19] many different GGAs have been developed over the years (Ma and Brueckner, 1968; Langreth and Mehl, 1983; Perdew, 1986; Perdew and Wang, 1986; Becke, 1986, 1988; Lee *et al.*, 1988; Perdew, 1991; Perdew *et al.*, 1996*a*).

Today there exists a whole zoo of hundreds of GGA functionals, and it is easy to get overwhelmed by this variety. For recent overviews of the performance of GGA functionals and for recommendations, see Kurth *et al.* (1999), Koch and Holthausen (2001), Sousa *et al.* (2007), and Rappaport *et al.* (2009). Here we list only a few examples of the most widely used GGAs.

The exchange functional of Becke (1988) is given by

$$E_{\rm x}^{\rm B88}[n_\uparrow, n_\downarrow] = E_x^{\rm LSDA}[n_\uparrow, n_\downarrow] - \beta \sum_\sigma \int d^3r\, n_\sigma^{4/3} \frac{x_\sigma^2}{1 + 6\beta x_\sigma \sinh^{-1}(x_\sigma)} , \tag{2.89}$$

where $x_\sigma = |\nabla n_\sigma(\mathbf{r})|/n_\sigma^{4/3}(\mathbf{r})$, and $\beta = 0.0042$ a.u. is an empirical parameter determined by fitting to atomic Hartree–Fock exchange energies.

[17] Notice that the condition $s \ll 1$ would not rule out small but rapidly oscillating density variations; an additional condition must be imposed for the gradient expansion to be valid, namely, $|\nabla_i \nabla_j n|/|\nabla n| k_F \ll 1$.

[18] The convergence of gradient expansions is a somewhat problematic issue; mathematically, they constitute so-called asymptotic series, which means that higher-order terms may lead to worse results.

[19] Second- and higher-order gradient terms of the GEA can be constructed using linear- and higher-order response theory. The mathematical effort required can be considerable.

The correlation energy functional of Lee, Yang, and Parr (Lee *et al.*, 1988), in its form for closed-shell systems, reads as follows:

$$E_{\rm c}^{\rm LYP}[n] = -a \int \frac{d^3r}{1+dn} \left\{ n + bn^{-1/3} \left[C_F n^{5/3} - 2t_W + \frac{1}{9} \left(t_W + \frac{\nabla^2}{2} n \right) e^{-cn^{-1/3}} \right] \right\}, \tag{2.90}$$

where

$$t_W = \frac{1}{8} \left(\frac{|\nabla n|^2}{n} - \nabla^2 n \right), \tag{2.91}$$

$C_F = 3/10(3\pi^2)^{2/3}$, and the fitting parameters are $a = 0.049$, $b = 0.132$, $c = 0.2533$, and $d = 0.349$. Combining the Becke and Lee–Yang–Parr functionals yields the popular GGA known as BLYP.

Many GGAs were developed over the years by John Perdew and coworkers, such as the PW91 functional used in Fig. 2.6 (Perdew, 1991). These functionals are nonempirical and are completely determined by the requirement that they satisfy as many exact properties as possible. The PBE functional can be viewed as the culmination of these efforts (Perdew *et al.*, 1996*a*). It consists of an exchange part,

$$E_{\rm x}^{\rm PBE}[n] = \int d^3r \, e_{\rm x}^h(n) \left[1 + \kappa - \frac{\kappa}{1 + \beta\pi^2 s^2/3\kappa} \right], \tag{2.92}$$

where $\kappa = 0.804$ and $\beta = 0.066725$, and s is defined in eqn (2.86). The spin-dependent version of the exchange functional can be obtained using the Oliver–Perdew spin-scaling relation (Oliver and Perdew, 1979)

$$E_{\rm x}[n_\uparrow, n_\downarrow] = \frac{1}{2} E_{\rm x}[2n_\uparrow] + \frac{1}{2} E_{\rm x}[2n_\downarrow] \,. \tag{2.93}$$

