Density Functionals for Non-Relativistic Coulomb Systems

John P. Perdew and Stefan Kurth

Department of Physics and Quantum Theory Group
Tulane University, New Orleans, LA 70118, USA

Outline

1 Introduction
1.1 Quantum Mechanical Many-Electron Problem
1.2 Summary of Kohn-Sham Spin-Density Functional Theory
2 Wavefunction Theory
2.1 Wavefunctions and Their Interpretation
2.2 Wavefunctions for Non-Interacting Electrons
2.3 Wavefunction Variational Principle
2.4 Hellmann-Feynman Theorem
2.5 Virial Theorem
3 Definitions of Density Functionals
3.1 Introduction to Density Functionals
3.2 Density Variational Principle
3.3 Kohn-Sham Non-Interacting System
3.4 Exchange Energy and Correlation Energy
3.5 Coupling-Constant Integration
4 Formal Properties of Functionals
4.1 Uniform Coordinate Scaling
4.2 Local Lower Bounds
4.3 Spin Scaling Relations
4.4 Size Consistency
4.5 Derivative Discontinuity
5 Uniform Electron Gas
5.1 Kinetic Energy
5.2 Exchange Energy
5.3 Correlation Energy
5.4 Linear Response
5.5 Clumping and Adiabatic Connection
6 Local and Semi-Local Approximations
6.1 Local Spin Density Approximation
6.2 Gradient Expansion
6.3 History of Several Generalized Gradient Approximations
6.4 Construction of a "GGA Made Simple"
6.5 GGA Nonlocality: Its Character, Origins, and Effects
References

1 Introduction

1.1 Quantum Mechanical Many-Electron Problem

The material world of everyday experience, as studied by chemistry and condensed-matter physics, is built up from electrons and a few (or at most a few hundred) kinds of nuclei . The basic interaction is electrostatic or Coulombic: An electron at position $\mathbf{r}$ is attracted to a nucleus of charge Z at $\mathbf{R}$ by the potential energy $-Z/|\mathbf{r}-\mathbf{R}|$, a pair of electrons at $\mathbf{r}$ and $\mathbf{r}'$ repel one another by the potential energy $1/|\mathbf{r}-\mathbf{r}'|$, and two nuclei at $\mathbf{R}$ and $\mathbf{R}'$ repel one another as $Z'Z/|\mathbf{R}-\mathbf{R}'|$. The electrons must be described by quantum mechanics, while the more massive nuclei can sometimes be regarded as classical particles. All of the

electrons in the lighter elements, and the chemically important valence electrons in most elements, move at speeds much less than the speed of light, and so are non-relativistic.

In essence, that is the simple story of practically everything. But there is still a long path from these general principles to theoretical prediction of the structures and properties of atoms, molecules, and solids , and eventually to the design of new chemicals or materials. If we restrict our focus to the important class of ground-state properties, we can take a shortcut through density functional theory.

These lectures present an introduction to density functionals for non-relativistic Coulomb systems. The reader is assumed to have a working knowledge of quantum mechanics at the level of one-particle wavefunctions $\psi(\mathbf{r})$ [1]. The many-electron wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ [2] is briefly introduced here, and then replaced as basic variable by the electron density $n(\mathbf{r})$. Various terms of the total energy are defined as functionals of the electron density, and some formal properties of these functionals are discussed. The most widely-used density functionals - the local spin density and generalized gradient approximations - are then introduced and discussed. At the end, the reader should be prepared to approach the broad literature of quantum chemistry and condensed-matter physics in which these density functionals are applied to predict diverse properties: the shapes and sizes of molecules, the crystal structures of solids, binding or atomization energies, ionization energies and electron affinities, the heights of energy barriers to various processes, static response functions, vibrational frequencies of nuclei, etc. Moreover, the reader's approach will be an informed and discerning one, based upon an understanding of where these functionals come from, why they work, and how they work.

These lectures are intended to teach at the introductory level, and not to serve as a comprehensive treatise. The reader who wants more can go to several excellent general sources [3, 4, 5] or to the original literature. Atomic units (in which all electromagnetic equations are written in cgs form, and the fundamental constants $\hbar$, e^2, and m are set to unity) have been used throughout.

1.2 Summary of Kohn-Sham Spin-Density Functional Theory

This introduction closes with a brief presentation of the Kohn-Sham [6] spin-density functional method , the most widely-used method of electronic-structure calculation in condensed-matter physics and one of the most widely-used methods in quantum chemistry. We seek the ground-state total energy E and spin densities $n_\uparrow(\mathbf{r})$, $n_\downarrow(\mathbf{r})$ for a collection of N electrons interacting with one another and with an external potential $v(\mathbf{r})$ (due to the nuclei in most practical cases). These are found by the selfconsistent solution of an auxiliary (fictitious) one-electron Schrödinger equation:

$$\left(-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + u([n];\mathbf{r}) + v_{\mathrm{xc}}^{\sigma}([n_\uparrow, n_\downarrow];\mathbf{r})\right)\psi_{\alpha\sigma}(\mathbf{r}) = \varepsilon_{\alpha\sigma}\psi_{\alpha\sigma}(\mathbf{r}) \ , \tag{1}$$

$$n_\sigma(\mathbf{r}) = \sum_\alpha \theta(\mu - \varepsilon_{\alpha\sigma})|\psi_{\alpha\sigma}(\mathbf{r})|^2 \ . \tag{2}$$

Here $\sigma = \uparrow$ or $\downarrow$ is the z-component of spin, and α stands for the set of remaining one-electron quantum numbers. The effective potential includes a classical Hartree potential

$$u([n];\mathbf{r}) = \int d^3r' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ , \tag{3}$$

$$n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r}) \ , \tag{4}$$

and a multiplicative spin-dependent exchange-correlation potential $v_{\rm xc}^\sigma([n_\uparrow,n_\downarrow];\mathbf{r})$, which is a functional of the spin densities. The step function $\theta(\mu-\varepsilon_{\alpha\sigma})$ in Eq. (2) ensures that all Kohn-Sham spin orbitals with $\varepsilon_{\alpha\sigma} < \mu$ are singly occupied, and those with $\varepsilon_{\alpha\sigma} > \mu$ are empty. The chemical potential μ is chosen to satisfy

$$\int d^3r\, n(\mathbf{r}) = N \ . \tag{5}$$

Because Eqs. (1) and (2) are interlinked, they can only be solved by iteration to selfconsistency.

The total energy is

$$E = T_s[n_\uparrow,n_\downarrow] + \int d^3r\, n(\mathbf{r})v(\mathbf{r}) + U[n] + E_{\rm xc}[n_\uparrow,n_\downarrow] \ , \tag{6}$$

where

$$T_s[n_\uparrow,n_\downarrow] = \sum_\sigma \sum_\alpha \theta(\mu-\varepsilon_{\alpha\sigma})\langle\psi_{\alpha\sigma}| -\frac{1}{2}\nabla^2|\psi_{\alpha\sigma}\rangle \tag{7}$$

is the non-interacting kinetic energy , a functional of the spin densities because (as we shall see) the external potential $v(\mathbf{r})$ and hence the Kohn-Sham orbitals are functionals of the spin densities. In our notation,

$$\langle\psi_{\alpha\sigma}|\hat{O}|\psi_{\alpha\sigma}\rangle = \int d^3r\, \psi^*_{\alpha\sigma}(\mathbf{r})\hat{O}\psi_{\alpha\sigma}(\mathbf{r}) \ . \tag{8}$$

The second term of Eq. (6) is the interaction of the electrons with the external potential. The third term of Eq. (6) is the Hartree electrostatic self-repulsion of the electron density

$$U[n] = \frac{1}{2}\int d^3r \int d^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{9}$$

The last term of Eq. (6) is the exchange-correlation energy , whose functional derivative (as explained later) yields the exchange-correlation potential

$$v_{\rm xc}^\sigma([n_\uparrow,n_\downarrow];\mathbf{r}) = \frac{\delta E_{\rm xc}}{\delta n_\sigma(\mathbf{r})} \ . \tag{10}$$

Not displayed in Eq. (6), but needed for a system of electrons and nuclei, is the electrostatic repulsion among the nuclei. $E_{\rm xc}$ is defined to include everything else omitted from the first three terms of Eq. (6).

If the exact dependence of $E_{\rm xc}$ upon $n_\uparrow$ and $n_\downarrow$ were known, these equations would predict the exact ground-state energy and spin-densities of a many-electron system. The forces on the nuclei , and their equilibrium positions, could then be found from $-\frac{\partial E}{\partial \mathbf{R}}$.

In practice, the exchange-correlation energy functional must be approximated. The local spin density [6, 7] (LSD) approximation has long been popular in solid state physics:

$$E_{\rm xc}^{\rm LSD}[n_\uparrow,n_\downarrow] = \int d^3r\, n(\mathbf{r}) e_{\rm xc}(n_\uparrow(\mathbf{r}),n_\downarrow(\mathbf{r})) \ , \tag{11}$$

where $e_{\rm xc}(n_\uparrow,n_\downarrow)$ is the known [8, 9, 10] exchange-correlation energy per particle for an electron gas of uniform spin densities $n_\uparrow$, $n_\downarrow$. More recently, generalized gradient approximations (GGA's) [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21] have become popular in quantum chemistry:

$$E_{\rm xc}^{\rm GGA}[n_\uparrow,n_\downarrow] = \int d^3r\, f(n_\uparrow,n_\downarrow,\nabla n_\uparrow,\nabla n_\downarrow) \ . \tag{12}$$

The input $e_{\rm xc}(n_\uparrow,n_\downarrow)$ to LSD is in principle unique, since there is a possible system in which $n_\uparrow$ and $n_\downarrow$ are constant and for which LSD is exact. At least in this sense, there is no unique input $f(n_\uparrow,n_\downarrow,\nabla n_\uparrow,\nabla n_\downarrow)$ to GGA. These lectures will stress a conservative "philosophy of approximation" [20, 21], in which we construct a nearly-unique GGA with all the known correct formal features of LSD, plus others.

The equations presented here are really all that we need to do a practical calculation for a many-electron system. They allow us to draw upon the intuition and experience we have developed for one-particle systems. The many-body effects are in $U[n]$ (trivially) and $E_{\rm xc}[n_\uparrow,n_\downarrow]$ (less trivially), but we shall also develop an intuitive appreciation for $E_{\rm xc}$.

While $E_{\rm xc}$ is often a relatively small fraction of the total energy of an atom, molecule, or solid (minus the work needed to break up the system into separated electrons and nuclei), the contribution from $E_{\rm xc}$ is typically about 100 % or more of the chemical bonding or atomization energy (the work needed to break up the system into separated neutral atoms). $E_{\rm xc}$ is a kind of "glue", without which atoms would bond weakly if at all. Thus, accurate approximations to $E_{\rm xc}$ are essential to the whole enterprise of density functional theory. Table 1 shows the typical relative errors we find from selfconsistent calculations within the LSD or GGA approximations of Eqs. (11) and (12). Table 2 shows the mean absolute errors in the atomization energies of 20 molecules when calculated by LSD, by GGA, and in the Hartree-Fock approximation . Hartree-Fock treats exchange exactly, but neglects correlation completely. While the Hartree-Fock total energy is an upper bound to the true ground-state total energy, the LSD and GGA energies are not.

Table 1. Typical errors for atoms, molecules, and solids from selfconsistent Kohn-Sham calculations within the LSD and GGA approximations of Eqs. (11) and (12). Note that there is typically some cancellation of errors between the exchange (E_x) and correlation (E_c) contributions to $E_{\rm xc}$. The "energy barrier" is the barrier to a chemical reaction that arises at a highly-bonded intermediate state.

Property	LSD	GGA
E_x	5 % (not negative enough)	0.5 %
E_c	100 % (too negative)	5 %
bond length	1 % (too short)	1 % (too long)
structure	overly favors close packing	more correct
energy barrier	100 % (too low)	30 % (too low)

In most cases we are only interested in small total-energy changes associated with re-arrangements of the outer or valence electrons , to which the inner or core electrons of the

Table 2. Mean absolute error of the atomization energies for 20 molecules, evaluated by various approximations. (1 hartree = 27.21 eV) (From Ref. [20].)

Approximation	Mean absolute error (eV)
Unrestricted Hartree-Fock	3.1 (underbinding)
LSD	1.3 (overbinding)
GGA	0.3 (mostly overbinding)
Desired "chemical accuracy"	0.05

atoms do not contribute. In these cases, we can replace each core by the pseudopotential [22] it presents to the valence electrons, and then expand the valence-electron orbitals in an economical and convenient basis of plane waves. Pseudopotentials are routinely combined with density functionals. Although the most realistic pseudopotentials are nonlocal operators and not simply local or multiplication operators, and although density functional theory in principle requires a local external potential, this inconsistency does not seem to cause any practical difficulties.

There are empirical versions of LSD and GGA, but these lectures will only discuss nonempirical versions. If every electronic-structure calculation were done at least twice, once with nonempirical LSD and once with nonempirical GGA, the results would be useful not only to those interested in the systems under consideration but also to those interested in the development and understanding of density functionals.

2 Wavefunction Theory

2.1 Wavefunctions and Their Interpretation

We begin with a brief review of one-particle quantum mechanics [1]. An electron has spin $s=\frac{1}{2}$ and z-component of spin $\sigma=+\frac{1}{2}$ ($\uparrow$) or $-\frac{1}{2}$ ($\downarrow$). The Hamiltonian or energy operator for one electron in the presence of an external potential $v(\mathbf{r})$ is

$$\hat{h}=-\frac{1}{2}\nabla^2+v(\mathbf{r}) \ . \tag{13}$$

The energy eigenstates $\psi_\alpha(\mathbf{r},\sigma)$ and eigenvalues ε_α are solutions of the time-independent Schrödinger equation

$$\hat{h}\psi_\alpha(\mathbf{r},\sigma)=\varepsilon_\alpha\psi_\alpha(\mathbf{r},\sigma) \ , \tag{14}$$

and $|\psi_\alpha(\mathbf{r},\sigma)|^2\mathrm{d}^3r$ is the probability to find the electron with spin σ in volume element d^3r at $\mathbf{r}$, given that it is in energy eigenstate ψ_α. Thus

$$\sum_\sigma\int\mathrm{d}^3r|\psi_\alpha(\mathbf{r},\sigma)|^2=\langle\psi|\psi\rangle=1 \ . \tag{15}$$

Since $\hat{h}$ commutes with $\hat{s}_z$, we can choose the ψ_α to be eigenstates of $\hat{s}_z$, i.e., we can choose $\sigma=\uparrow$ or $\downarrow$ as a one-electron quantum number.

The Hamiltonian for N electrons in the presence of an external potential $v(\mathbf{r})$ is [2]

$$\begin{aligned}\hat{H}&=\sum_{i=1}^N-\frac{1}{2}\nabla_i^2+\sum_{i=1}^N v(\mathbf{r}_i)+\frac{1}{2}\sum_i\sum_{j\neq i}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\\&=\hat{T}+\hat{V}_{\mathrm{ext}}+\hat{V}_{\mathrm{ee}} \ .\end{aligned} \tag{16}$$

The electron-electron repulsion $\hat{V}_{\mathrm{ee}}$ sums over distinct pairs of different electrons. The states of well-defined energy are the eigenstates of $\hat{H}$:

$$\hat{H}\Psi_k(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N)=E_k\Psi_k(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N) \ , \tag{17}$$

where k is a complete set of many-electron quantum numbers; we shall be interested mainly in the ground state or state of lowest energy, the zero-temperature equilibrium state for the electrons.

Because electrons are fermions, the only physical solutions of Eq. (17) are those wavefunctions that are antisymmetric [2] under exchange of two electron labels i and j:

$$\begin{aligned}&\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_i\sigma_i,\ldots,\mathbf{r}_j\sigma_j,\ldots,\mathbf{r}_N\sigma_N)=\\&-\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_j\sigma_j,\ldots,\mathbf{r}_i\sigma_i,\ldots,\mathbf{r}_N\sigma_N) \ .\end{aligned} \tag{18}$$

There are $N!$ distinct permutations of the labels $1,2,\ldots,N$, which by Eq. (18) all have the same $|\Psi|^2$. Thus $N!|\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N)|^2\mathrm{d}^3r_1\ldots\mathrm{d}^3r_N$ is the probability to find *any* electron with spin σ_1 in volume element d^3r_1, etc., and

$$\frac{1}{N!}\sum_{\sigma_1\ldots\sigma_N}\int\mathrm{d}^3r_1\ldots\int\mathrm{d}^3r_N N!|\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N)|^2=\int|\Psi|^2=\langle\Psi|\Psi\rangle=1 \ . \tag{19}$$

We define the electron spin density $n_\sigma(\mathbf{r})$ so that $n_\sigma(\mathbf{r})\mathrm{d}^3r$ is the probability to find an electron with spin σ in volume element d^3r at $\mathbf{r}$. We find $n_\sigma(\mathbf{r})$ by integrating over the coordinates and spins of the $(N-1)$ other electrons, i.e.,

$$\begin{aligned}n_\sigma(\mathbf{r})&=\frac{1}{(N-1)!}\sum_{\sigma_2\ldots\sigma_N}\int\mathrm{d}^3r_2\ldots\int\mathrm{d}^3r_N N!|\Psi(\mathbf{r}\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2\\&=N\sum_{\sigma_2\ldots\sigma_N}\int\mathrm{d}^3r_2\ldots\int\mathrm{d}^3r_N|\Psi(\mathbf{r}\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2 \ .\end{aligned} \tag{20}$$

Equations (19) and (20) yield

$$\sum_\sigma\int\mathrm{d}^3r n_\sigma(\mathbf{r})=N \ . \tag{21}$$

Based on the probability interpretation of $n_\sigma(\mathbf{r})$, we might have expected the right-hand side of Eq. (21) to be 1, but that is wrong; the sum of probabilities of all mutually-exclusive events equals 1, but finding an electron at $\mathbf{r}$ does *not* exclude the possibility of finding one at $\mathbf{r}'$, except in a one-electron system. Eq. (21) shows that $n_\sigma(\mathbf{r})\mathrm{d}^3r$ is the average number of electrons of spin σ in volume element d^3r. Moreover, the expectation value of the external potential is

$$\langle\hat{V}_{\mathrm{ext}}\rangle=\langle\Psi|\sum_{i=1}^N v(\mathbf{r}_i)|\Psi\rangle=\int\mathrm{d}^3r n(\mathbf{r})v(\mathbf{r}) \ , \tag{22}$$

with the electron density $n(\mathbf{r})$ given by Eq. (4).