The PBE correlation energy is given by

$$E_{\rm c}^{\rm PBE}[n,\zeta] = \int d^3r \left[e_{\rm c}^h(n,\zeta) + nc_0\phi^3 \ln \left\{ 1 + \frac{(1+At^2)\beta t^2/c_0}{1 + At^2 + A^2t^4} \right\} \right], \tag{2.94}$$

where $t = |\nabla n|/2k_s n$, with a screening wave vector $k_s = \sqrt{4k_F/\pi}$; $c_0 = 0.031091$, $\phi = [(1+\zeta)^{2/3} + (1-\zeta)^{2/3}]/2$, and

$$A = \frac{\beta/c_0}{\exp[-\widetilde{e}_{\rm c}(r_s,\zeta)/c_0\phi^3] - 1} \,. \tag{2.95}$$

Deciding which functional to choose is not always easy and may require some experience. A given xc functional may work very well for some properties and some classes of systems, but not in other situations. A meaningful assessment of approximate density functionals should therefore be done in a statistical sense, for instance by comparing the performance using large sets of benchmark systems, called test sets. Table 2.4 gives an example of such a systematic comparison (Staroverov *et al.*, 2003).

Compared with GGA xc energies, which typically agree very well with exact results, GGA xc *potentials* tend to be of much poorer quality.[20] An example is given

[20] Despite producing relatively poor approximations for the exact xc potential, most popular GGAs give excellent results for integrated quantities such as energies. As with the LDA, this is mainly thanks to positive side effects of error cancellation between exchange and correlation.

Table 2.4 Mean absolute errors in several molecular properties calculated for various test sets (Staroverov *et al.*, 2003).

	Formation enthalpy[a]	Ionization potential[b]	Equilibrium bond length[c]	Vibrational frequency[d]	H-bonded complexes[e]
HF	211.54	1.028	0.0249	136.2	2.77
LSDA	121.85	0.232	0.0131	48.9	5.78
BLYP	9.49	0.286	0.0223	55.2	0.64
BPW91	9.04	0.241	0.0168	41.4	0.99
BP86	26.33	0.215	0.0175	45.5	0.76
PW91	23.59	0.221	0.0145	39.8	1.43
PBE	22.22	0.235	0.0159	42.0	1.00
HCTH	7.17	0.232	0.0145	39.9	0.91
OLYP	5.88	0.288	0.0177	40.2	2.18
B3LYP	4.93	0.184	0.0104	33.5	0.43
B3PW91	3.90	0.190	0.0093	36.2	0.88
B3P86	26.14	0.551	0.0084	37.0	0.73
PBE0	6.66	0.199	0.0097	43.6	0.66
VSXC	3.46	0.226	0.0131	33.9	1.34
PKZB	6.98	0.310	0.0269	51.7	2.90
TPSS	5.81	0.242	0.0142	30.4	0.59
TPSSh	3.90	0.229	0.0101	26.9	0.78

[a] For a test set of 223 molecules (in kcal/mol).
[b] For a test set of 223 molecules (in eV), evaluated from the total-energy differences between the cation and the corresponding neutral, for their respective geometries.
[c] For a test set of 96 diatomic molecules (in Å).
[d] For a test set of 82 diatomic molecules (in cm^{-1}).
[e] For a test set of 10 hydrogen-bonded complexes (dissociation energies in kcal/mol).

in Fig. 2.6, where we have already pointed out that they do not have the correct asymptotic behavior $-1/r$. For instance, the exchange potential associated with the Becke B88 functional (2.89), $v_{\mathrm{x}}^{\mathrm{B88}}$, decays as $-k/r^2$, where k is some constant. There have been attempts to directly construct GGA expressions for xc potentials with the correct asymptotic behavior. One such example is the LB94 potential (van Leeuwen and Baerends, 1994),

$$v_{\mathrm{xc}\sigma}^{\mathrm{LB94}}(\mathbf{r}) = -\beta n_\sigma^{1/3}(\mathbf{r})\,\frac{x_\sigma^2}{1+3\beta x_\sigma \sinh^{-1}(x_\sigma)}\,, \tag{2.96}$$

with the fitting parameter $\beta = 0.05$. The LB94 potential gives a marked improvement

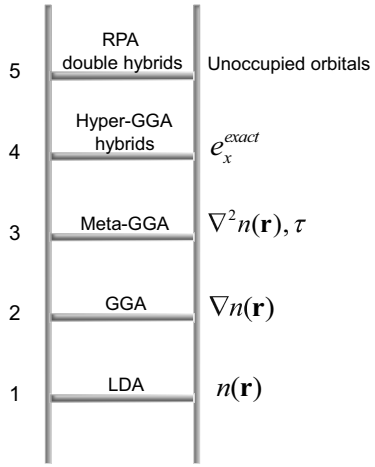

Fig. 2.7 The ladder of approximations which constitutes today's standard taxonomy of xc functionals in DFT. At each new rung of the ladder, additional elements enter into the construction of xc functionals.

in the Kohn–Sham eigenvalues and produces stable negative ions, but its overall use in electronic structure calculation has remained limited. In Section 9.2, we will see that it leads to considerable improvement in the description of Rydberg excitations.

2.3.3 Climbing the ladder of approximations

The LDA and a handful of the most popular GGAs have been instrumental in the success of DFT in the calculation of the electronic structures of materials. However, for many purposes they are not sufficiently accurate: for example, reliable prediction of chemical reactions requires so-called "chemical accuracy," which limits errors to below 1 kcal/mol (a common measure of energy in chemistry, which corresponds to 0.043 eV per particle). Numerous efforts have therefore been directed at going beyond the GGA to construct more accurate xc functionals.

A standard model of the taxonomy of xc functionals is to arrange them in a ladder of approximations,[21] as shown in Fig. 2.7. The ground level is Hartree theory; at the lowest two rungs we find the LDA and GGA. Let us now see what comes next.

Third rung: meta-GGAs. The GGAs (2.88) are density functionals which are constructed using the spin densities and their gradients as input. To achieve more flexibility, meta-GGAs include additional elements:

$$E_{\rm xc}^{\rm MGGA} = \int d^3r\, e_{\rm xc}^{\rm MGGA}(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow, \nabla^2 n_\uparrow, \nabla^2 n_\downarrow, \tau_\uparrow, \tau_\downarrow)\,. \tag{2.97}$$

In addition to the Laplacians of the spin densities, meta-GGAs also include the Kohn–Sham orbital kinetic-energy densities

[21] This picture was proposed by Perdew and Schmidt (2001) as the "Jacob's ladder of DFT," extending from the Hartree world to the heaven of chemical accuracy.

$$\tau_\sigma(\mathbf{r}) = \frac{1}{2}\sum_{j=1}^{\text{occ}} |\nabla\varphi_{j\sigma}(\mathbf{r})|^2 \,. \tag{2.98}$$

Some examples of meta-GGAs are the PKZB (Perdew *et al.*, 1999), TPSS (Tao *et al.*, 2003), and VSXC (Van Voorhis and Scuseria, 1998) functionals. A major advantage of these functionals is that they are partially free of self-interaction, and satisfy additional constraints not obeyed by standard GGAs (Staroverov *et al.*, 2004). The Laplacian of the density has the somewhat undesirable feature of diverging at the nucleus owing to the cusp condition (see Fig. 2.1); however, the PKZB and TPSS functionals do not use the Laplacian.

A more serious complication of meta-GGAs is that the kinetic-energy density τ is not an explicit functional of the density, which means that the functional derivative $v_{\text{xc}} = \delta E_{\text{xc}}/\delta n$ cannot be evaluated directly. Later, in Chapter 11, we will discuss how such functionals, which are implicit functionals of the density but explicit orbital functionals, can be formally treated within DFT, using the optimized-effective-potential (OEP) approach. Such a treatment, however, would be computationally expensive.