2.2 Wavefunctions for Non-Interacting Electrons

As an important special case, consider the Hamiltonian for N non-interacting electrons:

$$\hat{H}_{\rm non} = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 + v(\mathbf{r}_i) \right) \ . \tag{23}$$

The eigenfunctions of the one-electron problem of Eqs. (13) and (14) are spin orbitals which can be used to construct the antisymmetric eigenfunctions Φ of $\hat{H}_{non}$:

$$\hat{H}_{\rm non}\Phi = E_{\rm non}\Phi \ . \tag{24}$$

Let i stand for $\mathbf{r}_i, \sigma_i$ and construct the Slater determinant or antisymmetrized product [2]

$$\Phi = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \psi_{\alpha_1}(P1)\psi_{\alpha_2}(P2)\ldots\psi_{\alpha_N}(PN) \ , \tag{25}$$

where the quantum label α_i now includes the spin quantum number σ. Here P is any permutation of the labels $1, 2, \ldots, N$, and $(-1)^P$ equals $+1$ for an even permutation and -1 for an odd permutation. The total energy is

$$E_{\rm non} = \varepsilon_{\alpha_1} + \varepsilon_{\alpha_2} + \ldots + \varepsilon_{\alpha_N} \ , \tag{26}$$

and the density is given by the sum of $|\psi_{\alpha_i}(\mathbf{r})|^2$. If any α_i equals any α_j in Eq. (25), we find $\Phi = 0$, which is not a normalizable wavefunction. This is the Pauli exclusion principle : two or more non-interacting electrons may not occupy the same spin orbital.

As an example, consider the ground state for the non-interacting helium atom ($N = 2$). The occupied spin orbitals are

$$\psi_1(\mathbf{r}, \sigma) = \psi_{1s}(\mathbf{r})\delta_{\sigma,\uparrow} \ , \tag{27}$$

$$\psi_2(\mathbf{r}, \sigma) = \psi_{1s}(\mathbf{r})\delta_{\sigma,\downarrow} \ , \tag{28}$$

and the 2-electron Slater determinant is

$$\begin{aligned} \Phi(1,2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(\mathbf{r}_1, \sigma_1) & \psi_2(\mathbf{r}_1, \sigma_1) \\ \psi_1(\mathbf{r}_2, \sigma_2) & \psi_2(\mathbf{r}_2, \sigma_2) \end{vmatrix} \\ &= \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)\frac{1}{\sqrt{2}} \left(\delta_{\sigma_1,\uparrow}\delta_{\sigma_2,\downarrow} - \delta_{\sigma_2,\uparrow}\delta_{\sigma_1,\downarrow} \right) \ , \end{aligned} \tag{29}$$

which is symmetric in space but antisymmetric in spin (whence the total spin is $S = 0$).

If several different Slater determinants yield the same non-interacting energy $E_{\rm non}$, then a linear combination of them will be another antisymmetric eigenstate of $\hat{H}_{\rm non}$. More generally, the Slater-determinant eigenstates of $\hat{H}_{\rm non}$ define a complete orthonormal basis for expansion of the antisymmetric eigenstates of $\hat{H}$, the interacting Hamiltonian of Eq. (16).

2.3 Wavefunction Variational Principle

The Schrödinger equation (17) is equivalent to a wavefunction variational principle [2]: Extremize $\langle\Psi|\hat{H}|\Psi\rangle$ subject to the constraint $\langle\Psi|\Psi\rangle = 1$, i.e., set the following first variation to zero:

$$\delta \left\{ \langle\Psi|\hat{H}|\Psi\rangle / \langle\Psi|\Psi\rangle \right\} = 0 \ . \tag{30}$$

The ground state energy and wavefunction are found by minimizing the expression in curly brackets.

The Rayleigh-Ritz method finds the extrema or the minimum in a *restricted* space of wavefunctions. For example, the Hartree-Fock approximation to the ground-state wavefunction is the single Slater determinant Φ that minimizes $\langle\Phi|\hat{H}|\Phi\rangle/\langle\Phi|\Phi\rangle$. The configuration-interaction ground-state wavefunction [23] is an energy-minimizing linear combination of Slater determinants, restricted to certain kinds of excitations out of a reference determinant. The Quantum Monte Carlo method typically employs a trial wavefunction which is a single Slater determinant times a Jastrow pair-correlation factor [24]. Those widely-used many-electron wavefunction methods are both approximate and computationally demanding, especially for large systems where density functional methods are distinctly more efficient.

The unrestricted solution of Eq. (30) is equivalent by the method of Lagrange multipliers to the unconstrained solution of

$$\delta \left\{ \langle\Psi|\hat{H}|\Psi\rangle - E\langle\Psi|\Psi\rangle \right\} = 0 \ , \tag{31}$$

i.e.,

$$\langle\delta\Psi|(\hat{H} - E)|\Psi\rangle = 0 \ . \tag{32}$$

Since $\delta\Psi$ is an arbitrary variation, we recover the Schrödinger equation (17). Every eigenstate of $\hat{H}$ is an extremum of $\langle\Psi|\hat{H}|\Psi\rangle/\langle\Psi|\Psi\rangle$ and vice versa.

The wavefunction variational principle implies the Hellmann-Feynman and virial theorems below and also implies the Hohenberg-Kohn [25] density functional variational principle to be presented later.

2.4 Hellmann-Feynman Theorem

Often the Hamiltonian $\hat{H}_\lambda$ depends upon a parameter λ, and we want to know how the energy E_λ depends upon this parameter. For any normalized variational solution Ψ_λ (including in particular any eigenstate of $\hat{H}_\lambda$), we define

$$E_\lambda = \langle\Psi_\lambda|\hat{H}_\lambda|\Psi_\lambda\rangle \ . \tag{33}$$

Then

$$\frac{\mathrm{d}E_\lambda}{\mathrm{d}\lambda} = \frac{\mathrm{d}}{\mathrm{d}\lambda'} \langle\Psi_{\lambda'}|\hat{H}_\lambda|\Psi_{\lambda'}\rangle\Big|_{\lambda'=\lambda} + \langle\Psi_\lambda|\frac{\partial\hat{H}_\lambda}{\partial\lambda}|\Psi_\lambda\rangle \ . \tag{34}$$

The first term of Eq. (34) vanishes by the variational principle, and we find the Hellmann-Feynman theorem [26]

$$\frac{\mathrm{d}E_\lambda}{\mathrm{d}\lambda} = \langle\Psi_\lambda|\frac{\partial\hat{H}_\lambda}{\partial\lambda}|\Psi_\lambda\rangle \ . \tag{35}$$

Eq. (35) will be useful later for our understanding of $E_{\rm xc}$. For now, we shall use Eq. (35) to derive the electrostatic force theorem[26]. Let $\mathbf{r}_i$ be the position of the i-th electron, and $\mathbf{R}_I$ the position of the (static) nucleus I with atomic number Z_I. The Hamiltonian

$$\hat{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_i \sum_I \frac{-Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_i \sum_{j\neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_I \sum_{J\neq I} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (36)$$

depends parametrically upon the position $\mathbf{R}_I$, so the force on nucleus I is

$$\begin{aligned} -\frac{\partial E}{\partial \mathbf{R}_I} &= \left\langle \Psi \left| -\frac{\partial \hat{H}}{\partial \mathbf{R}_I} \right| \Psi \right\rangle \\ &= \int \mathrm{d}^3 r\, n(\mathbf{r}) \frac{Z_I(\mathbf{r} - \mathbf{R}_I)}{|\mathbf{r} - \mathbf{R}_I|^3} + \sum_{J\neq I} \frac{Z_I Z_J (\mathbf{R}_I - \mathbf{R}_J)}{|\mathbf{R}_I - \mathbf{R}_J|^3} \ , \end{aligned} \quad (37)$$

just as classical electrostatics would predict. Eq. (37) can be used to find the equilibrium geometries of a molecule or solid by varying all the $\mathbf{R}_I$ until the energy is a minimum and $-\partial E/\partial \mathbf{R}_I = 0$. Eq. (37) also forms the basis for a possible density functional molecular dynamics , in which the nuclei move under these forces by Newton's second law. In principle, all we need for either application is an accurate electron density for each set of nuclear positions.

2.5 Virial Theorem

The density scaling relations to be presented in section 4, which constitute important constraints on the density functionals, are rooted in the same wavefunction scaling that will be used here to derive the virial theorem [26].

Let $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ be any extremum of $\langle \Psi|\hat{H}|\Psi\rangle$ over normalized wavefunctions, i.e., any eigenstate or optimized restricted trial wavefunction (where irrelevant spin variables have been suppressed). For any scale parameter $\gamma > 0$, define the uniformly-scaled wavefunction

$$\Psi_\gamma(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \gamma^{3N/2}\Psi(\gamma\mathbf{r}_1, \ldots, \gamma\mathbf{r}_N) \quad (38)$$

and observe that

$$\langle \Psi_\gamma|\Psi_\gamma\rangle = \langle \Psi|\Psi\rangle = 1 \ . \quad (39)$$

The density corresponding to the scaled wavefunction is the scaled density

$$n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r}) \ , \quad (40)$$

which clearly conserves the electron number:

$$\int \mathrm{d}^3 r\, n_\gamma(\mathbf{r}) = \int \mathrm{d}^3 r\, n(\mathbf{r}) = N \ . \quad (41)$$

$\gamma > 1$ leads to densities $n_\gamma(\mathbf{r})$ that are higher (on average) and more contracted than $n(\mathbf{r})$, while $\gamma < 1$ produces densities that are lower and more expanded.

Now consider what happens to $\langle \hat{H}\rangle = \langle \hat{T} + \hat{V}\rangle$ under scaling. By definition of Ψ,

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\langle \Psi_\gamma|\hat{T} + \hat{V}|\Psi_\gamma\rangle\Big|_{\gamma=1} = 0 \ . \quad (42)$$

But $\hat{T}$ is homogeneous of degree -2 in $\mathbf{r}$, so

$$\langle \Psi_\gamma|\hat{T}|\Psi_\gamma\rangle = \gamma^2\langle \Psi|\hat{T}|\Psi\rangle \ , \quad (43)$$

and Eq. (42) becomes

$$2\langle \Psi|\hat{T}|\Psi\rangle + \frac{\mathrm{d}}{\mathrm{d}\gamma}\langle \Psi_\gamma|\hat{V}|\Psi_\gamma\rangle\Big|_{\gamma=1} = 0 \ , \quad (44)$$

or

$$2\langle \hat{T}\rangle - \langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \frac{\partial \hat{V}}{\partial \mathbf{r}_i}\rangle = 0 \ . \quad (45)$$

If the potential energy $\hat{V}$ is homogeneous of degree n, i.e., if

$$V(\gamma\mathbf{r}_i, \ldots, \gamma\mathbf{r}_N) = \gamma^n V(\mathbf{r}_i, \ldots, \mathbf{r}_N) \ , \quad (46)$$

then

$$\langle \Psi_\gamma|\hat{V}|\Psi_\gamma\rangle = \gamma^{-n}\langle \Psi|\hat{V}|\Psi\rangle \ , \quad (47)$$

and Eq. (44) becomes simply

$$2\langle \Psi|\hat{T}|\Psi\rangle - n\langle \Psi|\hat{V}|\Psi\rangle = 0 \ . \quad (48)$$

For example, $n = -1$ for the Hamiltonian of Eq. (36) in the presence of a single nucleus, or more generally when the Hellmann-Feynman forces of Eq. (37) vanish for the state Ψ.

3 Definitions of Density Functionals

3.1 Introduction to Density Functionals

The many-electron wavefunction $\Psi(\mathbf{r}_1\sigma_1, \ldots, \mathbf{r}_N\sigma_N)$ contains a great deal of information - all we could ever have, but more than we usually want. Because it is a function of many variables, it is not easy to calculate, store, apply or even think about. Often we want no more than the total energy E (and its changes), or perhaps also the spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$, for the ground state. As we shall see, we can formally replace Ψ by the observables $n_\uparrow$ and $n_\downarrow$ as the basic variational objects.

While a *function* is a rule which assigns a number $f(x)$ to a number x, a *functional* is a rule which assigns a number $F[f]$ to a function f. For example, $h[\Psi] = \langle \Psi|\hat{H}|\Psi\rangle$ is a functional of the trial wavefunction Ψ, given the Hamiltonian $\hat{H}$. $U[n]$ of Eq. (9) is a functional of the density $n(\mathbf{r})$, as is the local density approximation for the exchange energy:

$$E_{\rm x}^{\rm LDA}[n] = A_{\rm x}\int \mathrm{d}^3 r\, n(\mathbf{r})^{4/3} \ . \quad (49)$$

The *functional derivative* $\delta F/\delta n(\mathbf{r})$ tells us how the functional $F[n]$ changes under a small variation $\delta n(\mathbf{r})$:

$$\delta F = \int \mathrm{d}^3 r \left(\frac{\delta F}{\delta n(\mathbf{r})}\right)\delta n(\mathbf{r}) \ . \quad (50)$$

For example,

$$\begin{aligned}\delta E_{\mathrm{x}}^{\mathrm{LDA}} &= A_{\mathrm{x}}\int \mathrm{d}^3r \left\{[n(\mathbf{r})+\delta n(\mathbf{r})]^{4/3} - n(\mathbf{r})^{4/3}\right\} \\ &= A_{\mathrm{x}}\int \mathrm{d}^3r \frac{4}{3} n(\mathbf{r})^{1/3}\delta n(\mathbf{r}) \ ,\end{aligned}$$

so

$$\frac{\delta E_{\mathrm{x}}^{\mathrm{LDA}}}{\delta n(\mathbf{r})} = A_{\mathrm{x}}\frac{4}{3}n(\mathbf{r})^{1/3} \ . \tag{51}$$

Similarly,

$$\frac{\delta U[n]}{\delta n(\mathbf{r})} = u([n];\mathbf{r}) \ , \tag{52}$$

where the right hand side is given by Eq. (3). Functional derivatives of various orders can be linked through the translational and rotational symmetries of empty space [27].

3.2 Density Variational Principle

We seek a density functional analog of Eq. (30). Instead of the original derivation of Hohenberg, Kohn and Sham [25, 6] , which was based upon "reductio ad absurdum", we follow the "constrained search" approach of Levy [28], which is in some respects simpler and more constructive.

Eq. (30) tells us that the ground state energy can be found by minimizing $\langle\Psi|\hat{H}|\Psi\rangle$ over all normalized, antisymmetric N-particle wavefunctions:

$$E = \min_{\Psi}\langle\Psi|\hat{H}|\Psi\rangle \ . \tag{53}$$

We now separate the minimization of Eq. (53) into two steps. First we consider all wavefunctions Ψ which yield a given density $n(\mathbf{r})$, and minimize over those wavefunctions:

$$\min_{\Psi\to n}\langle\Psi|\hat{H}|\Psi\rangle = \min_{\Psi\to n}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi\rangle + \int \mathrm{d}^3r\, v(\mathbf{r})n(\mathbf{r}) \ , \tag{54}$$

where we have exploited the fact that all wavefunctions that yield the same $n(\mathbf{r})$ also yield the same $\langle\Psi|\hat{V}_{\mathrm{ext}}|\Psi\rangle$. Then we define the universal functional

$$F[n] = \min_{\Psi\to n}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi\rangle = \langle\Psi_n^{\min}|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi_n^{\min}\rangle \tag{55}$$

where $\Psi_n^{\min}$ is that wavefunction which delivers the minimum for a given n. Finally we minimize over all N-electron densities $n(\mathbf{r})$:

$$\begin{aligned}E &= \min_n E_v[n] \\ &= \min_n \left\{F[n] + \int \mathrm{d}^3r\, v(\mathbf{r})n(\mathbf{r})\right\} \ ,\end{aligned} \tag{56}$$

where of course $v(\mathbf{r})$ is held fixed during the minimization. The minimizing density is then the ground-state density.

The constraint of fixed N can be handled formally through introduction of a Lagrange multiplier μ:

$$\delta\left\{F[n] + \int \mathrm{d}^3r\, v(\mathbf{r})n(\mathbf{r}) - \mu\int \mathrm{d}^3r\, n(\mathbf{r})\right\} = 0 \ , \tag{57}$$

which is equivalent to the Euler equation

$$\frac{\delta F}{\delta n(\mathbf{r})} + v(\mathbf{r}) = \mu \ . \tag{58}$$

μ is to be adjusted until Eq. (5) is satisfied. Eq. (58) shows that the external potential $v(\mathbf{r})$ is uniquely determined by the ground state density (or by any one of them, if the ground state is degenerate).

The functional $F[n]$ is defined via Eq. (55) for all densities $n(\mathbf{r})$ which are "N-representable", i.e., come from an antisymmetric N-electron wavefunction. We shall discuss the extension from wavefunctions to ensembles in section 4.5. The functional derivative $\delta F/\delta n(\mathbf{r})$ is defined via Eq. (58) for all densities which are "v-representable", i.e., come from antisymmetric N-electron ground-state wavefunctions for some choice of external potential $v(\mathbf{r})$.

This formal development requires only the total density of Eq. (4), and not the separate spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$. However, it is clear how to get to a spin-density functional theory: just replace the constraint of fixed n in Eq. (54) and subsequent equations by that of fixed $n_\uparrow$ and $n_\downarrow$. There are two practical reasons to do so: (1) This extension is required when the external potential is spin-dependent, i.e., $v(\mathbf{r}) \to v_\sigma(\mathbf{r})$, as when an external magnetic field couples to the z-component of electron spin. (If this field also couples to the current density $\mathbf{j}(\mathbf{r})$, then we must resort to a current-density functional theory.) (2) Even when $v(\mathbf{r})$ is spin-independent, we may be interested in the physical spin magnetization (e.g., in magnetic materials). (3) Even when neither (1) nor (2) applies, our local and semi-local approximations (Eqs. (11) and (12)) typically work better when we use $n_\uparrow$ and $n_\downarrow$ instead of n.

3.3 Kohn-Sham Non-Interacting System

For a system of non-interacting electrons, $\hat{V}_{\mathrm{ee}}$ of Eq. (16) vanishes so $F[n]$ of Eq. (55) reduces to

$$T_{\mathrm{s}}[n] = \min_{\Psi\to n}\langle\Psi|\hat{T}|\Psi\rangle = \langle\Phi_n^{\min}|\hat{T}|\Phi_n^{\min}\rangle \ . \tag{59}$$

Although we can search over all antisymmetric N-electron wavefunctions in Eq. (59), the minimizing wavefunction $\Phi_n^{\min}$ for a given density will be a non-interacting wavefunction (a single Slater determinant or a linear combination of a few) for some external potential $\hat{V}_{\mathrm{s}}$ such that

$$\frac{\delta T_{\mathrm{s}}}{\delta n(\mathbf{r})} + v_{\mathrm{s}}(\mathbf{r}) = \mu \ , \tag{60}$$

as in Eq. (58). In Eq. (60), the Kohn-Sham potential $v_{\mathrm{s}}(\mathbf{r})$ is a functional of $n(\mathbf{r})$. If there were any difference between μ and μ_{s}, the chemical potentials for interacting and non-interacting systems of the same density, it could be absorbed into $v_{\mathrm{s}}(\mathbf{r})$. We have assumed that $n(\mathbf{r})$ is both interacting and non-interacting v-representable.

Now we *define* the exchange-correlation energy $E_{\mathrm{xc}}[n]$ by

$$F[n] = T_{\mathrm{s}}[n] + U[n] + E_{\mathrm{xc}}[n] \ , \tag{61}$$

where $U[n]$ is given by Eq. (9). The Euler equations (58) and (60) are consistent with one another if and only if

$$v_{\mathrm{s}}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta U[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} \ . \tag{62}$$

Thus we have derived the Kohn-Sham method [6] of section 1.2.