In practice, meta-GGAs are usually implemented by going outside the strict limits of Kohn–Sham theory and making the total energy stationary with respect to orbital variations. This yields a differential operator in place of the local multiplicative Kohn–Sham potential (Staroverov *et al.*, 2003). Alternatively, meta-GGAs and other implicit density functionals can also be applied in a non-self-consistent manner by evaluating total energies with GGA Kohn–Sham orbitals. In this way, the total computational costs with meta-GGAs are not significantly higher than for ordinary GGAs.

Fourth rung: hyper-GGAs and hybrid functionals. The next level of refinement is achieved by incorporating the exact exchange-energy density,

$$e_{\text{x}}^{\text{exact}}(\mathbf{r}) = -\frac{1}{2}\sum_\sigma \sum_{i,j=1}^{N_\sigma} \int d^3r' \, \frac{\varphi^*_{i\sigma}(\mathbf{r}')\varphi_{j\sigma}(\mathbf{r}')\varphi_{i\sigma}(\mathbf{r})\varphi^*_{j\sigma}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \,. \tag{2.99}$$

Like the kinetic-energy density (2.98), $e_{\text{x}}^{\text{exact}}$ is an implicit density functional, but an explicit orbital functional. This raises issues for the construction of the Kohn–Sham potential similar to those for the meta-GGAs, which will be resolved in Chapter 11.

The class of approximate xc functionals which (in addition to the density, density gradients, and kinetic-energy density) depend on $e_{\text{x}}^{\text{exact}}$ is known as hyper-GGAs (Perdew *et al.*, 2008; Odashima and Capelle, 2009). While general hyper-GGAs (which use the exchange energy density to construct approximate correlation functionals) have not yet found widespread use, a particular subgroup has been enormously successful: the hybrid functionals.[22]

The basic idea is simple: hybrid functionals are constructed by mixing a fraction of the exact exchange energy functional with a standard LDA or GGA, in the following generic way:

$$E_{\text{xc}}^{\text{hybrid}} = aE_{\text{x}}^{\text{exact}} + (1-a)E_{\text{x}}^{\text{GGA}} + E_{\text{c}}^{\text{GGA}} \,, \tag{2.100}$$

[22] A detailed account of the historic development and the properties of hybrid functionals, with many examples, can be found in Koch and Holthausen (2001).

where the semiempirical constant a has a typical value of around 0.25 (Becke, 1993, 1996; Perdew *et al.*, 1996*b*). The PBE0 functional is obtained by substituting the PBE functional into eqn (2.100) (Adamo and Barone, 1999). The most popular hybrid functional today is the three-parameter functional known as the B3LYP functional (Stephens *et al.*, 1994):

$$E_{\rm xc}^{\rm B3LYP} = (1-a)E_{\rm x}^{\rm LDA} + aE_{\rm x}^{\rm exact} + bE_{\rm x}^{\rm B88} + cE_{\rm c}^{\rm LYP} + (1-c)E_{\rm c}^{\rm LDA}\,, \qquad (2.101)$$

where $a = 0.20$, $b = 0.72$, and $c = 0.81$. According to a recent survey (Sousa *et al.*, 2007), the B3LYP functional is used today for about 80% of all DFT applications. Taking a look at Table 2.4, we see that B3LYP indeed outperforms most other GGAs and meta-GGAs for the structural and energetic properties of molecules close to equilibrium. This has been confirmed in a myriad of studies in theoretical chemistry over the past decade. From a formal standpoint, the success of B3LYP is somewhat unexpected: the construction is based on pragmatism rather than compelling theoretical arguments. However, B3LYP does not work for everything: for instance, since the LYP correlation functional does not reproduce the correct limit for homogeneous systems, B3LYP fails for "free-electron-like" metallic systems (Paier *et al.*, 2007).

A particular class of hybrid functionals, called *range-separated hybrids*, has attracted much interest lately (Baer *et al.*, 2010). The basic idea is to separate the Coulomb interaction into a short-range (SR) and a long-range (LR) part:

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{f(\mu|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} + \frac{1-f(\mu|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}\,, \qquad (2.102)$$

where the function f has the properties $f(\mu x \to 0) = 1$ and $f(\mu x \to \infty) = 0$. Common examples are

$$f(\mu x) = e^{-\mu x} \qquad \text{and} \qquad f(\mu x) = \mathrm{erfc}(\mu x) = \frac{2}{\sqrt{\pi}}\int_x^\infty dx' e^{-x'^2}\,. \qquad (2.103)$$

The separation parameter μ is determined either empirically (Iikura *et al.*, 2001; Baer and Neuhauser, 2005; Gerber and Ángyán, 2005) or using physical arguments (Livshits and Baer, 2007; Baer *et al.*, 2010). The resulting range-separated hybrid xc functional then has the following generic form (Vydrov and Scuseria, 2006):

$$E_{\rm xc} = E_{\rm x}^{\rm SR-DFA} + E_{\rm x}^{\rm LR-HF} + E_{\rm c}^{\rm DFA}\,. \qquad (2.104)$$

Here, DFA stands for any standard density-functional approximation to the exchange or correlation energy, such as the LDA or GGA. The correlation part is taken to be the same as for the standard (nonhybrid) case, but the exchange parts have to be recomputed since they depend on the particular choice of separation (2.102).[23]

The main strength of range-separated hybrid functionals is that they produce effective potentials with the correct long-range asymptotic behavior. This, in turn, leads to a significant improvement in properties such as the polarizabilities of long-chain molecules, bond dissociation, and, particularly importantly for TDDFT, Rydberg and charge-transfer excitations. The latter will be discussed in Chapter 9.

[23]The local exchange energy of an electron liquid with a Yukawa or erfc interaction (2.103) was calculated in Iikura *et al.* (2001) and Robinson *et al.* (1962).

Fifth rung: unoccupied orbitals. In the development of hybrid functionals (the third and fourth rungs of the ladder), we have seen that full exact exchange cannot be combined with density-based approximate GGA correlation functionals for all $\mathbf{r}$, since we lose the advantage of error cancellation. This is why hybrids incorporate only partial exact exchange; but this does not reproduce the correct asymptotic behavior of the potential. Range-separated hybrids work in the asymptotic range, but problems at short distances remain. So what can we do to get a better description of correlation?

We know the exact exchange energy (2.99) as a functional of the *occupied* orbitals. No simple closed expression for the exact correlation energy is available; we have to rely on many-body perturbation theory for the next-higher-order terms beyond exchange, on partial resummations of certain diagrams in the perturbation series, or on other sophisticated many-body schemes (see Chapter 13).

The fifth rung of the ladder of approximations therefore makes a major leap forward and includes *unoccupied* orbitals in the construction of correlation functionals. Remember that all Kohn–Sham orbitals, whether occupied or empty, are implicit functionals of the density; therefore, such schemes are still formally in the domain of DFT (and require, in principle, the OEP approach). However, treating correlation partially exactly in this manner comes with a significant increase in computational cost, and applications in electronic-structure theory are only slowly beginning to emerge.

For instance, Grimme (2006) has constructed a double hybrid functional (B2PLYP) which incorporates correlation at a second-order perturbative level (MP2):

$$E_{\rm xc} = (1-a)E_{\rm x}^{\rm GGA} + aE_{\rm x}^{\rm exact} + bE_{\rm c}^{\rm GGA} + cE_{\rm c}^{\rm MP2}\,. \tag{2.105}$$

Significant improvements over both B3LYP and MP2 in organic main-group chemistry were achieved with this approach, whose computational cost tends to be a factor of two or so higher than that of B3LYP.