The Kohn-Sham method treats $T_{\mathrm{s}}[n]$ exactly, leaving only $E_{\mathrm{xc}}[n]$ to be approximated. This makes good sense, for several reasons: (1) T_{s} is typically a very large part of the energy, while E_{xc} is a smaller part. (2) T_{s} is largely responsible for density oscillations of the shell structure and Friedel types, which are accurately described by the Kohn-Sham method. (3) E is somewhat better suited to the local and semi-local approximations than is $T_{\mathrm{s}}[n]$, for reasons to be discussed later. The price to be paid for these benefits is the appearance of orbitals. If we had a very accurate approximation for T_{s} directly in terms of n, we could dispense with the orbitals and solve the Euler equation (60) directly for $n(\mathbf{r})$.

The total energy of Eq. (6) may also be written as

$$E = \sum_{\alpha\sigma} \theta(\mu - \varepsilon_{\alpha\sigma})\varepsilon_{\alpha\sigma} - U[n] - \int \mathrm{d}^3 r\, n(\mathbf{r}) v_{\mathrm{xc}}([n];\mathbf{r}) + E_{\mathrm{xc}}[n] \ , \tag{63}$$

where the second and third terms on the right-hand-side simply remove contributions to the first term which do not belong in the total energy. The first term on the right of Eq. (63), the non-interacting energy E_{non}, is the only term that appears in the semi-empirical Hückel theory [26]. This first term includes most of the electronic shell structure effects which arise when $T_{\mathrm{s}}[n]$ is treated exactly (but not when $T_{\mathrm{s}}[n]$ is treated in a continuum model like the Thomas-Fermi approximation or the gradient expansion).

3.4 Exchange Energy and Correlation Energy

$E_{\mathrm{xc}}[n]$ is the sum of distinct exchange and correlation terms:

$$E_{\mathrm{xc}}[n] = E_{\mathrm{x}}[n] + E_{\mathrm{c}}[n] \ , \tag{64}$$

where [29]

$$E_{\mathrm{x}}[n] = \langle \Phi_n^{\mathrm{min}} | \hat{V}_{\mathrm{ee}} | \Phi_n^{\mathrm{min}} \rangle - U[n] \ . \tag{65}$$

When Φ_n^{min} is a single Slater determinant, Eq. (65) is just the usual Fock integral applied to the Kohn-Sham orbitals, i.e., it differs from the Hartree-Fock exchange energy only to the extent that the Kohn-Sham orbitals differ from the Hartree-Fock orbitals for a given system or density (in the same way that $T_s[n]$ differs from the Hartree-Fock kinetic energy). We note that

$$\langle \Phi_n^{\mathrm{min}} | \hat{T} + \hat{V}_{\mathrm{ee}} | \Phi_n^{\mathrm{min}} \rangle = T_{\mathrm{s}}[n] + U[n] + E_{\mathrm{x}}[n] \ , \tag{66}$$

and that, in the one-electron ($\hat{V}_{\mathrm{ee}} = 0$) limit [9],

$$E_{\mathrm{x}}[n] = -U[n] \qquad (N=1) \ . \tag{67}$$

The correlation energy is

$$\begin{aligned} E_{\mathrm{c}}[n] &= F[n] - \{T_{\mathrm{s}}[n] + U[n] + E_{\mathrm{x}}[n]\} \\ &= \langle \Psi_n^{\mathrm{min}} | \hat{T} + \hat{V}_{\mathrm{ee}} | \Psi_n^{\mathrm{min}} \rangle - \langle \Phi_n^{\mathrm{min}} | \hat{T} + \hat{V}_{\mathrm{ee}} | \Phi_n^{\mathrm{min}} \rangle \ . \end{aligned} \tag{68}$$

Since Ψ_n^{min} is that wavefunction which yields density n and minimizes $\langle \hat{T} + \hat{V}_{\mathrm{ee}} \rangle$, Eq. (68) shows that

$$E_{\mathrm{c}}[n] \leq 0 \ . \tag{69}$$

Since Φ_n^{min} is that wavefunction which yields density n and minimizes $\langle \hat{T} \rangle$, Eq. (68) shows that $E_{\mathrm{c}}[n]$ is the sum of a positive kinetic energy piece and a negative potential energy piece. These pieces of E_c contribute respectively to the first and second terms of the virial theorem, Eq. (45). Clearly for any one-electron system [9]

$$E_{\mathrm{c}}[n] = 0 \qquad (N=1) \ . \tag{70}$$

Eqs. (67) and (70) show that the exchange-correlation energy of a one-electron system simply cancels the spurious self-interaction $U[n]$. In the same way, the exchange-correlation potential cancels the spurious self-interaction in the Kohn-Sham potential [9]

$$\frac{\delta E_{\mathrm{x}}}{\delta n(\mathbf{r})} = -u([n];\mathbf{r}) \qquad (N=1) \ , \tag{71}$$

$$\frac{\delta E_{\mathrm{c}}}{\delta n(\mathbf{r})} = 0 \qquad (N=1) \ . \tag{72}$$

Thus

$$\lim_{r\to\infty} \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} = -\frac{1}{r} \qquad (N=1) \ . \tag{73}$$

The extension of these one-electron results to spin-density functional theory is straightforward, since a one-electron system is fully spin-polarized.

3.5 Coupling-Constant Integration

The definitions (65) and (68) are formal ones, and do not provide much intuitive or physical insight into the exchange and correlation energies, or much guidance for the approximation of their density functionals. These insights are provided by the coupling-constant integration [30, 31, 32, 33] to be derived below.

Let us define $\Psi_n^{\mathrm{min},\lambda}$ as that normalized, antisymmetric wavefunction which yields density $n(\mathbf{r})$ and minimizes the expectation value of $\hat{T} + \lambda \hat{V}_{\mathrm{ee}}$, where we have introduced a non-negative coupling constant λ. When $\lambda = 1$, $\Psi_n^{\mathrm{min},\lambda}$ is Ψ_n^{min}, the interacting ground-state wavefunction for density n. When $\lambda = 0$, $\Psi_n^{\mathrm{min},\lambda}$ is Φ_n^{min}, the non-interacting or Kohn-Sham wavefunction for density n. Varying λ at fixed $n(\mathbf{r})$ amounts to varying the external potential $v_\lambda(\mathbf{r})$: At $\lambda = 1$, $v_\lambda(\mathbf{r})$ is the true external potential, while at $\lambda = 0$

it is the Kohn-Sham effective potential $v_s(\mathbf{r})$. We normally assume a smooth, "adiabatic connection" between the interacting and non-interacting ground states as λ is reduced from 1 to 0.

Now we write Eqs. (64), (65) and (68) as

$$\begin{aligned} &E_{\rm xc}[n] \\ &= \langle \Psi_n^{\min,\lambda}|\hat{T}+\lambda\hat{V}_{\rm ee}|\Psi_n^{\min,\lambda}\rangle\Big|_{\lambda=1} - \langle \Psi_n^{\min,\lambda}|\hat{T}+\lambda\hat{V}_{\rm ee}|\Psi_n^{\min,\lambda}\rangle\Big|_{\lambda=0} - U[n] \\ &= \int_0^1 \mathrm{d}\lambda \frac{\mathrm{d}}{\mathrm{d}\lambda}\langle \Psi_n^{\min,\lambda}|\hat{T}+\lambda\hat{V}_{\rm ee}|\Psi_n^{\min,\lambda}\rangle - U[n] \ . \end{aligned} \tag{74}$$

The Hellmann-Feynman theorem of section 2.4 allows us to simplify Eq. (74) to

$$E_{\rm xc}[n] = \int_0^1 \mathrm{d}\lambda \langle \Psi_n^{\min,\lambda}|\hat{V}_{\rm ee}|\Psi_n^{\min,\lambda}\rangle - U[n] \ . \tag{75}$$

Eq. (75) "looks like" a potential energy; the kinetic energy contribution to $E_{\rm xc}$ has been subsumed by the coupling-constant integration. We should remember, of course, that only $\lambda = 1$ is real or physical. The Kohn-Sham system at $\lambda = 0$, and all the intermediate values of λ, are convenient mathematical fictions.

To make further progress, we need to know how to evaluate the N-electron expectation value of a sum of one-body operators like $\hat{T}$, or a sum of two-body operators like $\hat{V}_{\rm ee}$. For this purpose, we introduce one-electron (ρ_1) and two-electron (ρ_2) reduced density matrices [34] :

$$\begin{aligned} &\rho_1(\mathbf{r}'\sigma,\mathbf{r}\sigma) \equiv N \sum_{\sigma_2\ldots\sigma_N} \\ &\int \mathrm{d}^3r_2 \ldots \int \mathrm{d}^3r_N \Psi^*(\mathbf{r}'\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)\Psi(\mathbf{r}\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N) \ , \end{aligned} \tag{76}$$

$$\rho_2(\mathbf{r}',\mathbf{r}) \equiv N(N-1)\sum_{\sigma_1\ldots\sigma_N}\int \mathrm{d}^3r_3 \ldots \int \mathrm{d}^3r_N |\Psi(\mathbf{r}'\sigma_1,\mathbf{r}\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2 \ . \tag{77}$$

¿From Eq. (20),

$$n_\sigma(\mathbf{r}) = \rho_1(\mathbf{r}\sigma,\mathbf{r}\sigma) \ . \tag{78}$$

Clearly also

$$\langle \hat{T}\rangle = \sum_\sigma \int \mathrm{d}^3r - \frac{1}{2}\frac{\partial}{\partial\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{r}}\rho_1(\mathbf{r}'\sigma,\mathbf{r}\sigma)\Big|_{\mathbf{r}'=\mathbf{r}} \ , \tag{79}$$

$$\langle \hat{V}_{\rm ee}\rangle = \frac{1}{2}\int \mathrm{d}^3r \int \mathrm{d}^3r' \frac{\rho_2(\mathbf{r}',\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{80}$$

We interpret the positive number $\rho_2(\mathbf{r}',\mathbf{r})\mathrm{d}^3r'\mathrm{d}^3r$ as the joint probability of finding an electron in volume element d^3r' at $\mathbf{r}'$, *and* an electron in d^3r at $\mathbf{r}$. By standard probability theory, this is the product of the probability of finding an electron in d^3r ($n(\mathbf{r})\mathrm{d}^3r$) and the conditional probability of finding an electron in d^3r', given that there is one at $\mathbf{r}$ ($n_2(\mathbf{r},\mathbf{r}')\mathrm{d}^3r'$):

$$\rho_2(\mathbf{r}',\mathbf{r}) = n(\mathbf{r})n_2(\mathbf{r},\mathbf{r}') \ . \tag{81}$$

By arguments similar to those used in section 2.1, we interpret $n_2(\mathbf{r},\mathbf{r}')$ as the average density of electrons at $\mathbf{r}'$, given that there is an electron at $\mathbf{r}$. Clearly then

$$\int \mathrm{d}^3r' n_2(\mathbf{r},\mathbf{r}') = N-1 \ . \tag{82}$$

For the wavefunction $\Psi_n^{\min,\lambda}$, we write

$$n_2(\mathbf{r},\mathbf{r}') = n(\mathbf{r}') + n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r}') \ , \tag{83}$$

an equation which defines $n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r}')$, the density at $\mathbf{r}'$ of the exchange-correlation hole [33] about an electron at $\mathbf{r}$. Eqs. (5) and (83) imply that

$$\int \mathrm{d}^3r' n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r}') = -1 \ , \tag{84}$$

which says that, if an electron is definitely at $\mathbf{r}$, it is missing from the rest of the system.

Because the Coulomb interaction $1/u$ is singular as $u = |\mathbf{r}-\mathbf{r}'| \to 0$, the exchange-correlation hole density has a cusp [35, 34] around $u = 0$:

$$\frac{\partial}{\partial u}\int \frac{\mathrm{d}\Omega_{\mathbf{u}}}{4\pi} n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r}+\mathbf{u})\Big|_{u=0} = \lambda\left[n(\mathbf{r}) + n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r})\right] \ , \tag{85}$$

where $\int \mathrm{d}\Omega_{\mathbf{u}}/(4\pi)$ is an angular average. This cusp vanishes when $\lambda = 0$, and also in the fully-spin-polarized and low-density limits, in which all other electrons are excluded from the position of a given electron: $n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r}) = -n(\mathbf{r})$.

We can now rewrite Eq. (75) as [33]

$$E_{\rm xc}[n] = \frac{1}{2}\int \mathrm{d}^3r \int \mathrm{d}^3r' \frac{n(\mathbf{r})\bar{n}_{\rm xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ , \tag{86}$$

where

$$\bar{n}_{\rm xc}(\mathbf{r},\mathbf{r}') = \int_0^1 \mathrm{d}\lambda n_{\rm xc}^\lambda(\mathbf{r},\mathbf{r}') \tag{87}$$

is the coupling-constant averaged hole density. The exchange-correlation energy is just the electrostatic interaction between each electron and the coupling-constant-averaged exchange-correlation hole which surrounds it. The hole is created by three effects: (1) self-interaction correction, a classical effect which guarantees that an electron cannot interact with itself, (2) the Pauli exclusion principle, which tends to keep two electrons with parallel spins apart in space, and (3) the Coulomb repulsion, which tends to keep any two electrons apart in space. Effects (1) and (2) are responsible for the exchange energy, which is present even at $\lambda = 0$, while effect (3) is responsible for the correlation energy, and arises only for $\lambda \neq 0$.

If $\Psi_n^{\min,\lambda=0}$ is a single Slater determinant, as it typically is, then the one- and two-electron density matrices at $\lambda = 0$ can be constructed explicitly from the Kohn-Sham spin orbitals $\psi_{\alpha\sigma}(\mathbf{r})$:

$$\rho_1^{\lambda=0}(\mathbf{r}'\sigma,\mathbf{r}\sigma) = \sum_\alpha \theta(\mu-\varepsilon_{\alpha\sigma})\psi_{\alpha\sigma}^*(\mathbf{r}')\psi_{\alpha\sigma}(\mathbf{r}) \ , \tag{88}$$

$$\rho_2^{\lambda=0}(\mathbf{r}',\mathbf{r}) = n(\mathbf{r})n(\mathbf{r}') + n(\mathbf{r})n_{\rm x}(\mathbf{r},\mathbf{r}') \ , \tag{89}$$

where

$$n_{\mathrm{x}}(\mathbf{r},\mathbf{r}') = n_{\mathrm{xc}}^{\lambda=0}(\mathbf{r},\mathbf{r}') = -\sum_{\sigma} \frac{|\rho_1^{\lambda=0}(\mathbf{r}'\sigma,\mathbf{r}\sigma)|^2}{n(\mathbf{r})} \tag{90}$$

is the exact exchange-hole density. Eq. (90) shows that

$$n_{\mathrm{x}}(\mathbf{r},\mathbf{r}') \leq 0 \ , \tag{91}$$

so the exact exchange energy

$$E_{\mathrm{x}}[n] = \frac{1}{2}\int \mathrm{d}^3r \int \mathrm{d}^3r' \frac{n(\mathbf{r})n_{\mathrm{x}}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{92}$$

is also negative, and can be written as the sum of up-spin and down-spin contributions:

$$E_{\mathrm{x}} = E_{\mathrm{x}}^{\uparrow} + E_{\mathrm{x}}^{\downarrow} < 0 \ . \tag{93}$$

Eq. (84) provides a sum rule for the exchange hole:

$$\int \mathrm{d}^3r' n_{\mathrm{x}}(\mathbf{r},\mathbf{r}') = -1 \ . \tag{94}$$

Eqs. (90) and (78) show that the "on-top" exchange hole density is [36]

$$n_{\mathrm{x}}(\mathbf{r},\mathbf{r}) = -\frac{n_{\uparrow}^2(\mathbf{r}) + n_{\downarrow}^2(\mathbf{r})}{n(\mathbf{r})} \ , \tag{95}$$

which is determined by just the local spin densities at position $\mathbf{r}$ - suggesting a reason why local spin density approximations work better than local density approximations.

The correlation hole density is defined by

$$\bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = n_{\mathrm{x}}(\mathbf{r},\mathbf{r}') + \bar{n}_{\mathrm{c}}(\mathbf{r},\mathbf{r}') \ , \tag{96}$$

and satisfies the sum rule

$$\int \mathrm{d}^3r' \bar{n}_{\mathrm{c}}(\mathbf{r},\mathbf{r}') = 0 \ , \tag{97}$$

which says that Coulomb repulsion changes the shape of the hole but not its integral. In fact, this repulsion typically makes the hole deeper but more short-ranged, with a negative on-top correlation hole density:

$$\bar{n}_{\mathrm{c}}(\mathbf{r},\mathbf{r}) \leq 0 \ . \tag{98}$$

The positivity of Eq. (77) is equivalent via Eqs. (81) and (83) to the inequality

$$\bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') \geq -n(\mathbf{r}') \ , \tag{99}$$

which asserts that the hole cannot take away electrons that weren't there initially. By the sum rule (97), the correlation hole density $\bar{n}_c(\mathbf{r},\mathbf{r}')$ must have positive as well as negative contributions. Moreover, unlike the exchange hole density $n_x(\mathbf{r},\mathbf{r}')$, the exchange-correlation hole density $\bar{n}_{xc}(\mathbf{r},\mathbf{r}')$ can be positive.

To better understand E_{xc}, we can simplify Eq. (86) to the "real-space analysis" [37]

$$E_{\mathrm{xc}}[n] = \frac{N}{2}\int_0^{\infty} \mathrm{d}u \ 4\pi u^2 \frac{\langle \bar{n}_{\mathrm{xc}}(u)\rangle}{u} \ , \tag{100}$$

where

$$\langle \bar{n}_{\mathrm{xc}}(u)\rangle = \frac{1}{N}\int \mathrm{d}^3r\, n(\mathbf{r}) \int \frac{\mathrm{d}\Omega_{\mathbf{u}}}{4\pi} \bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}+\mathbf{u}) \tag{101}$$

is the system- and spherical-average of the coupling-constant-averaged hole density. The sum rule of Eq. (84) becomes

$$\int_0^{\infty} \mathrm{d}u \ 4\pi u^2 \langle \bar{n}_{\mathrm{xc}}(u)\rangle = -1 \ . \tag{102}$$

As u increases from 0, $\langle n_{\mathrm{x}}(u)\rangle$ rises analytically like $\langle n_{\mathrm{x}}(0)\rangle + \mathcal{O}(u^2)$, while $\langle \bar{n}_{\mathrm{c}}(u)\rangle$ rises like $\langle \bar{n}_{\mathrm{c}}(0)\rangle + \mathcal{O}(|u|)$ as a consequence of the cusp of Eq. (85). Because of the constraint of Eq. (102) and because of the factor $1/u$ in Eq. (100), E_{xc} typically becomes more negative as the on-top hole density $\langle \bar{n}_{\mathrm{xc}}(u)\rangle$ gets more negative.

4 Formal Properties of Functionals

4.1 Uniform Coordinate Scaling

The more we know of the exact properties of the density functionals $E_{xc}[n]$ and $T_s[n]$, the better we shall understand and be able to approximate these functionals. We start with the behavior of the functionals under a uniform coordinate scaling of the density, Eq. (40).

The Hartree electrostatic self-repulsion of the electrons is known exactly (Eq. (9)), and has a simple coordinate scaling:

$$\begin{aligned} U[n_{\gamma}] &= \frac{1}{2}\int \mathrm{d}^3(\gamma r) \int \mathrm{d}^3(\gamma r') \frac{n(\gamma\mathbf{r})n(\gamma\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &= \gamma \frac{1}{2}\int \mathrm{d}^3r_1 \int \mathrm{d}^3r_1' \frac{n(\mathbf{r}_1)n(\mathbf{r}_1')}{|\mathbf{r}_1-\mathbf{r}_1'|} = \gamma U[n] \ , \end{aligned} \tag{103}$$

where $\mathbf{r}_1 = \gamma\mathbf{r}$ and $\mathbf{r}_1' = \gamma\mathbf{r}'$.

Next consider the non-interacting kinetic energy of Eq. (59). Scaling all the wavefunctions Ψ in the constrained search as in Eq. (38) will scale the density as in Eq. (40) and scale each kinetic energy expectation value as in Eq. (43). Thus the constrained search for the unscaled density maps into the constrained search for the scaled density, and [38]

$$T_{\mathrm{s}}[n_{\gamma}] = \gamma^2 \, T_{\mathrm{s}}[n] \ . \tag{104}$$

We turn now to the exchange energy of Eq. (65). By the argument of the last paragraph, $\Phi_{n_\gamma}^{\min}$ is the scaled version of $\Phi_n^{\min}$. Since also

$$\hat{V}_{\mathrm{ee}}(\gamma\mathbf{r}_1,\ldots,\gamma\mathbf{r}_N) = \gamma^{-1}\, \hat{V}_{\mathrm{ee}}(\mathbf{r}_1,\ldots,\mathbf{r}_N) \ , \tag{105}$$

and with the help of Eq. (103), we find [38]

$$E_{\mathrm{x}}[n_{\gamma}] = \gamma E_{\mathrm{x}}[n] \ . \tag{106}$$

In the high-density ($\gamma \to \infty$) limit, $T_s[n_\gamma]$ dominates $U[n_\gamma]$ and $E_{\mathrm{x}}[n_\gamma]$. An example would be an ion with a fixed number of electrons N and a nuclear charge Z which tends to infinity; in this limit, the density and energy become essentially hydrogenic, and the effects of U and E_{x} become relatively negligible. In the low-density ($\gamma \to 0$) limit, $U[n_\gamma]$ and $E_{\mathrm{x}}[n_\gamma]$ dominate $T_{\mathrm{s}}[n_\gamma]$.

We can use coordinate scaling relations to fix the form of a local density approximation

$$F[n] = \int \mathrm{d}^3 r f(n(\mathbf{r})) \ . \tag{107}$$

If $F[n_\lambda] = \lambda^p F[n]$, then

$$\lambda^{-3} \int \mathrm{d}^3(\lambda \mathbf{r}) f\left(\lambda^3 n(\lambda \mathbf{r})\right) = \lambda^p \int \mathrm{d}^3 r f(n(\mathbf{r})) \ , \tag{108}$$

or $f(\lambda^3 n) = \lambda^{p+3} f(n)$, whence

$$f(n) = n^{1+p/3} \ . \tag{109}$$

For the exchange energy of Eq. (106), $p = 1$ so Eqs. (107) and (109) imply Eq. (49). For the non-interacting kinetic energy of Eq. (104), $p = 2$ so Eqs. (107) and (109) imply the Thomas-Fermi approximation

$$T_0[n] = A_s \int \mathrm{d}^3 r n^{5/3}(\mathbf{r}) \ . \tag{110}$$

$U[n]$ of Eq. (9) is too strongly nonlocal for any local approximation.

While $T_{\mathrm{s}}[n]$, $U[n]$ and $E_{\mathrm{x}}[n]$ have simple scalings, $E_{\mathrm{c}}[n]$ of Eq. (68) does not. This is because $\Psi^{\min}_{n_\gamma}$, the wavefunction which via Eq. (55) yields the scaled density $n_\gamma(\mathbf{r})$ and minimizes the expectation value of $\hat{T} + \hat{V}_{\mathrm{ee}}$, is *not* the scaled wavefunction $\gamma^{3N/2} \Psi^{\min}_n(\gamma \mathbf{r}_1, \ldots, \gamma \mathbf{r}_N)$. The scaled wavefunction yields $n_\gamma(\mathbf{r})$ but minimizes the expectation value of $\hat{T} + \gamma \hat{V}_{\mathrm{ee}}$, and it is this latter expectation value which scales like γ^2 under wavefunction scaling. Thus [39]

$$E_{\mathrm{c}}[n_\gamma] = \gamma^2 E_{\mathrm{c}}^{1/\gamma}[n] \ , \tag{111}$$

where $E_{\mathrm{c}}^{1/\gamma}[n]$ is the density functional for the correlation energy in a system for which the electron-electron interaction is not $\hat{V}_{\mathrm{ee}}$ but $\gamma^{-1} \hat{V}_{\mathrm{ee}}$.

To understand these results, let us assume that the Kohn-Sham non-interacting Hamiltonian has a non-degenerate ground state. In the high-density limit ($\gamma \to \infty$), $\Psi^{\min}_{n_\gamma}$ minimizes just $\langle \hat{T} \rangle$ and reduces to $\Phi^{\min}_{n_\gamma}$. Now we treat

$$\Delta \equiv \hat{V}_{\mathrm{ee}} - \sum_{i=1}^{N} \left(\frac{\delta U}{\delta n(\mathbf{r}_i)} + \frac{\delta E_x}{\delta n(\mathbf{r}_i)} \right) \tag{112}$$

as a weak perturbation [40, 41] on the Kohn-Sham non-interacting Hamiltonian, and find

$$E_{\mathrm{c}} = \sum_{n \neq 0} \frac{|\langle n | \Delta | 0 \rangle|^2}{E_0 - E_n} \ , \tag{113}$$

where the $|n\rangle$ are the eigenfunctions of the Kohn-Sham non-interacting Hamiltonian, and $|0\rangle$ is its ground state. Both the numerator and the denominator of Eq. (113) scale like γ^2, so [42]

$$\lim_{\gamma \to \infty} E_{\mathrm{c}}[n_\gamma] = constant \ . \tag{114}$$

In the low-density limit, $\Psi^{\min}_{n_\gamma}$ minimizes just $\langle \hat{V}_{\mathrm{ee}} \rangle$, and Eq. (68) then shows that [43]

$$E_{\mathrm{c}}[n_\gamma] \approx \gamma D[n] \qquad (\gamma \to 0) \ . \tag{115}$$

Generally, we have a scaling inequality [38]

$$E_{\mathrm{c}}[n_\gamma] > \gamma E_{\mathrm{c}}[n] \qquad (\gamma > 1) \ , \tag{116}$$

$$E_{\mathrm{c}}[n_\gamma] < \gamma E_{\mathrm{c}}[n] \qquad (\gamma < 1) \ . \tag{117}$$

If we choose a density n, we can plot $E_{\mathrm{c}}[n_\gamma]$ versus γ, and compare the result to the straight line $\gamma E_{\mathrm{c}}[n]$. These two curves will drop away from zero as γ increases from zero (with different initial slopes), then cross at $\gamma = 1$. The convex $E_{\mathrm{c}}[n_\gamma]$ will then approach a negative constant as $\gamma \to \infty$.

4.2 Local Lower Bounds

Because of the importance of local and semilocal approximations like Eqs. (11) and (12), bounds on the exact functionals are especially useful when the bounds are themselves local functionals.

Lieb and Thirring [44] have conjectured that $T_{\mathrm{s}}[n]$ is bounded from below by the Thomas-Fermi functional

$$T_{\mathrm{s}}[n] \geq T_0[n] \ , \tag{118}$$

where $T_0[n]$ is given by Eqs. (110) with

$$A_{\mathrm{s}} = \frac{3}{10} (3\pi^2)^{2/3} \ . \tag{119}$$

We have already established that

$$E_{\mathrm{x}}[n] \geq E_{\mathrm{xc}}[n] \geq E_{\mathrm{xc}}^{\lambda=1}[n] \ , \tag{120}$$

where the final term of Eq. (120) is the integrand $E_{\mathrm{xc}}^{\lambda}[n]$ of the coupling-constant integration of Eq. (75) ,

$$E_{\mathrm{xc}}^{\lambda}[n] = \langle \Psi_n^{\min,\lambda} | \hat{V}_{\mathrm{ee}} | \Psi_n^{\min,\lambda} \rangle - U[n] \ , \tag{121}$$

evaluated at the upper limit $\lambda = 1$. Lieb and Oxford [45] have proved that

$$E_{\mathrm{xc}}^{\lambda=1}[n] \geq 2.273 E_{\mathrm{x}}^{\mathrm{LDA}}[n] \ , \tag{122}$$

where $E_{\mathrm{x}}^{\mathrm{LDA}}[n]$ is the local density approximation for the exchange energy, Eq. (49), with

$$A_{\mathrm{x}} = -\frac{3}{4\pi} (3\pi^2)^{1/3} \ . \tag{123}$$

4.3 Spin Scaling Relations

Spin scaling relations can be used to convert density functionals into spin-density functionals.

For example, the non-interacting kinetic energy is the sum of the separate kinetic energies of the spin-up and spin-down electrons:

$$T_s[n_\uparrow, n_\downarrow] = T_s[n_\uparrow, 0] + T_s[0, n_\downarrow] \ . \tag{124}$$

The corresponding density functional, appropriate to a spin-unpolarized system, is [46]

$$T_s[n] = T_s[n/2, n/2] = 2T_s[n/2, 0] \ , \tag{125}$$

whence $T_s[n/2, 0] = \frac{1}{2}T_s[n]$ and Eq. (124) becomes

$$T_s[n_\uparrow, n_\downarrow] = \frac{1}{2}T_s[2n_\uparrow] + \frac{1}{2}T_s[2n_\downarrow] \ . \tag{126}$$

Similarly, Eq. (93) implies [46]

$$E_x[n_\uparrow, n_\downarrow] = \frac{1}{2}E_x[2n_\uparrow] + \frac{1}{2}E_x[2n_\downarrow] \ . \tag{127}$$

For example, we can start with the local density approximations (110) and (49), then apply (126) and (127) to generate the corresponding local spin density approximations.

Because two electrons of anti-parallel spin repel one another Coulombically, making an important contribution to the correlation energy, there is no simple spin scaling relation for E_c.

4.4 Size Consistency

Common sense tells us that the total energy E and density $n(\mathbf{r})$ for a system, comprised of two well-separated subsystems with energies E_1 and E_2 and densities $n_1(\mathbf{r})$ and $n_2(\mathbf{r})$, must be $E = E_1 + E_2$ and $n(\mathbf{r}) = n_1(\mathbf{r}) + n_2(\mathbf{r})$. Approximations which satisfy this expectation, such as the LSD of Eq. (11) or the GGA of Eq. (12), are properly size consistent [47]. Size consistency is not only a principle of physics, it is almost a principle of epistemology: How could we analyze or understand complex systems, if they could not be separated into simpler components?

Density functionals which are *not* size consistent are to be avoided. An example is the Fermi-Amaldi [48] approximation for the exchange energy,

$$E_x^{\mathrm{FA}}[n] = -U[n/N] \ , \tag{128}$$

where N is given by Eq. (5), which was constructed to satisfy Eq. (67).

4.5 Derivative Discontinuity

In section 3, our density functionals were defined as constrained searches over wavefunctions. Because all wavefunctions searched have the same electron number, there is no way to make a number-nonconserving density variation $\delta n(\mathbf{r})$. The functional derivatives are defined only up to an arbitrary constant, which has no effect on Eq. (50) when $\int d^3r\,\delta n(\mathbf{r}) = 0$.

To complete the definition of the functional derivatives and of the chemical potential μ, we extend the constrained search from wavefunctions to ensembles [49, 50] . An ensemble or mixed state is a set of wavefunctions or pure states and their respective probabilities. By including wavefunctions with different electron numbers in the same ensemble, we can develop a density functional theory for non-integer particle number . Fractional particle numbers can arise in an open system that shares electrons with its environment, and in which the electron number fluctuates between integers.

The upshot is that the ground-state energy $E(N)$ varies linearly between two adjacent integers, and has a derivative discontinuity at each integer. This discontinuity arises in part from the exchange-correlation energy (and entirely so in cases for which the integer does not fall on the boundary of an electronic shell or subshell, e.g., for $N = 6$ in the carbon atom but not for $N = 10$ in the neon atom).

By Janak's theorem [51], the highest partly-occupied Kohn-Sham eigenvalue $\varepsilon_{\mathrm{HO}}$ equals $\partial E/\partial N = \mu$, and so changes discontinuously [49, 50] at an integer Z:

$$\varepsilon_{\mathrm{HO}} = \begin{cases} -I_Z & (Z-1 < N < Z) \\ -A_Z & (Z < N < Z+1) \end{cases} \ , \tag{129}$$

where I_Z is the first ionization energy of the Z-electron system (i.e., the least energy needed to remove an electron from this system), and A_Z is the electron affinity of the Z-electron system (i.e., $A_Z = I_{Z+1}$). If Z does not fall on the boundary of an electronic shell or subshell, all of the difference between $-I_Z$ and $-A_Z$ must arise from a discontinous jump in the exchange-correlation potential $\delta E_{\mathrm{xc}}/\delta n(\mathbf{r})$ as the electron number N crosses the integer Z.

Since the asymptotic decay of the density of a finite system with Z electrons is controlled by I_Z, we can show that the exchange-correlation potential tends to zero as $|\mathbf{r}| \to \infty$ [52]:

$$\lim_{|\mathbf{r}|\to\infty} \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} = 0 \quad (Z-1 < N < Z) \ , \tag{130}$$

or more precisely

$$\lim_{|\mathbf{r}|\to\infty} \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} = -\frac{1}{r} \quad (Z-1 < N < Z) \ . \tag{131}$$

As N increases through the integer Z, $\delta E_{\mathrm{xc}}/\delta n(\mathbf{r})$ jumps up by a positive additive constant. With further increases in N above Z, this "constant" vanishes, first at very large $|\mathbf{r}|$ and then at smaller and smaller $|\mathbf{r}|$, until it is all gone in the limit where N approaches the integer $Z + 1$ from below.

Simple continuum approximations to $E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]$, such as the LSD of Eq. (11) or the GGA of Eq. (12), miss much or all the derivative discontinuity, and can at best average over it. For example, the highest occupied orbital energy for a neutral atom becomes approximately $-\frac{1}{2}(I_Z + A_Z)$, the average of Eq. (129) from the electron-deficient and electron-rich sides of neutrality. We must never forget, when we make these approximations, that we are fitting a round peg into a square hole. The areas (integrated properties) of a circle and a square can be matched, but their perimeters (differential properties) will remain stubbornly different.

5 Uniform Electron Gas

5.1 Kinetic Energy

Simple systems play an important paradigmatic role in science. For example, the hydrogen atom is a paradigm for all of atomic physics. In the same way, the uniform electron gas [24] is a paradigm for solid-state physics, and also for density functional theory. In this system, the electron density $n(\mathbf{r})$ is uniform or constant over space, and thus the electron number is infinite. The negative charge of the electrons is neutralized by a rigid uniform positive background. We could imagine creating such a system by starting with a simple metal, regarded as a perfect crystal of valence electrons and ions, and then smearing out the ions to make the uniform background of positive charge. In fact, the simple metal sodium is physically very much like a uniform electron gas.

We begin by evaluating the non-interacting kinetic energy (this section) and exchange energy (next section) per electron for a spin-unpolarized electron gas of uniform density n. The corresponding energies for the spin-polarized case can then be found from Eqs. (126) and (127).

By symmetry, the Kohn-Sham potential $v_s(\mathbf{r})$ must be uniform or constant, and we take it to be zero. We impose boundary conditions within a cube of volume $\mathcal{V} \to \infty$, i.e., we require that the orbitals repeat from one face of the cube to its opposite face. (Presumably any choice of boundary conditions would give the same answer as $\mathcal{V} \to \infty$.) The Kohn-Sham orbitals are then plane waves $\exp(i\mathbf{k}\cdot\mathbf{r})/\sqrt{\mathcal{V}}$, with momenta or wavevectors $\mathbf{k}$ and energies $k^2/2$. The number of orbitals of both spins in a volume d^3k of wavevector space is $2[\mathcal{V}/(2\pi)^3]\mathrm{d}^3k$, by an elementary geometrical argument [53].

Let $N = n\mathcal{V}$ be the number of electrons in volume $\mathcal{V}$. These electrons occupy the N lowest Kohn-Sham spin orbitals, i.e., those with $k < k_{\mathrm{F}}$:

$$N = 2\sum_{\mathbf{k}} \theta(k_{\mathrm{F}} - k) = 2\frac{\mathcal{V}}{(2\pi)^3}\int_0^{k_{\mathrm{F}}} \mathrm{d}k\; 4\pi k^2 = \mathcal{V}\frac{k_{\mathrm{F}}^3}{3\pi^2} \;, \tag{132}$$

where k_{F} is called the Fermi wavevector. The Fermi wavelength $2\pi/k_{\mathrm{F}}$ is the shortest de Broglie wavelength for the non-interacting electrons. Clearly

$$n = \frac{k_{\mathrm{F}}^3}{3\pi^2} = \frac{3}{4\pi r_{\mathrm{s}}^3} \;, \tag{133}$$

where we have introduced the Seitz radius r_{s} - the radius of a sphere which on average contains one electron.

The kinetic energy of an orbital is $k^2/2$, and the average kinetic energy per electron is

$$t_{\mathrm{s}}(n) = \frac{2}{N}\sum_{\mathbf{k}} \theta(k_{\mathrm{F}} - k)\frac{k^2}{2} = \frac{2\mathcal{V}}{N(2\pi)^3}\int_0^{k_{\mathrm{F}}} \mathrm{d}k\; 4\pi k^2 \frac{k^2}{2} = \frac{3}{5}\frac{k_{\mathrm{F}}^2}{2} \;, \tag{134}$$

or 3/5 of the Fermi energy. In other notation,

$$t_{\mathrm{s}}(n) = \frac{3}{10}(3\pi^2 n)^{2/3} = \frac{3}{10}\frac{(9\pi/4)^{2/3}}{r_{\mathrm{s}}^2} \;. \tag{135}$$

All of this kinetic energy follows from the Pauli exclusion principle, i.e., from the fermion character of the electron.