While MP2-based double hybrids have become popular in the chemistry community, the preferred approach in physics is the random-phase approximation (RPA). The RPA was first brought into the DFT context through the work of Langreth and Perdew (1977, 1980). In the language of quantum chemistry, the RPA method is equivalent to a ring coupled-cluster doubles approach (Scuseria *et al.*, 2008). The RPA will turn out to be very important for the description of dispersion forces, and we will discuss it in detail in Chapter 14. As we will see, the RPA is based on the fluctuation–dissipation theorem of linear-response theory, which will be introduced in Chapter 7.

2.3.4 Other approximations

Most approximate xc functionals used today fall within the "standard model" of Fig. 2.7, but there are many other examples which lie outside these categories. For instance, an occasionally encountered class of nonlocal density functionals is that of the average-density approximation and the weighted-density approximation (ADA and WDA), which will not be discussed further here (Dreizler and Gross, 1990).

In solid-state physics, many methodologies have been developed which combine the principles of DFT with non-DFT many-body methods. Examples are the LDA+U and GW methods. We will come back to the latter in Chapter 13.

Lastly, let us discuss in a bit more detail another important example. As we pointed out in Section 2.2.2, the exact xc energy functional is free of self-interaction. This requirement can be expressed in compact form via the condition (2.50). Approximate xc energy functionals $E_{\rm xc}^{\rm app}$ often do not satisfy this condition, i.e.,

$$E_{\rm H}[n_{j\sigma}] + E_{\rm xc}^{\rm app}[n_{j\sigma}, 0] \neq 0\,. \tag{2.106}$$

This has several undesirable consequences in practice, the most severe of them being an asymptotic behavior different from $-1/r$.

Perdew and Zunger (1981) proposed a simple and straightforward cure for this problem, known as the self-interaction correction (SIC).[24] The idea of the SIC is to subtract the self-interaction error (2.106) for each individual orbital:

$$\begin{aligned} E_{\rm xc}^{\rm SIC} &= E_{\rm xc}^{\rm app}[n_\uparrow, n_\downarrow] \\ &\quad - \sum_j \left(E_{\rm H}[n_{j\uparrow}] + E_{\rm xc}^{\rm app}[n_{j\uparrow}, 0]\right) - \sum_j \left(E_{\rm H}[n_{j\downarrow}] + E_{\rm xc}^{\rm app}[0, n_{j\downarrow}]\right)\,. \end{aligned} \tag{2.107}$$

The self-interaction corrected xc energy thus becomes a functional of the orbital densities $n_{j\sigma}$. This SIC prescription can be implemented in a self-consistent manner by minimizing $E_{\rm xc}^{\rm SIC}$ with respect to the orbitals; this has the formal difficulty that it leads to state-dependent single-particle potentials. The resulting single-particle orbitals are no longer orthogonal to each other and must, at least in principle, be reorthogonalized. Furthermore, they are not invariant under unitary transformations.

Like the meta-GGAs, the SIC belongs to the class of implicit density functionals and explicit orbital functionals, and we will discuss in Chapter 11 how to use the OEP scheme to construct the Kohn–Sham xc potential (which is not state-dependent) for functionals of this type.

2.3.5 Lower-dimensional systems

The vast majority of applications of DFT take place in the three-dimensional world of molecules and solids. However, low-dimensional systems have become of increasing interest in nanoscale science, and DFT has made important contributions for a wide range of 2D and 1D systems such as quantum wells, quantum dots, quantum wires, and atomic and molecular chains. We therefore briefly conclude this chapter with some recent developments in the field of approximate xc functionals in lower dimensions.

In 2D, the LSDA is constructed in the same way as in 3D (see Exercise 2.10). There has been some recent work on going beyond the LSDA and developing gradient-corrected exchange and correlation functionals Pittalis *et al.* (2009, 2010). These functionals have been applied to study the electronic structure of quantum dots.