5.2 Exchange Energy

To evaluate the exchange energy, we need the Kohn-Sham one-matrix for electrons of spin σ, as defined in Eq. (88):

$$\begin{aligned}
\rho_1^{\lambda=0}(\mathbf{r}+\mathbf{u}\sigma, \mathbf{r}\sigma) &= \sum_{\mathbf{k}} \theta(k_{\mathrm{F}} - k)\frac{\exp(-i\mathbf{k}\cdot(\mathbf{r}+\mathbf{u}))}{\sqrt{\mathcal{V}}}\frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{\sqrt{\mathcal{V}}} \\
&= \frac{1}{(2\pi)^3}\int_0^{k_{\mathrm{F}}} \mathrm{d}k\; 4\pi k^2 \int \frac{\mathrm{d}\Omega_{\mathbf{k}}}{4\pi}\exp(-i\mathbf{k}\cdot\mathbf{u}) \\
&= \frac{1}{2\pi^2}\int_0^{k_{\mathrm{F}}} \mathrm{d}k\; k^2 \frac{\sin(ku)}{ku} \\
&= \frac{k_{\mathrm{F}}^3}{2\pi^2}\frac{\sin(k_{\mathrm{F}}u) - k_{\mathrm{F}}u\cos(k_{\mathrm{F}}u)}{(k_{\mathrm{F}}u)^3} \;.
\end{aligned} \tag{136}$$

The exchange hole density at distance u from an electron is, by Eq. (90),

$$n_{\mathrm{x}}(u) = -2\frac{|\rho_1^{\lambda=0}(\mathbf{r}+\mathbf{u}\sigma, \mathbf{r}\sigma)|^2}{n} \;, \tag{137}$$

which ranges from $-n/2$ at $u = 0$ (where all other electrons of the same spin are excluded by the Pauli principle) to 0 (like $1/u^4$) as $u \to \infty$. The exchange energy per electron is

$$e_{\mathrm{x}}(n) = \int_0^\infty \mathrm{d}u\; 2\pi u n_{\mathrm{x}}(u) = -\frac{3}{4\pi}k_{\mathrm{F}} \;. \tag{138}$$

In other notation,

$$e_{\mathrm{x}}(n) = -\frac{3}{4\pi}(3\pi^2 n)^{1/3} = -\frac{3}{4\pi}\frac{(9\pi/4)^{1/3}}{r_{\mathrm{s}}} \;. \tag{139}$$

Since the self-interaction correction vanishes for the diffuse orbitals of the uniform gas, all of this exchange energy is due to the Pauli exclusion principle.

5.3 Correlation Energy

Exact analytic expressions for $e_c(n)$, the correlation energy per electron of the uniform gas, are known only in extreme limits. The high-density ($r_s \to 0$) limit is also the weak-coupling limit, in which

$$e_c(n) = c_0 \ln r_s - c_1 + c_2 r_s \ln r_s - c_3 r_s + \ldots \quad (r_s \to 0) \tag{140}$$

from many-body perturbation theory [54]. The positive constants $c_0 = 0.031091$ [54] and $c_1 = 0.046644$ [55] are known. Eq. (140) does not quite tend to a constant when $r_s \to 0$, as Eq. (114) would suggest, because the excited states of the non-interacting system lie arbitrarily close in energy to the ground state.

The low-density ($r_s \to \infty$) limit is also the strong coupling limit in which the uniform fluid phase is unstable against the formation of a close-packed Wigner lattice of localized electrons. Because the energies of these two phases remain nearly degenerate as $r_s \to \infty$, they have the same kind of dependence upon r_s [56]:

$$e_c(n) \to -\frac{d_0}{r_s} + \frac{d_1}{r_s^{3/2}} + \ldots \quad (r_s \to \infty) \ . \tag{141}$$

The constants d_0 and d_1 in Eq. (141) can be estimated from the Madelung electrostatic and zero-point vibrational energies of the Wigner crystal, respectively. The estimate

$$d_0 \approx -\frac{9}{10} \tag{142}$$

can be found from the electrostatic energy of a neutral spherical cell: just add the electrostatic self-repulsion $3/5r_s$ of a sphere of uniform positive background (with radius r_s) to the interaction $-3/2r_s$ between this background and the electron at its center. The origin of the $r_s^{-3/2}$ term in Eq. (141) is also simple: Think of the potential energy of the electron at small distance u from the center of the sphere as $-3/2r_s + \frac{1}{2}ku^2$, where k is a spring constant. Since this potential energy must vanish for $u \approx r_s$, we find that $k \sim r_s^{-3}$ and thus the zero-point vibrational energy is $3\omega/2 = 1.5\sqrt{k/m} \sim r_s^{-3/2}$.

An expression which encompasses both limits (140)and (141) is [8]

$$e_c(n) = -2c_0(1+\alpha_1 r_s) \ln\left[1 + \frac{1}{2c_0(\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_4 r_s^2)}\right] , \tag{143}$$

where

$$\beta_1 = \frac{1}{2c_0} \exp\left(-\frac{c_1}{2c_0}\right) , \tag{144}$$

$$\beta_2 = 2c_0 \beta_1^2 \ . \tag{145}$$

The coefficients $\alpha_1 = 0.21370$, $\beta_3 = 1.6382$, and $\beta_4 = 0.49294$ are found by fitting to accurate Quantum Monte Carlo correlation energies [57] for r_s =2, 5, 10, 20, 50, and 100.

The uniform electron gas is in equilibrium when the density n minimizes the total energy per electron, i.e., when

$$\frac{\partial}{\partial n}[t_s(n) + e_x(n) + e_c(n)] = 0 \ . \tag{146}$$

This condition is met at $r_s = 4.1$, close to the observed valence electron density of sodium. At any r_s, we have

$$\frac{\delta T_s}{\delta n(\mathbf{r})} = \frac{\partial}{\partial n}[n t_s(n)] = \frac{1}{2}k_F^2 , \tag{147}$$

$$\frac{\delta E_x}{\delta n(\mathbf{r})} = \frac{\partial}{\partial n}[n e_x(n)] = -\frac{1}{\pi}k_F \ . \tag{148}$$

Eq. (143) with the parameters listed above provides a representation of $e_c(n_\uparrow, n_\downarrow)$ for $n_\uparrow = n_\downarrow = n/2$; other accurate representations are also available [9, 10]. Eq. (143) with different parameters ($c_0 = 0.015545$, $c_1 = 0.025599$, $\alpha_1 = 0.20548$, $\beta_3 = 3.3662$, $\beta_4 = 0.62517$) can represent $e_c(n_\uparrow, n_\downarrow)$ for $n_\uparrow = n$ and $n_\downarrow = 0$, the correlation energy per electron for a fully spin-polarized uniform gas. But we shall need $e_c(n_\uparrow, n_\downarrow)$ for arbitrary relative spin polarization

$$\zeta = \frac{(n_\uparrow - n_\downarrow)}{(n_\uparrow + n_\downarrow)} , \tag{149}$$

which ranges from 0 for an unpolarized system to ± 1 for a fully-spin-polarized system. A useful interpolation formula, based upon a study of the random phase approximation, is [10]

$$\begin{aligned} e_c(n_\uparrow, n_\downarrow) &= e_c(n) + \alpha_c(n)\frac{f(\zeta)}{f''(0)}(1-\zeta^4) + [e_c(n,0) - e_c(n)]f(\zeta)\zeta^4 \\ &= e_c(n) + \alpha_c(n)\zeta^2 + \mathcal{O}(\zeta^4) , \end{aligned} \tag{150}$$

where

$$f(\zeta) = \frac{[(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2]}{(2^{4/3} - 2)} \ . \tag{151}$$

In Eq. (150), $\alpha_c(n)$ is the correlation contribution to the spin stiffness. Roughly $\alpha_c(n) \approx e_c(n,0) - e_c(n)$, but more precisely $-\alpha_c(n)$ can be parametrized in the form of Eq. (143) (with $c_0 = 0.016887$, $c_1 = 0.035475$, $\alpha_1 = 0.11125$, $\beta_3 = 0.88026$, $\beta_4 = 0.49671$).

For completeness, we note that the spin-scaling relations (126) and (127) imply that

$$e_x(n_\uparrow, n_\downarrow) = e_x(n)\frac{[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}]}{2} , \tag{152}$$

$$t_s(n_\uparrow, n_\downarrow) = t_s(n)\frac{[(1+\zeta)^{5/3} + (1-\zeta)^{5/3}]}{2} \ . \tag{153}$$

The exchange-hole density of Eq. (137) can also be spin scaled. Expressions for the exchange and correlation holes for arbitrary r_s and ζ are given in Ref. [58].

5.4 Linear Response

We now discuss the linear response of the spin-unpolarized uniform electron gas to a weak, static, external potential $\delta v(\mathbf{r})$. This is a well-studied problem [59], and a practical one for the local-pseudopotential description of a simple metal [60].

Because the unperturbed system is homogeneous, we find that, to first order in $\delta v(\mathbf{r})$, the electron density response is

$$\delta n(\mathbf{r}) = \int \mathrm{d}^3 r' \chi(|\mathbf{r}-\mathbf{r}'|)\delta v(\mathbf{r}') \tag{154}$$

where χ is a linear response function. If

$$\delta v(\mathbf{r}) = \delta v(\mathbf{q}) \exp(i\mathbf{q}\cdot\mathbf{r}) \tag{155}$$

is a wave of wavevector $\mathbf{q}$ and small amplitude $\delta v(\mathbf{q})$, then Eq. (154) becomes $\delta n(\mathbf{r}) = \delta n(\mathbf{q}) \exp(i\mathbf{q}\cdot\mathbf{r})$, where

$$\delta n(\mathbf{q}) = \chi(\mathbf{q})\delta v(\mathbf{q}) \ , \tag{156}$$

and

$$\chi(\mathbf{q}) = \int \mathrm{d}^3 x \exp(-i\mathbf{q}\cdot\mathbf{x})\chi(|\mathbf{x}|) \tag{157}$$

is the Fourier transform of $\chi(|\mathbf{r}-\mathbf{r}'|)$ with respect to $\mathbf{x}=\mathbf{r}-\mathbf{r}'$. (In Eq. (155), the real part of the complex exponential $\exp(i\alpha) = \cos(\alpha) + i\sin(\alpha)$ is understood.)

By the Kohn-Sham theorem, we also have

$$\delta n(\mathbf{q}) = \chi_{\mathrm{s}}(q)\delta v_{\mathrm{s}}(\mathbf{q}) \ , \tag{158}$$

where $\delta v_{\mathrm{s}}(\mathbf{q})$ is the change in the Kohn-Sham effective one-electron potential of Eq. (62), and

$$\chi_{\mathrm{s}}(q) = -\frac{k_{\mathrm{F}}}{\pi^2} F(q/2k_{\mathrm{F}}) \tag{159}$$

is the density response function for the non-interacting uniform electron gas. The Lindhard function

$$F(x) = \frac{1}{2} + \frac{1-x^2}{4x} \ln\left|\frac{1+x}{1-x}\right| \tag{160}$$

equals $1 - x^2/3 - x^4/15$ as $x \to 0$, $1/2$ at $x=1$, and $1/(3x^2) + 1/(15x^4)$ as $x \to \infty$. $\mathrm{d}F/\mathrm{d}x$ diverges logarithmically as $x \to 1$.

Besides $\delta v(\mathbf{r})$, the other contributions to $\delta v_{\mathrm{s}}(\mathbf{r})$ of Eq. (62) are

$$\delta\left(\frac{\delta U}{\delta n(\mathbf{r})}\right) = \int \mathrm{d}^3 r' \frac{\delta n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ , \tag{161}$$

$$\delta\left(\frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})}\right) = \int \mathrm{d}^3 r' \frac{\delta^2 E_{\mathrm{xc}}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} \delta n(\mathbf{r}') \ . \tag{162}$$

In other words,

$$\delta v_{\mathrm{s}}(\mathbf{q}) = \delta v(\mathbf{q}) + \frac{4\pi}{q^2}\delta n(\mathbf{q}) - \frac{\pi}{k_F^2}\gamma_{\mathrm{xc}}(q)\delta n(\mathbf{q}) \ , \tag{163}$$

where the coefficient of the first $\delta n(\mathbf{q})$ is the Fourier transform of the Coulomb interaction $1/|\mathbf{r}-\mathbf{r}'|$, and the coefficient of the second $\delta n(\mathbf{q})$ is the Fourier transform of $\delta^2 E_{\mathrm{xc}}/\delta n(\mathbf{r})\delta n(\mathbf{r}')$.

We re-write Eq. (163) as

$$\delta v_{\mathrm{s}}(\mathbf{q}) = \delta v(\mathbf{q}) + \frac{4\pi}{q^2}\left[1 - G_{\mathrm{xc}}(q)\right]\delta n(\mathbf{q}) \ , \tag{164}$$

where

$$G_{\mathrm{xc}}(q) = \gamma_{\mathrm{xc}}(q)\left(\frac{q}{2k_{\mathrm{F}}}\right)^2 \tag{165}$$

is the so-called local-field factor. Then we insert Eq. (158) into Eq. (164) and find

$$\delta v_{\mathrm{s}}(\mathbf{q}) = \frac{\delta v(\mathbf{q})}{\epsilon_{\mathrm{s}}(q)} \tag{166}$$

where

$$\epsilon_{\mathrm{s}}(q) = 1 - \frac{4\pi}{q^2}\left[1 - G_{\mathrm{xc}}(q)\right]\chi_{\mathrm{s}}(q) \ . \tag{167}$$

In other words, the density response function of the interacting uniform electron gas is

$$\chi(q) = \frac{\chi_{\mathrm{s}}(q)}{\epsilon_{\mathrm{s}}(q)} \ . \tag{168}$$

These results are particularly simple in the long-wavelength ($q \to 0$) limit, in which $\gamma_{\mathrm{xc}}(q)$ tends to a constant and

$$\epsilon_{\mathrm{s}}(q) \to 1 - \frac{\gamma_{\mathrm{xc}}(q=0)}{\pi k_{\mathrm{F}}} + \frac{k_{\mathrm{s}}^2}{q^2} \quad (q \to 0) \ , \tag{169}$$

where

$$k_{\mathrm{s}} = \left(\frac{4k_{\mathrm{F}}}{\pi}\right)^{1/2} = \left(\frac{4}{\pi}\right)^{1/2}\left(\frac{9\pi}{4}\right)^{1/6}\frac{1}{r_{\mathrm{s}}^{1/2}} \tag{170}$$

is the inverse of the Thomas-Fermi screening length - the characteristic distance over which an external perturbation is screened out. Eqs. (166) and (167) show that a slowly-varying external perturbation $\delta v(\mathbf{q})$ is strongly "screened out" by the uniform electron gas, leaving only a very weak Kohn-Sham potential $\delta v_{\mathrm{s}}(\mathbf{q})$. Eq. (168) shows that the response function $\chi(q)$ is weaker than $\chi_{\mathrm{s}}(q)$ by a factor $(q/k_{\mathrm{s}})^2$ in the limit $q \to 0$.

In Eq. (166), $\epsilon_{\mathrm{s}}(q)$ is a kind of dielectric function , but it is not the standard dielectric function $\epsilon(q)$ which predicts the response of the electrostatic potential alone:

$$\delta v(\mathbf{q}) + \frac{4\pi}{q^2}\delta n(\mathbf{q}) = \frac{\delta v(\mathbf{q})}{\epsilon(q)} \ . \tag{171}$$

By inserting Eq. (156) into Eq. (171), we find

$$\frac{1}{\epsilon(q)} = 1 + \frac{4\pi}{q^2}\chi(q) \ . \tag{172}$$

It is only when we neglect exchange and correlation that we find the simple Lindhard result

$$\epsilon(q) \to \epsilon_s(q) \to \epsilon_L(q) = 1 - \frac{4\pi}{q^2}\chi_s(q) \quad (\gamma_{xc} \to 0) \ . \tag{173}$$

Neglecting correlation, γ_x is a numerically-tabulated function of $(q/2k_F)$ with the small-q expansion [61]

$$\gamma_x(q) = 1 + \frac{5}{9}\left(\frac{q}{2k_F}\right)^2 + \frac{73}{225}\left(\frac{q}{2k_F}\right)^4 \quad (q \to 0) \ . \tag{174}$$

When correlation is included, $\gamma_{xc}(q)$ depends upon r_s as well as $(q/2k_F)$, in a way that is known from Quantum Monte Carlo studies [62] of the weakly-perturbed uniform gas.

The second-order change δE in the total energy may be found from the Hellmann-Feynman theorem of Sect. 2.4. Replace $\delta v(\mathbf{r})$ by $v_\lambda(\mathbf{r}) = \lambda\delta v(\mathbf{r})$ and $\delta n(\mathbf{r})$ by $\lambda\delta n(\mathbf{r})$, to find

$$\begin{aligned}\delta E &= \int_0^1 \mathrm{d}\lambda \int \mathrm{d}^3r\, n_\lambda(\mathbf{r}) \frac{\mathrm{d}}{\mathrm{d}\lambda} v_\lambda(\mathbf{r}) \\ &= \int_0^1 \mathrm{d}\lambda \int \mathrm{d}^3r [n + \lambda\delta n(\mathbf{r})]\delta v(\mathbf{r}) \\ &= \frac{1}{2}\int \mathrm{d}^3r\, \delta n(\mathbf{r})\delta v(\mathbf{r}) \\ &= \frac{1}{2}\delta n(-\mathbf{q})\delta v(\mathbf{q}) \ .\end{aligned} \tag{175}$$

5.5 Clumping and Adiabatic Connection

The uniform electron gas for $r_s \lesssim 30$ provides a nice example of the adiabatic connection discussed in Sect. 3.5. As the coupling constant λ turns on from 0 to 1, the ground state wavefunction evolves continuously from the Kohn-Sham determinant of plane waves to the ground state of interacting electrons in the presence of the external potential, while the density remains fixed. (One should of course regard the infinite system as the infinite-volume limit of a finite chunk of uniform background neutralized by electrons.)

The adiabatic connection between non-interacting and interacting uniform-density ground states could be destroyed by any tendency of the density to clump. A fictitious attractive interaction between electrons would yield such a tendency. Even in the absence of attractive interactions, clumping appears in the very-low-density electron gas as a charge density wave or Wigner crystallization [56, 59]. Then there is probably no external potential which will hold the interacting system in a uniform-density ground state, but one can still find the energy of the uniform state by imposing density uniformity as a constraint on a trial interacting wavefunction.

The uniform phase becomes unstable against a charge density wave of wavevector $\mathbf{q}$ and infinitesimal amplitude when $\epsilon_s(q)$ of Eq. (167) vanishes [59]. This instability for $q \approx 2k_F$ arises at low density as a consequence of exchange and correlation.

Table 3. Exchange-correlation energies of atoms, in hartree.

Atom	LSD	GGA	Exact
H	-0.29	-0.31	-0.31
He	-1.00	-1.06	-1.09
Li	-1.69	-1.81	-1.83
Be	-2.54	-2.72	-2.76
N	-6.32	-6.73	-6.78
Ne	-11.78	-12.42	-12.50

Table 4. Atomization energies of molecules, in eV. (1 hartree = 27.21 eV). ¿From Ref. [20].

Molecule	LSD	GGA	Exact
H_2	4.9	4.6	4.7
CH_4	20.0	18.2	18.2
NH_3	14.6	13.1	12.9
H_2O	11.6	10.1	10.1
CO	13.0	11.7	11.2
O_2	7.6	6.2	5.2

6 Local and Semi-Local Approximations

6.1 Local Spin Density Approximation

The local spin density approximation (LSD) for the exchange-correlation energy, Eq. (11), was proposed in the original work of Kohn and Sham [6], and has proved to be remarkably accurate, useful, and hard to improve upon. The generalized gradient approximation (GGA) of Eq. (12), a kind of simple extension of LSD, is now more widely used in quantum chemistry, but LSD remains the most popular way to do electronic-structure calculations in solid state physics. Tables 1 and 2 provide a summary of typical errors for LSD and GGA, while Tables 3 and 4 make this comparison for a few specific atoms and molecules. The LSD is parametrized as in Sect. 5, while the GGA is the non-empirical one of Perdew, Burke, and Ernzerhof [20], to be presented later.

The LSD approximation to any energy component G is

$$G^{\mathrm{LSD}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3r\, n(\mathbf{r}) g(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})) \ , \tag{176}$$

where $g(n_\uparrow, n_\downarrow)$ is that energy component per particle in an electron gas with uniform spin densities $n_\uparrow$ and $n_\downarrow$, and $n(\mathbf{r})\mathrm{d}^3r$ is the average number of electrons in volume element d^3r. Sections 5.1-5.3 provide the ingredients for $T_s^{\mathrm{LSD}} = T_0$, E_x^{LSD}, and E_c^{LSD}. The functional derivative of Eq. (176) is

$$\frac{\delta G^{\mathrm{LSD}}}{\delta n_\sigma(\mathbf{r})} = \frac{\partial}{\partial n_\sigma}[(n_\uparrow + n_\downarrow) g(n_\uparrow, n_\downarrow)] \ . \tag{177}$$

By construction, LSD is exact for a uniform density, or more generally for a density that varies slowly over space [6]. More precisely, LSD should be valid when the length scale of the density variation is large in comparison with length scales set by the local density, such as the Fermi wavelength $2\pi/k_F$ or the screening length $1/k_s$. This condition is rarely satisfied in real electronic systems, so we must look elsewhere to understand why LSD works.

We need to understand why LSD works, for three reasons: to justify LSD calculations, to understand the physics, and to develop improved density functional approximations. Thus we will start with the good news about LSD, proceed to the mixed good/bad news, and close with the bad news.

LSD has *many* correct formal features. It is exact for uniform densities and nearly-exact for slowly-varying ones, a feature that makes LSD well suited at least to the description of the crystalline simple metals. It satisfies the inequalities $E_{\mathrm{x}} < 0$ (Eq. (93)) and $E_{\mathrm{c}} < 0$ (Eq. (69)), the correct uniform coordinate scaling of E_{x} (Eq. (106)), the correct spin scaling of E_{x} (Eq. (127)), the correct coordinate scaling for E_{c} (Eqs. (111), (116), (117)), the correct low-density behavior of E_{c} (Eq. (115)), and the correct Lieb-Oxford bound on E_{xc} (Eqs. (120) and (122)). LSD is properly size-consistent (Sect. 4.4).

LSD provides a surprisingly good account of the linear response of the spin-unpolarized uniform electron gas (Sect. 5.4). Since

$$\frac{\delta^2 E_{\mathrm{xc}}^{\mathrm{LSD}}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} = \delta(\mathbf{r}-\mathbf{r}')\frac{\partial^2[ne_{\mathrm{xc}}(n)]}{\partial n^2} \ , \tag{178}$$

where $\delta(\mathbf{r}-\mathbf{r}')$ is the Dirac delta function, we find

$$\gamma_{\mathrm{xc}}^{\mathrm{LSD}}(q) = 1 - \frac{k_{\mathrm{F}}^2}{\pi}\frac{\partial^2}{\partial n^2}[ne_{\mathrm{c}}(n)] \ , \tag{179}$$

a constant independent of q, which must be the exact $q \to 0$ or slowly-varying limit of $\gamma_{\mathrm{xc}}(q)$. Figure 1 of Ref. [20] shows that the "exact" $\gamma_{\mathrm{xc}}(q)$ from a Quantum Monte Carlo calculation [62] for $r_{\mathrm{s}} = 4$ is remarkably close to the LSD prediction for $q \lesssim 2k_{\mathrm{F}}$. The same is true over the whole valence-electron density range $2 \lesssim r_{\mathrm{s}} \lesssim 5$, and results from a strong cancellation between the nonlocalities of exchange and correlation. Indeed the exact result for exchange (neglecting correlation), Eq. (174), is strongly q-dependent or nonlocal. The displayed terms of Eq. (174) suffice for $q \lesssim 2k_{\mathrm{F}}$.

Powerful reasons for the success of LSD are provided by the coupling constant integration of Sect. 3.5. Comparison of Eqs. (86) and (11) reveals that the LSD approximations for the exchange and correlation holes of an inhomogeneous system are

$$n_{\mathrm{x}}^{\mathrm{LSD}}(\mathbf{r},\mathbf{r}') = n_{\mathrm{x}}^{\mathrm{unif}}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}); |\mathbf{r}-\mathbf{r}'|) \ , \tag{180}$$

$$n_{\mathrm{c}}^{\mathrm{LSD}}(\mathbf{r},\mathbf{r}') = n_{\mathrm{c}}^{\mathrm{unif}}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}); |\mathbf{r}-\mathbf{r}'|) \ , \tag{181}$$

where $n_{\mathrm{xc}}^{\mathrm{unif}}(n_\uparrow, n_\downarrow; u)$ is the hole in an electron gas with uniform spin densities $n_\uparrow$ and $n_\downarrow$. Since the uniform gas is a possible physical system, Eqs. (180) and (181) obey the exact constraints of Eqs. (91) (negativity of n_{x}), (94) (sum rule on n_{x}), (95), (97) (sum rule on $\bar{n}_{\mathrm{c}}$), (98), and (85) (cusp condition).

By Eq. (95), the LSD on-top exchange hole $n_{\mathrm{x}}^{\mathrm{LSD}}(\mathbf{r},\mathbf{r})$ is exact, at least when the Kohn-Sham wavefunction is a single Slater determinant. The LSD on-top correlation hole $\bar{n}_{\mathrm{c}}^{\mathrm{LSD}}(\mathbf{r},\mathbf{r})$ is *not* exact [63] (except in the high-density, low-density, fully spin-polarized, or slowly-varying limit), but it is often quite realistic [64]. By Eq. (85), its cusp is then also realistic.

Because it satisfies all these constraints, the LSD model for the system-, spherically-, and coupling-constant-averaged hole of Eq. (101),

$$\langle \bar{n}_{\mathrm{xc}}^{\mathrm{LSD}}(u)\rangle = \frac{1}{N}\int \mathrm{d}^3r\, n(\mathbf{r})\bar{n}_{\mathrm{xc}}^{\mathrm{unif}}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}); u) \ , \tag{182}$$

can be very physical. Moreover, the system average in Eq. (182) "unweights" regions of space where LSD is expected to be least reliable, such as near a nucleus or in the evanescent tail of the electron density [65, 64].

Since correlation makes $\langle \bar{n}_{\mathrm{xc}}(u=0)\rangle$ deeper, and thus by Eq. (102) makes $\langle \bar{n}_{\mathrm{xc}}(u)\rangle$ more short-ranged, E_{xc} can be "more local" than either E_{x} or E_{c}. In other words, LSD often benefits from a cancellation of errors between exchange and correlation.

Mixed good and bad news about LSD is the fact that selfconsistent LSD calculations can break exact spin symmetries. As an example, consider "stretched H_2", the hydrogen molecule ($N = 2$) with a very large separation between the two nuclei. The exact ground state is a spin singlet ($S = 0$), with $n_\uparrow(\mathbf{r}) = n_\downarrow(\mathbf{r}) = n(\mathbf{r})/2$. But the LSD ground state localizes all of the spin-up density on one of the nuclei, and all of the spin-down density on the other. Although (or rather *because*) the LSD spin densities are wrong, the LSD total energy is correctly the sum of the energies of two isolated hydrogen atoms, so this symmetry breaking is by no means entirely a bad thing [66, 67]. The selfconsistent LSD on-top hole density $\langle \bar{n}_{\mathrm{xc}}(0)\rangle = -\langle n\rangle$ is also right: Heitler-London correlation ensues that two electrons are never found near one another, or on the same nucleus at the same time.

Finally, we present the bad news about LSD: (1) LSD does not incorporate known inhomogeneity or gradient corrections to the exchange-correlation hole near the electron (Sect. 6.2) (2) It does not satisfy the high-density correlation scaling requirement of Eq. (114), but shows a $\ln\gamma$ divergence associated with the $\ln r_{\mathrm{s}}$ term of Eq. (140). (3) LSD is not exact in the one-electron limit, i.e., does not satisfy Eqs. (67), and (70)- (73). Although the "self-interaction error" is small for the exchange-correlation energy, it is more substantial for the exchange-correlation potential and orbital eigenvalues. (4) As a "continuum approximation", based as it is on the uniform electron gas and its continuous one-electron energy spectrum, LSD misses the derivative discontinuity of Sect. 4.5. Effectively, LSD averages over the discontinuity, so its highest occupied orbital energy for a Z-electron system is not Eq. (129) but $\varepsilon^{\mathrm{HO}} \approx -(I_Z + A_Z)/2$. A second consequence is that LSD predicts an incorrect dissociation of a hetero-nuclear molecule or solid to fractionally charged fragments. (In LSD calculations of atomization energies, the dissociation products are constrained to be neutral atoms, and not these unphysical fragments.) (5) LSD does not guarantee satisfaction of Eq. (99), an inherently nonlocal constraint.

The GGA to be derived in Sect. 6.4 will preserve all the good or mixed features of LSD listed above, while eliminating bad features (1) and (2) but not (3) - (5). Elimination of (3) - (5) will probably require the construction of $E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]$ from the Kohn-Sham orbitals (which are themselves highly-nonlocal functionals of the density). For example, the self-interaction correction [9, 68] to LSD eliminates most of the bad features (3) and (4), but not in an entirely satisfactory way.

6.2 Gradient Expansion

Gradient expansions [6, 69], which offer systematic corrections to LSD for electron densities that vary slowly over space, might appear to be the natural next step beyond LSD. As we shall see, they are not; understanding why not will light the path to the generalized gradient approximations of Sect. 6.3.

As a first measure of inhomogeneity, we define the reduced density gradient

$$s = \frac{|\nabla n|}{2k_{\mathrm{F}}n} = \frac{|\nabla n|}{2(3\pi)^{1/3}n^{4/3}} = \frac{3}{2}\left(\frac{4}{9\pi}\right)^{1/3}|\nabla r_{\mathrm{s}}| \ , \tag{183}$$

which measures how fast and how much the density varies on the scale of the local Fermi wavelength $2\pi/k_{\mathrm{F}}$. For the energy of an atom, molecule, or solid, the range $0 \lesssim s \lesssim 1$ is very important. The range $1 \lesssim s \lesssim 3$ is somewhat important, more so in atoms than in solids, while $s > 3$ (as in the exponential tail of the density) is unimportant [70, 71].

Other measures of density inhomogeneity, such as $p = \nabla^2 n/(2k_{\mathrm{F}})^2 n$, are also possible. Note that s and p are small not only for a slow density variation but also for a density variation of small amplitude (as in Sect. 5.4). The slowly-varying limit is one in which p/s is also small [6].

Under the uniform density scaling of Eq. (40), $s(\mathbf{r}) \to s_\gamma(\mathbf{r}) = s(\gamma\mathbf{r})$. The functionals $T_{\mathrm{s}}[n]$ and $E_{\mathrm{x}}[n]$ must scale as in Eqs. (104) and (106), so their gradient expansions are

$$T_{\mathrm{s}}[n] = A_{\mathrm{s}}\int \mathrm{d}^3r\, n^{5/3}[1 + \alpha s^2 + \ldots] \ , \tag{184}$$

$$E_{\mathrm{x}}[n] = A_{\mathrm{x}}\int \mathrm{d}^3r\, n^{4/3}[1 + \mu s^2 + \ldots] \ , \tag{185}$$

Because there is no special direction in the uniform electron gas, there can be no term linear in ∇n. Moreover, terms linear in $\nabla^2 n$ can be recast as s^2 terms, since

$$\int \mathrm{d}^3r\, f(n)\nabla^2 n = -\int \mathrm{d}^3r \left(\frac{\partial f}{\partial n}\right)|\nabla n|^2 \tag{186}$$

via integration by parts. Neglecting the dotted terms in Eqs. (184) and (185), which are fourth or higher-order in ∇, amounts to the second-order gradient expansion, which we call the gradient expansion approximation (GEA).

Correlation introduces a second length scale, the screening length $1/k_{\mathrm{s}}$, and thus another reduced density gradient

$$t = \frac{|\nabla n|}{2k_{\mathrm{s}}n} = \left(\frac{\pi}{4}\right)^{1/2}\left(\frac{9\pi}{4}\right)^{1/6}\frac{s}{r_{\mathrm{s}}^{1/2}} \ . \tag{187}$$

In the high-density ($r_{\mathrm{s}} \to 0$) limit, the screening length ($1/k_{\mathrm{s}} \sim r_{\mathrm{s}}^{1/2}$) is the only important length scale for the correlation hole. The gradient expansion of the correlation energy is

$$E_{\mathrm{c}}[n] = \int \mathrm{d}^3r\, n\left[e_{\mathrm{c}}(n) + \beta(n)t^2 + \ldots\right] \ . \tag{188}$$

While $e_{\mathrm{c}}(n)$ does not quite approach a constant as $n \to \infty$, $\beta(n)$ does [69].

While the form of the gradient expansion is easy to guess, the coefficients can only be calculated by hard work. Start with the uniform electron gas, in either its non-interacting (T_{s},E_{x}) or interacting (E_{c}) ground state, and apply a weak external perturbation $\delta v_{\mathrm{s}}(\mathbf{q})\exp(i\mathbf{q}\cdot\mathbf{r})$ or $\delta v(\mathbf{q})\exp(i\mathbf{q}\cdot\mathbf{r})$, respectively. Find the linear response $\delta n(\mathbf{q})$ of the density, and the second-order response δG of the energy component G of interest. Use the linear response of the density (as in Eqs. (157) or (156)) to express δG entirely in terms of $\delta n(\mathbf{q})$. Finally, expand δG in powers of q^2, observing that $|\nabla n|^2 \sim q^2|\delta n(\mathbf{q})|^2$, and extract the gradient coefficient.

In this way, Kirzhnits [72] found the gradient coefficient for T_{s},

$$\alpha = \frac{5}{27} \tag{189}$$

(which respects the conjectured bound of Eq. (118)), Sham [73] found the coefficient of E_{x},

$$\mu_{\mathrm{Sham}} = \frac{7}{81} \ , \tag{190}$$

and Ma and Brueckner [69] found the high-density limit of $\beta(n)$:

$$\beta_{\mathrm{MB}} = 0.066725 \ . \tag{191}$$

The weak density dependence of $\beta(n)$ is also known [74], as is its spin-dependence [75]. Neglecting small $\nabla\zeta$ contributions, the gradient coefficients (coefficients of $|\nabla n|^2/n^{4/3}$) for both exchange and correlation at arbitrary relative spin polarization ζ are found from those for $\zeta = 0$ through multiplication by [76]

$$\phi(\zeta) = \frac{1}{2}\left[(1+\zeta)^{2/3} + (1-\zeta)^{2/3}\right] \ . \tag{192}$$

For exchange, this is easily verified by applying the spin-scaling relation of Eq. (127) to Eqs. (185) and (183).

There is another interesting similarity between the gradient coefficients for exchange and correlation. Generalize the definition of t (Eq. (187)) to

$$t = \frac{|\nabla n|}{2\phi k_{\mathrm{s}}n} = \left(\frac{\pi}{4}\right)^{1/2}\left(\frac{9\pi}{4}\right)^{1/6}\frac{s}{\phi r_{\mathrm{s}}^{1/2}} \ . \tag{193}$$

Then

$$\beta_{\mathrm{MB}}\phi^3 n t^2 = \mu C_{\mathrm{x}}\phi n^{4/3}s^2 \ , \tag{194}$$

where

$$\mu = \beta_{\mathrm{MB}}\frac{\pi^2}{3} = 0.21951 \ . \tag{195}$$

Sham's derivation [73] of Eq. (190) starts with a screened Coulomb interaction $(1/u)\exp(-\kappa u)$, and takes the limit $\kappa \to 0$ at the end of the calculation. Antoniewicz and Kleinman [77] showed that the correct gradient coefficient for the unscreened Coulomb interaction is not μ_{Sham} but

$$\mu_{\mathrm{AK}} = \frac{10}{81} \quad . \tag{196}$$

It is believed [78] that a similar order-of limits problem exists for β, in such a way that the combination of Sham's exchange coefficient with the Ma-Brueckner [69] correlation coefficient yields the correct gradient expansion of E_{xc} in the slowly-varying high-density limit.

Numerical tests of these gradient expansions for atoms show that the second-order gradient term provides a useful correction to the Thomas-Fermi or local density approximation for T_{s}, and a modestly useful correction to the local density approximation for E_{x}, but seriously worsens the local spin density results for E_{c} and E_{xc}. In fact, the GEA correlation energies are positive! The latter fact was pointed out in the original work of Ma and Brueckner [69], who suggested the first generalized gradient approximation as a remedy.

The local spin density approximation to E_{xc}, which is the leading term of the gradient expansion, provides rather realistic results for atoms, molecules, and solids. But the second-order term, which is the next systematic correction for slowly-varying densities, makes E_{xc} worse.

There are two answers to the seeming paradox of the previous paragraph. The first is that realistic electron densities are *not* very close to the slowly-varying limit ($s \ll 1$, $p/s \ll 1$, $t \ll 1$, etc.). The second is this: The LSD approximation to the exchange-correlation hole is the hole of a possible physical system, the uniform electron gas, and so satisfies many exact constraints, as discussed in Sect. 6.1. The second-order gradient expansion or GEA approximation to the hole is not, and does not.

The second-order gradient expansion or GEA models are known for both the exchange hole [12, 13] $n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ and the correlation hole $\bar{n}_{\mathrm{c}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ [79]. They appear to be more realistic than the corresponding LSD models at small u, but far less realistic at large u, where several spurious features appear: $n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}+\mathbf{u})_{\mathrm{GEA}}$ has an undamped $\cos(2k_{\mathrm{F}}u)$ oscillation which violates the negativity constraint of Eq. (91), and integrates to -1 (Eq. (94)) only with the help of a convergence factor $\exp(-\kappa u)$ ($\kappa \to 0$). $\bar{n}_{\mathrm{c}}(\mathbf{r}, \mathbf{r}+\mathbf{u})_{\mathrm{GEA}}$ has a positive u^{-4} tail, and integrates not to zero (Eq. (97)) but to a positive number $\sim s^2$. These spurious large-u behaviors are sampled by the long range of the Coulomb interaction $1/u$, leading to unsatisfactory energies for real systems.