In 1D, the situation is more complicated: the Coulomb interaction involves some mathematical subtleties, and electrons form Luttinger liquids instead of Fermi liquids (Giuliani and Vignale, 2005). Hubbard-type 1D lattice systems are an important class of interacting 1D systems for which a DFT treatment has recently become possible (Lima *et al.*, 2003). Applications to strongly correlated 1D impurity systems, quantum spin chains, and trapped cold atoms are beginning to emerge (Silva *et al.*, 2005; Xianlong *et al.*, 2006; Abedinpour *et al.*, 2007; Alcaraz and Capelle, 2007).

[24] Other SIC prescriptions exist in the literature, but eqn (2.107) is by far the most widely used.

Exercise 2.1 To get an impression of the "exponential wall" one encounters when dealing with many-body wave functions, let us do a back-of-the envelope calculation. Imagine that we represent a wave function $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ (where we ignore spin) with 10 parameters for each spatial coordinate; these could be grid points, or coefficients of basis functions. Without any spatial symmetry, a one-electron system therefore requires $10^3 = 1000$ numbers to fully represent it. In single precision, this corresponds to 1 kilobyte of data that needs to be stored.

How many bytes would you need to store the wave function of a two-electron system, and how mamy for a 10-electron system? Assume that a hard drive has a capacity of 1 Terabyte; how many hard drives would you need? How large could N be before the storage requirement exceeds the number of atoms in the universe (about 10^{80})?

Exercise 2.2 Verify that the density of a noninteracting system is given by the sum over the squares of single-particle orbitals in eqn (2.37) or (2.41), by inserting a single Slater determinant in the definition (C.1) of the one-particle density.

Exercise 2.3 Review Appendix C, and verify all definitions of densities and density matrices for the case where the wave functions are two- and three-electron Slater determinants.

Exercise 2.4 Check Kato's theorem for the case of the hydrogen atom. Show that the slope of the hydrogenic density at the cusp determines the nuclear charge.

Exercise 2.5 Derive the Hartree potential $v_{\rm H}(\mathbf{r}) = \int d^3r'\, n(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$ by evaluating the functional derivative (see Appendix B) of the Hartree energy (2.31) with respect to the density.

Exercise 2.6 Calculate the Hartree potential of the hydrogen atom and reproduce Fig. 2.2. Convince yourself that the correlation potential has a $-1/r$ asymptotic behavior.

Exercise 2.7 Consider a two-electron system with a doubly occupied Kohn–Sham orbital (e.g., the helium atom). Show that the exact exchange energy (2.51) is minus one-half of the Hartree energy in this case.

Exercise 2.8 (a) Find that potential of a one-dimensional system which corresponds to the ground-state density $n_0(x) = \pi^{-1/2} e^{-x^2}$, using the procedure of Appendix E.

(b) Calculate the density $n_1(x)$ of the first excited state of a single particle in a harmonic-oscillator potential, and use again the inversion procedure of Appendix E to determine the potential that produces it. Compare this with what you got in part (a), and discuss.

Exercise 2.9 Calculate the kinetic energy per unit volume of a noninteracting homogeneous electron liquid, $\tau_s^h(n) = (3/10)(3\pi^2)^{2/3} n^{5/3}$.

Exercise 2.10 Find explicit forms of the LSDA exchange potentials in 2D and 3D by using the exchange energy densities given in Section 2.2.4.

Exercise 2.11 Check the Oliver–Perdew spin-scaling relation (2.93) for the LSDA.

Exercise 2.12 Calculate the LDA and the LSDA exchange energies and potentials for the hydrogen atom. Do you notice a difference? How do they compare with the exact answer?

Exercise 2.13 Calculate the B88 exchange potential from eqn (2.89), and find its asymptotic behavior. Compare this with the asymptotic behavior of the LB94 xc potential, eqn (2.96).