The gradient expansion for the exchange hole density is known [80] to third order in ∇, and suggests the following interpretation of the gradient expansion: When the density does not vary too rapidly over space (e.g., in the weak-pseudopotential description of a simple metal), the addition of each successive order of the gradient expansion improves the description of the hole at small u while worsening it at large u. The bad large-u behavior thwarts our expectation that the hole will remain normalized to each order in ∇.

The non-interacting kinetic energy T_{s} does not sample the spurious large-u part of the gradient expansion, so its gradient expansion (Eqs. (184) and (189)) works reasonably well

even for realistic electron densities. In fact, we can use Eq. (79) to show that

$$T_{\mathrm{s}}[n] = \sum_\sigma \int \mathrm{d}^3 r \, -\frac{1}{2}\frac{\partial}{\partial \mathbf{r}} \cdot \frac{\partial}{\partial \mathbf{r}} \rho_1^{\lambda=0}(\mathbf{r}'\sigma, \mathbf{r}\sigma)\Big|_{\mathbf{r}'=\mathbf{r}} \tag{197}$$

samples only the small-u part of the gradient expansion of the Kohn-Sham one-electron reduced density matrix, while $E_{\mathrm{x}}[n]$ of Eqs. (90) and (92) also samples large values of u. The GEA for $T_{\mathrm{s}}[n]$ is, in a sense, its own GGA [81]. Moreover, the sixth-order gradient expansion of T_{s} is also known: it diverges for finite systems, but provides accurate monovacancy formation energies for jellium [82].

The GEA form of Eqs. (184), (185), and (188) is a special case of the GGA form of Eq. (12). To find the functional derivative, note that

$$\begin{aligned}
\delta F &= \int \mathrm{d}^3 r \, \delta f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow) \\
&= \sum_\sigma \int \mathrm{d}^3 r \left[\frac{\partial f}{\partial n_\sigma(\mathbf{r})} \delta n_\sigma(\mathbf{r}) + \frac{\partial f}{\partial \nabla n_\sigma(\mathbf{r})} \cdot \nabla \delta n_\sigma(\mathbf{r}) \right] \\
&= \sum_\sigma \int \mathrm{d}^3 r \frac{\delta F}{\delta n_\sigma(\mathbf{r})} \delta n_\sigma(\mathbf{r}) \quad .
\end{aligned} \tag{198}$$

Integration by parts gives

$$\frac{\delta F}{\delta n_\sigma(\mathbf{r})} = \frac{\partial f}{\partial n_\sigma(\mathbf{r})} - \nabla \cdot \frac{\partial f}{\partial \nabla n_\sigma(\mathbf{r})} \quad . \tag{199}$$

For example, the functional derivative of the gradient term in the spin-unpolarized high-density limit is

$$\frac{\delta}{\delta n(\mathbf{r})} \int \mathrm{d}^3 r \, C_{\mathrm{xc}} \frac{|\nabla n(\mathbf{r})|^2}{n^{4/3}} = C_{\mathrm{xc}} \left[\frac{4}{3} \frac{|\nabla n(\mathbf{r})|^2}{n^{7/3}} - 2 \frac{\nabla^2 n}{n^{4/3}} \right] \quad , \tag{200}$$

which involves second as well as first derivatives of the density.

The GEA for the linear response function $\gamma_{\mathrm{xc}}(q)$ of Eq. (163) is found by inserting $n(\mathbf{r}) = n + \delta n(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r})$ into Eq. (199) and linearizing in $\delta n(\mathbf{q})$:

$$\gamma_{\mathrm{xc}}^{\mathrm{GEA}}(q) = \gamma_{\mathrm{xc}}^{\mathrm{LSD}} - 24\pi (3\pi^2)^{1/3} C_{\mathrm{xc}} \left(\frac{q}{2k_{\mathrm{F}}} \right)^2 \quad . \tag{201}$$

For example, the Antoniewicz-Kleinman gradient coefficient [77] for exchange of Eq. (196), inserted into Eqs. (200) and (201), yields the q^2 term of Eq. (174).

6.3 History of Several Generalized Gradient Approximations

In 1968, Ma and Brueckner [69] derived the second-order gradient expansion for the correlation energy in the high-density limit, Eqs. (188) and (191). In numerical tests, they found that it led to improperly positive correlation energies for atoms, because of the large size of the positive gradient term. As a remedy, they proposed the first GGA,

$$E_{\mathrm{c}}^{\mathrm{MB}}[n] = \int \mathrm{d}^3 r \, n e_{\mathrm{c}}(n) \left[1 - \frac{\beta_{\mathrm{MB}} t^2}{\nu n e_{\mathrm{c}}(n)} \right]^{-\nu} \quad , \tag{202}$$

where $\nu \approx 0.32$ was fitted to known correlation energies. Eq. (202) reduces to Eqs. (188) and (191) in the slowly-varying ($t \to 0$) limit, but provides a strictly negative "energy density" which tends to zero as $t \to \infty$. In this respect, it is strikingly like the nonempirical GGA's that were developed in 1991 or later, differing from them mainly in the presence of an empirical parameter, the absence of a spin-density generalization, and a less satisfactory high-density limit.

Under the uniform scaling of Eq. (40), $n(\mathbf{r}) \to n_\gamma(\mathbf{r})$, we find $r_s(\mathbf{r}) \to \gamma^{-1} r_s(\gamma\mathbf{r})$, $\zeta(\mathbf{r}) \to \zeta(\gamma\mathbf{r})$, $s(\mathbf{r}) \to s(\gamma\mathbf{r})$, and $t(\mathbf{r}) \to \gamma^{1/2} t(\gamma\mathbf{r})$. Thus $E_c^{\mathrm{MB}}[n_\gamma]$ tends to $E_c^{\mathrm{LSD}}[n_\gamma]$ as $\gamma \to \infty$, and not to a negative constant as required by Eq. (114).

In 1980, Langreth and Perdew [83] explained the failure of the second-order gradient expansion (GEA) for E_c. They made a complete wavevector analysis of E_{xc}, i.e., they replaced the Coulomb interaction $1/u$ in Eq. (100) by its Fourier transform and found

$$E_{xc}[n] = \frac{N}{2}\int_0^\infty \mathrm{d}k \frac{4\pi k^2}{(2\pi)^3}\langle \bar{n}_{xc}(k)\rangle \frac{4\pi}{k^2} \ , \tag{203}$$

where

$$\langle \bar{n}_{xc}(k)\rangle = \int_0^\infty \mathrm{d}u\ 4\pi u^2 \langle \bar{n}_{xc}(u)\rangle \frac{\sin(ku)}{ku} \tag{204}$$

is the Fourier transform of the system- and spherically-averaged exchange-correlation hole. In Eq. (203), E_{xc} is decomposed into contributions from dynamic density fluctuations of various wavevectors k.

The sum rule of Eq. (102) should emerge from Eq. (204) in the $k \to 0$ limit (since $\sin(x)/x \to 1$ as $x \to 0$), and does so for the exchange energy at the GEA level. But the $k \to 0$ limit of $\bar{n}_c^{\mathrm{GEA}}(k)$ turns out to be a positive number proportional to t^2, and not zero. The reason seems to be that the GEA correlation hole is only a truncated expansion, and not the exact hole for any physical system, so it can and does violate the sum rule.

Langreth and Mehl [11] (1983) proposed a GGA based upon the wavevector analysis of Eq. (203). They introduced a sharp cutoff of the spurious small-k contributions to E_c^{GEA}: all contributions were set to zero for $k < k_c = f|\nabla n/n|$, where $f \approx 0.15$ is only semi-empirical since $f \approx 1/6$ was estimated theoretically. Extension of the Langreth-Mehl E_c^{GGA} beyond the random phase approximation was made by Perdew [14] in 1986.

The errors of the GEA for the exchange energy are best revealed in real space (Eq. (100)), not in wavevector space (Eq. (203)). In 1985, Perdew [12] showed that the GEA for the exchange hole density $n_x(\mathbf{r}, \mathbf{r}+\mathbf{u})$ contains a spurious undamped $\cos(2k_F u)$ oscillation as $u \to \infty$, which violates the negativity constraint of Eq. (91) and respects the sum rule of Eq. (94) only with the help of a convergence factor (e.g., $\exp(-\kappa u)$ as $\kappa \to 0$). This suggested that the required cutoffs should be done in real space, not in wavevector space. The GEA hole density $n_x^{\mathrm{GEA}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ was replaced by zero for all $\mathbf{u}$ where n_x^{GEA} was positive, and for all $u > u_x(\mathbf{r})$ where the cutoff radius $u_x(\mathbf{r})$ was chosen to recover Eq. (94). Eq. (92) then provided a numerically-defined GGA for E_x, which turned out to be more accurate than either LSD or GEA. In 1986, Perdew and Wang [13] simplified this GGA in two ways: (1) They replaced $n_x^{\mathrm{GEA}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$, which depends upon both first and second

derivatives of $n(\mathbf{r})$, by $\tilde{n}_x^{\mathrm{GEA}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$, an equivalent expression found through integration by parts, which depends only upon $\nabla n(\mathbf{r})$. (2) The resulting numerical GGA has the form

$$E_x^{\mathrm{GGA}}[n] = A_x \int \mathrm{d}^3 r\, n^{4/3} F_x(s) \ , \tag{205}$$

which scales properly as in Eq. (106). The function $F_x(s)$ was plotted and fitted by an analytic form. The spin-scaling relation (127) was used to generate a spin-density generalization. Perdew and Wang [13] also coined the term "generalized gradient approximation".

A parallel but more empirical line of GGA development arose in quantum chemistry around 1986. Becke [15, 16] showed that a GGA for E_x could be constructed with the help of one or two parameters fitted to exchange energies of atoms, and demonstrated numerically that these functionals could greatly reduce the LSD overestimate of atomization energies of molecules. Lee, Yang, and Parr [17] transformed the Colle-Salvetti [84] expression for the correlation energy from a functional of the Kohn-Sham one-particle density matrix into a functional of the density. This functional contains one empirical parameter and works well in conjunction with Becke [16] exchange for many atoms and molecules, although it underestimates the correlation energy of the uniform electron gas by about a factor of two at valence-electron densities.

The real-space cutoff of the GEA hole provides a powerful nonempirical way to construct GGA's. Since exchange and correlation should be treated in a balanced way, there was a need to extend the 1986 real-space cutoff construction [13] from exchange to correlation with the help of a second cutoff radius $u_c(\mathbf{r})$ chosen to satisfy Eq. (97). Without accurate formulas for the correlation hole of the uniform electron gas, this extension had to wait until 1991, when it led to the Perdew-Wang 1991 (PW91) [18, 79] GGA for E_{xc}. For most practical purposes, PW91 is equivalent to the Perdew-Burke-Ernzerhof [20, 21] (PBE) "GGA made simple", which will be derived, presented, and discussed in the next two sections.

6.4 Construction of a "GGA Made Simple"

The PW91 GGA and its construction [18, 79] are simple in principle, but complicated in practice by a mass of detail. In 1996, Perdew, Burke and Ernzerhof [20, 21] (PBE) showed how to construct essentially the same GGA in a much simpler form and with a much simpler derivation.

Ideally, an approximate density functional $E_{xc}[n_\uparrow, n_\downarrow]$ should have all of the following features: (1) a non-empirical derivation, since the principles of quantum mechanics are well-known and sufficient; (2) universality, since in principle one functional should work for diverse systems (atoms, molecules, solids) with different bonding characters (covalent, ionic, metallic, hydrogen, and van der Waals); (3) simplicity, since this is our only hope for intuitive understanding and our best hope for practical calculation; and (4) accuracy enough to be useful in calculations for real systems.

The LSD of Eq. (11) and the non-empirical GGA of Eq. (12) nicely balance these desiderata. Both are exact only for the electron gas of uniform density, and represent controlled extrapolations away from the slowly-varying limit (unlike the GEA of Sect.

6.2, which is an uncontrolled extrapolation). LSD is a controlled extrapolation because, even when applied to a density that varies rapidly over space, it preserves many features of the exact E_{xc}, as discussed in Sect. 6.1. LSD has worked well in solid state applications for thirty years.

Our conservative philosophy of GGA construction is to try to retain all the correct features of LSD, while adding others. In particular, we retain the correct uniform-gas limit, for two reasons: (1) This is the only limit in which the restricted GGA form *can* be exact. (2) Nature's data set includes the crystalline simple metals like Na and Al. The success of the stabilized jellium model [85] reaffirms that the valence electrons in these systems are correlated very much as in a uniform gas. Among the welter of possible conditions which could be imposed to construct a GGA, the most natural and important are those respected by LSD or by the real-space cutoff construction of PW91, and these are the conditions chosen in the PBE derivation [20] below. The resulting GGA is one in which all parameters (other than those in LSD) are fundamental constants.

We start by writing the correlation energy in the form

$$E_{\mathrm{c}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3 r n [e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta) + H(r_{\mathrm{s}}, \zeta, t)] \ , \tag{206}$$

where the local density parameters r_{s} and ζ are defined in Eqs. (133) and (149), and the reduced density gradient t in Eq. (193). The small-t behavior of nH should be given by the left-hand side of Eq. (194), which emerges naturally from the real-space cutoff construction of PW91 [79]. In the opposite or $t \to \infty$ limit, we expect that $H \to -e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta)$, the correlation energy per electron of the uniform gas, as it does in the PW91 construction or in the Ma-Brueckner GGA of Eq. (202). Finally, under the uniform scaling of Eq. (40) to the high-density ($\gamma \to \infty$) limit, Eq. (206) should tend to a negative constant, as in Eq. (114) or in the numerically-constructed PW91. This means that H must cancel the logarithmic singularity of e_{c} (Eq. (140)) in this limit.

A simple function which meets these expectations is

$$H = c_0 \phi^3 \ln \left\{ 1 + \frac{\beta_{\mathrm{MB}}}{c_0} t^2 \left[\frac{1 + At^2}{1 + At^2 + A^2 t^4} \right] \right\} \ , \tag{207}$$

where ϕ is given by Eq. (192) and

$$A = \frac{\beta_{\mathrm{MB}}}{c_0} \frac{1}{\exp\left[-e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta)/c_0 \phi^3\right] - 1} \ . \tag{208}$$

We now check the required limits:

$$\begin{aligned} t \to 0: \quad H &\to c_0 \phi^3 \ln \left\{ 1 + \frac{\beta_{\mathrm{MB}}}{c_0} t^2 \right\} \\ &\to \beta_{\mathrm{MB}} \phi^3 t^2 \ . \end{aligned} \tag{209}$$

$$\begin{aligned} t \to \infty: \quad H &\to c_0 \phi^3 \ln \left\{ 1 + \frac{\beta_{\mathrm{MB}}}{c_0 A} \right\} \\ &\to c_0 \phi^3 \ln \left\{ \exp\left[-\frac{e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta)}{c_0 \phi^3} \right] \right\} \\ &\to -e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta) \ . \end{aligned} \tag{210}$$

$$r_{\mathrm{s}} \to 0 \text{ at fixed } s: \quad H \to c_0 \phi^3 \ln t^2 \to -c_0 \phi^3 \ln r_{\mathrm{s}} \ . \tag{211}$$

To a good approximation, Eq. (140) can be generalized to

$$e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta) = \phi^3 [c_0 \ln r_{\mathrm{s}} - c_1 + \ldots] \ , \tag{212}$$

which cancels the log singularity of Eq. (211).

Under uniform density scaling to the high-density limit, we find

$$\gamma \to \infty: \quad E_{\mathrm{c}}^{\mathrm{GGA}}[n_\gamma] \to -c_0 \int \mathrm{d}^3 r n \phi^3 \ln \left[1 + \frac{1}{\chi s^2/\phi^2 + (\chi s^2/\phi^2)^2} \right] \tag{213}$$

(where s is defined by Eq. (183)), a negative constant as required by Eq. (114), with

$$\chi = \left(\frac{3\pi^2}{16} \right)^{2/3} \frac{\beta_{\mathrm{MB}}}{c_0} \exp(-c_1/c_0) \ . \tag{214}$$

For a two-electron ion of nuclear charge Z in the limit $Z \to \infty$, Eq. (213) is -0.0479 hartree and the exact value is -0.0467. Realistic results from Eq. (213) in the $Z \to \infty$ limit have also be found [86] for ions with 3, 9, 10, and 11 electrons.

Now we turn to the construction of a GGA for the exchange energy. Because of the spin-scaling relation (127), we only need to construct $E_{\mathrm{x}}^{\mathrm{GGA}}[n]$, which must be of the form of Eq. (205). To recover the good LSD description of the linear response of the uniform gas (Sect. 5.4), we choose the gradient coefficient for exchange to cancel that for correlation, i.e., we take advantage of Eq. (194) to write

$$s \to 0: \quad F_{\mathrm{x}}(s) = 1 + \mu s^2 \ . \tag{215}$$

Then the gradient coefficients for exchange and correlation will cancel for *all* r_{s} and ζ, apart from small $\nabla\zeta$ contributions to $E_{\mathrm{x}}^{\mathrm{GGA}}$, as discussed in the next section.

The value of μ of Eq. (195) is 1.78 times bigger than μ_{AK} of Eq. (196), the proper gradient coefficient for exchange in the slowly-varying limit. But this choice can be justified in two other ways as well: (a) It provides a decent fit to the results of the real-space cutoff construction [79] of the PW91 exchange energy, which does *not* recover μ_{AK} in the slowly-varying limit. (b) It provides a reasonable emulation of the exact-exchange linear response function of Eq. (174) over the important range of $0 < q/2k_{\mathrm{F}} \lesssim 1$ (but not of course in the limit $q \to 0$, where μ_{AK} is needed).

Finally, we want to satisfy the Lieb-Oxford bound of Eqs. (120) and (122), which LSD respects. We can achieve this, and also recover the limit of Eq. (215), with the simple form

$$F_{\mathrm{x}}(s) = 1 + \kappa - \frac{\kappa}{(1 + \mu s^2/\kappa)} \ , \tag{216}$$

where κ is a constant less than or equal to 0.804. Taking $\kappa = 0.804$ gives a GGA which is virtually identical to PW91 over the range of densities and reduced density gradients important in most real systems. We shall complete the discussion of this paragraph in the next section.

6.5 GGA Nonlocality: Its Character, Origins, and Effects

A useful way to visualize and think about gradient-corrected nonlocality, or to compare one GGA with another, is to write [19, 87]

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] \approx \int \mathrm{d}^3 r n \left(-\frac{c}{r_\mathrm{s}}\right) F_{\mathrm{xc}}(r_\mathrm{s}, \zeta, s) \ , \tag{217}$$

where $c = (3/4\pi)(9\pi/4)^{1/3}$ and $-c/r_\mathrm{s} = e_\mathrm{x}(r_\mathrm{s}, \zeta = 0)$ is the exchange energy per electron of a spin-unpolarized uniform electron gas. The enhancement factor $F_{\mathrm{xc}}(r_\mathrm{s}, \zeta, s)$ shows the effects of correlation (through its r_s dependence), spin polarization (ζ), and inhomogeneity or nonlocality (s). F_{xc} is the analog of $3\alpha/2$ in Slater's $X\alpha$ method [88], so its variation is bounded and plottable. Figure 1 shows $F_{\mathrm{xc}}(r_\mathrm{s}, \zeta = 0, s)$, the enhancement factor for a spin-unpolarized system. Figure 2 shows $F_{\mathrm{xc}}(r_\mathrm{s}, \zeta = 1, s) - F_{\mathrm{xc}}(r_\mathrm{s}, \zeta = 0, s)$, the enhancement factor for the spin polarization energy. (Roughly, $F_{\mathrm{xc}}(r_\mathrm{s}, \zeta, s) \approx F_{\mathrm{xc}}(r_\mathrm{s}, \zeta = 0, s) + \zeta^2 [F_{\mathrm{xc}}(r_\mathrm{s}, \zeta = 1, s) - F_{\mathrm{xc}}(r_\mathrm{s}, \zeta = 0, s)]$). The nonlocality is the s-dependence, and

$$F_{\mathrm{xc}}^{\mathrm{LSD}}(r_\mathrm{s}, \zeta, s) = F_{\mathrm{xc}}(r_\mathrm{s}, \zeta, s = 0) \tag{218}$$

is visualized as a set of horizontal straight lines coinciding with the GGA curves in the limit $s \to 0$.

Clearly, the correlation energy of Eq. (206) can be written in the form of Eq. (217). To get the exchange energy into this form, apply the spin-scaling relation (127) to Eq. (205), then drop small ∇s contributions to find

$$\begin{aligned} F_\mathrm{x}(\zeta, s) &= \frac{1}{2}(1+\zeta)^{4/3} F_\mathrm{x}\left(s/(1+\zeta)^{1/3}\right) + \frac{1}{2}(1-\zeta)^{4/3} F_\mathrm{x}\left(s/(1-\zeta)^{1/3}\right) \\ &= \frac{1}{2}\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right] + \mu\phi s^2 + \ldots \end{aligned} \tag{219}$$

Now

$$F_{\mathrm{xc}}(r_\mathrm{s}, \zeta, s) = F_\mathrm{x}(\zeta, s) + F_\mathrm{c}(r_\mathrm{s}, \zeta, s) \ , \tag{220}$$

where

$$\lim_{r_\mathrm{s} \to 0} F_\mathrm{c}(r_\mathrm{s}, \zeta, s) = 0 \tag{221}$$

by Eqs. (106) and (114). Thus the $r_\mathrm{s} = 0$ or high-density-limit curve in each figure is the exchange-only enhancement factor. Clearly $F_\mathrm{x} > 0$, $F_\mathrm{c} > 0$, and $F_\mathrm{x}(\zeta = 0, s = 0) = 1$ by definition.

The Lieb-Oxford bound of Eq. (122) will be satisfied for all densities $n(\mathbf{r})$ if and only if

$$F_{\mathrm{xc}}(r_\mathrm{s}, \zeta, s) \geq 2.273 \ . \tag{222}$$

For the PBE GGA of Eqs. (206) and (216), this requires that

$$2^{1/3} F_\mathrm{x}(s/2^{1/3}) \leq 2.273 \ , \tag{223}$$

or

$$\kappa \leq 2.273/2^{1/3} - 1 = 0.804 \ , \tag{224}$$

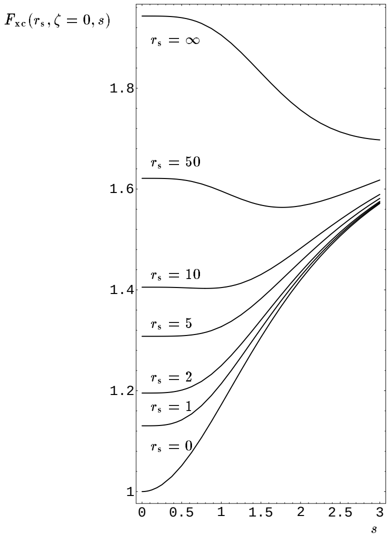

Fig. 1. The enhancment factor F_{xc} of Eq. (217) for the GGA of Perdew, Burke, and Ernzerhof [20], as a function of the reduced density gradient s of Eq. (183), for $\zeta = 0$. The local density parameter r_s and the relative spin polarization ζ are defined in Eqs. (133) and (149), respectively.

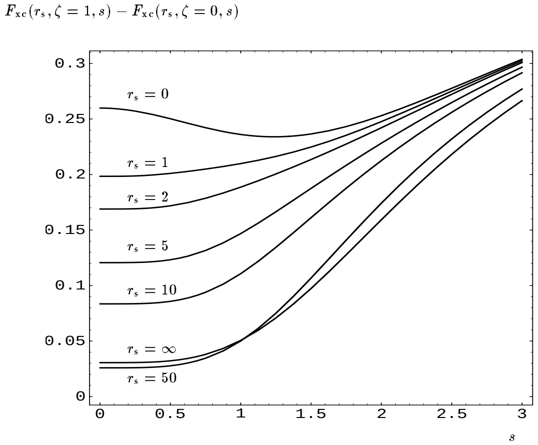

Fig. 2. Same as Fig. 1, but for the difference between the fully spin-polarized ($\zeta = 1$) and unpolarized ($\zeta = 0$) enhancement factors.

as stated in Sect. 6.4.

There is much to be seen and explained [21] in Eq. (217) and Figs. 1 and 2. However, the main qualitative features are simply stated: When we make a density variation in which r_s decreases, ζ increases, or s increases everywhere, we find that $|E_x|$ increases and $|E_c/E_x|$ decreases.

To understand this pattern [21], we note that the second-order gradient expansion for the non-interacting kinetic energy $T_s[n_\uparrow, n_\downarrow]$, which is arguably its own GGA [81], can be written as

$$T_s^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3r\, n \frac{3}{10} \frac{\left(\frac{9\pi}{4}\right)^{2/3}}{r_s^2} G(\zeta, s) \ , \tag{225}$$

$$G(\zeta, s) = \frac{1}{2}\left[(1+\zeta)^{5/3} + (1-\zeta)^{5/3}\right] + \frac{5}{27}s^2 \ , \tag{226}$$

using approximate spin scaling (Eq. (126) plus neglect of $\nabla\zeta$ contributions). Eqs. (225) and (226) respect Eq. (104) and confirm our intuition based upon the Pauli exclusion and uncertainty principles: Under a density variation in which r_s decreases, ζ increases, or s increases everywhere, we find that $T_s[n_\uparrow, n_\downarrow]$ increases.

The first effect of such an increase in T_s is an increase in $|E_x|$. T_s and $|E_x|$ are "conjoint" [89], in the sense that both can be constructed from the occupied Kohn-Sham orbitals (Eqs. (7), (88), (90) and (92)). With more kinetic energy, these occupied orbitals will have shorter de Broglie wavelengths. By the uncertainty principle, they can then dig a more short-ranged and deeper exchange hole with a more negative exchange energy. Thus exchange turns on when we decrease r_s, increase ζ, or increase s.

The second effect of such an increase in T_s is to strengthen the Kohn-Sham Hamiltonian which holds non-interacting electrons at the spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$. This makes the electron-electron repulsion of Eq. (112) a relatively weaker perturbation on the Kohn-Sham problem, and so reduces the ratio $|E_c/E_x|$. Thus correlation turns off relative to exchange when we decrease r_s, increase ζ, or increase s.

We note in particular that $F_x(r_s, \zeta, s)$ increases while $F_c(r_s, \zeta, s)$ decreases with increasing s. The nonlocalities of exchange and correlation are opposite, and tend to cancel for valence-electron densities ($1 \lesssim r_s \lesssim 10$) in the range $0 \lesssim s \lesssim 1$. The same remarkable cancellation occurs [62, 21] in the linear response function for the uniform gas of Eq. (163), i.e., $\gamma_{xc}(q) \approx \gamma_{xc}^{\mathrm{LSD}}(q) = \gamma_{xc}(q=0)$ for $0 \le q/2k_F \lesssim 1$.

The core electrons in any system, and the valence electrons in solids, sample primarily the range $0 \lesssim s \lesssim 1$. The high-density core electrons see a strong, exchange-like nonlocality of E_{xc} which provides an important correction to the LSD total energy. But the valence electrons in solids see an almost-complete cancellation between the nonlocalities of exchange and correlation. This helps to explain why LSD has been so successful in solid state physics, and why the small residue of GGA nonlocality in solids does not provide a universally-better description than LSD.

The valence electrons in atoms and molecules see $0 \lesssim s \lesssim \infty$, when s diverges in the exponential tail of the density, but the energetically-important range is $0 \lesssim s \lesssim 3$

[70, 71]. Figs. 1 and 2 show that GGA nonlocality *is* important in this range, so GGA is almost-always better than LSD for atoms and molecules.

For $r_s \lesssim 10$, the residual GGA nonlocality is exchange-like, i.e., exchange and correlation together turn on stronger with increasing inhomogeneity. It can then be seen from Eq. (217) that gradient corrections will favor greater density inhomogeneity and higher density [70]. Defining average density parameters $\langle r_s \rangle$, $\langle \zeta \rangle$, and $\langle s \rangle$ as in Ref. [70], we find that gradient corrections favor changes $\mathrm{d}\langle s \rangle > 0$ and $\mathrm{d}\langle r_s \rangle < 0$. Gradient corrections tend to drive a process forward when [70]

$$\frac{\mathrm{d}\langle s \rangle}{\langle s \rangle} \ge \frac{\mathrm{d}\langle r_s \rangle}{\langle r_s \rangle} \ . \tag{227}$$

In a typical process (bond stretching, transition to a more open structure, fragmentation, or atomization), one has $\mathrm{d}\langle s \rangle > 0$ and $\mathrm{d}\langle r_s \rangle > 0$. Thus, by Eq. (227), these effects compete - another reason why LSD has met with some success. In most such cases, the left-hand side of Eq. (227) is bigger than the right, so typically gradient corrections favor larger bond lengths or lattice constants (and thus softer vibration frequencies), more open structures, fragmentation of a highly-bonded transition state, or atomization of a molecule. In the case of bond stretching in H_2, however, the right hand side of Eq. (227) exceeds the left, so gradient corrections actually and correctly shrink the equilibrium bond length relative to LSD.

There have been many interesting tests and applications of GGA to a wide range of atoms, molecules, and solids. Some references will be found in Refs. [19, 90, 79, 21].

We close by discussing those situations in which LSD or GGA can fail badly. They seem to be of two types: (1) When the Kohn-Sham non-interacting wavefunction is not a single Slater determinant, or when the non-interacting energies are nearly degenerate, the LSD and GGA exchange-correlation holes can be unrealistic even very close to or on top of the electron [36, 91, 66]. (2) In an extended system, the exact hole may display a diffuse long-range tail which is not properly captured by either LSD or GGA. To a limited extent, this effect could be mimicked by reducing the parameter κ in Eq. (216). An example of a diffuse hole arises in the calculation of the surface energy of a metal [19, 32]: When an electron wanders out into the vacuum region, the exchange-correlation hole around it can extend significantly backward into the interior of the metal. A more extreme example is "stretched H_2^+", the ground state of one electron in the presence of two protons at very large separation: Half of the exact hole is localized on each proton, a situation which has no analog in the electron gas of uniform or slowly-varying density, and for which LSD and GGA make large self-interaction errors [9, 92, 68].

"Stretched H_2^+" and related systems are of course unusual. In most systems, the exact exchange-correlation hole is reasonably localized around its electron, as it is in LSD or GGA - and that fact is one of the reasons [93] why LSD and GGA work as well as they do.

Acknowledgements: Work supported in part by the U.S. National Science Foundation under Grant No. DMR95-21353, and in part by the Deutsche Forschungsgemeinschaft. We thank Matthias Ernzerhof for help with the figures.

References

1. D. Park, *Introduction to the Quantum Theory* (McGraw-Hill, New York, 1974).
2. H.A. Bethe, *Intermediate Quantum Mechanics* (Benjamin, New York, 1964).
3. R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
4. R.M. Dreizler and E.K.U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).
5. R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
6. W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).
7. U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).
8. J.P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).
9. J.P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
10. S.H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
11. D.C. Langreth and M.J. Mehl, Phys. Rev. B **28**, 1809 (1983).
12. J.P. Perdew, Phys. Rev. Lett. **55**, 1665 (1985).
13. J.P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
14. J.P. Perdew, Phys. Rev. B **33**, 8822 (1986).
15. A.D. Becke, J. Chem. Phys. **84**, 4524 (1986).
16. A.D. Becke, Phys. Rev. A **38**, 3098 (1988).
17. C. Lee, W. Yang, and R.G. Parr, Phys. Rev. B **37**, 785 (1988).
18. J.P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.
19. J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, and C. Fiolhais , Phys. Rev. B **46**, 6671 (1992), ibid., **48**, 4978 (1993)(E).
20. J.P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996), ibid, **78**, 1396 (1997)(E).
21. J.P. Perdew, M. Ernzerhof, A. Zupan, and K. Burke, submitted to J. Chem. Phys.
22. J. Hafner, *From Hamiltonians to Phase Diagrams* (Dover, New York, 1988).
23. A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry* (Macmillan Publishing, New York, 1982).
24. P. Fulde, *Electron Correlations in Molecules and Solids* (Springer, Berlin, 1993).
25. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
26. I.N. Levine, *Quantum Chemistry* (Allyn and Bacon, Boston, 1974).
27. D.P. Joubert, Int. J. Quantum Chem. **61**, 355 (1997).
28. M. Levy, Proc. Natl. Acad. Sci. USA **76**, 6062 (1979).
29. M. Levy, in *Recent Developments and Applications of Modern Density Functional Theory*, edited by J.M. Seminario (Elsevier, Amsterdam, 1996), p. 3.
30. J. Harris and R.O. Jones, J. Phys. F **4**, 1170 (1974).
31. D.C. Langreth and J.P. Perdew, Solid State Commun. **17**, 1425 (1975).
32. D.C. Langreth and J.P. Perdew, Phys. Rev. B **15**, 2884 (1977).
33. O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
34. E.R. Davidson, *Reduced Density Matrices in Quantum Chemistry* (Academic Press, New York, 1976).
35. J.C. Kimball, Phys. Rev. A **7**, 1648 (1973).
36. T. Ziegler, A. Rauk, and E.J. Baerends, Theoret. Chim. Acta **43**, 261 (1977).
37. K. Burke and J.P. Perdew, Int. J. Quantum Chem. **56**, 199 (1995).
38. M. Levy and J.P. Perdew, Phys. Rev. A **32**, 2010 (1985).
39. M. Levy, in *Single-Particle Density in Physics and Chemistry*, edited by N.H.March and B.M. Deb (Academic, London, 1987), p. 45.
40. A. Görling and M. Levy, Phys. Rev. B **47**, 13105 (1993).
41. A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994).
42. M. Levy, Phys. Rev. A **43**, 4637 (1991).
43. M. Levy and J.P. Perdew, Phys. Rev. B **48**, 11638 (1993).
44. E.H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).
45. E.H. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).
46. G.L. Oliver and J.P. Perdew, Phys. Rev. A **20**, 397 (1979).
47. J.P. Perdew, in *Density Functional Theory of Many-Fermion Systems*, Vol. 21 of *Advances in Quantum Chemistry*, edited by S.B. Trickey (Academic, New York, 1990), p. 113.
48. E. Fermi and E. Amaldi, Accad. Ital. Rome **6**, 119 (1934).
49. J.P. Perdew, R.G. Parr, M. Levy, and J.L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).
50. J.P. Perdew, in *Density Functional Methods in Physics*, Vol. B123 of *NATO ASI Series*, edited by R.M. Dreizler and J. da Providencia (Plenum Press, New York, 1985), p. 265.
51. J.F. Janak, Phys. Rev. B **18**, 7165 (1978).
52. M. Levy, J.P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).
53. N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Holt, Rinehart, and Winston, New York, 1976).

54. M. Gell-Mann and K.A. Brueckner, Phys. Rev. **106**, 364 (1957).
55. L. Onsager, L. Mittag, and M.J. Stephen, Ann. Phys. (Leipzig) **18**, 71 (1966).
56. R.A. Coldwell-Horsfall and A.A. Maradudin, J. Math. Phys. **1**, 395 (1960).
57. D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
58. J.P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).
59. J.P. Perdew and T. Datta, Phys. Stat. Sol. (b) **102**, 283 (1980).
60. C. Fiolhais, J.P. Perdew, S.Q. Armster, J.M. MacLaren, and M. Brajczewska, Phys. Rev. B **51**, 14001 (1995), ibid., **53**, 13193 (1996)(E).
61. E. Engel and S.H. Vosko, Phys. Rev. B **42**, 4940 (1990).
62. S. Moroni, D.M. Ceperley, and G. Senatore, Phys. Rev. Lett. **75**, 689 (1995).
63. K. Burke, J.P. Perdew, and D.C. Langreth, Phys. Rev. Lett. **73**, 1283 (1994).
64. K. Burke, J.P. Perdew, and M. Ernzerhof, in preparation.
65. M. Ernzerhof, in *Macmillan Encyclopedia of Physics* (Macmillan Publishing Company, New York, 1996), Vol. 2, p. 733.
66. J.P. Perdew, A. Savin, and K. Burke, Phys. Rev. A **51**, 4531 (1995).
67. J.P. Perdew, M. Ernzerhof, K. Burke, and A. Savin, Int. J. Quantum Chem. Symp. **61**, 197 (1997).
68. J.P. Perdew and M. Ernzerhof, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J.F. Dobson, G. Vignale, and M. P. Das (Plenum Press, New York, 1997).
69. S.-K. Ma and K.A. Brueckner, Phys. Rev. **165**, 18 (1968).
70. A. Zupan, K. Burke, M. Ernzerhof, and J.P. Perdew, J. Chem. Phys. **106**, 10184 (1997).
71. A. Zupan, J.P. Perdew, K. Burke, and M. Causá, Int. J. Quantum Chem. **61**, 287 (1997).
72. D.A. Kirzhnits, Sov. Phys. JETP **5**, 64 (1957).
73. L.J. Sham, in *Computational Methods in Band Theory*, edited by P.M. Marcus, J.F. Janak, and A.R. Williams (Plenum Press, New York, 1971), p. 458.
74. M. Rasolt and D.J.W. Geldart, Phys. Rev. B **34**, 1325 (1986).
75. M. Rasolt and H.L. Davis, Phys. Lett. A **86**, 45 (1981).
76. Y. Wang and J.P. Perdew, Phys. Rev. B **43**, 8911 (1991).
77. P.A. Antoniewicz and L. Kleinman, Phys. Rev. B **31**, 6779 (1985).
78. D.C. Langreth and S.H. Vosko, Phys. Rev. Lett. **59**, 497 (1987), ibid. **60**, 1984 (1988).
79. J.P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16533 (1996).
80. Y. Wang, J.P. Perdew, J.A. Chevary, L.D. MacDonald, and S.H. Vosko, Phys. Rev. A **41**, 78 (1990).
81. J.P. Perdew, Phys. Lett. A **165**, 79 (1992).
82. Z. Yan, J.P. Perdew, T. Korhonen, and P. Ziesche, Phys. Rev. A **55**, 4601 (1997).
83. D.C. Langreth and J.P. Perdew, Phys. Rev. B **21**, 5469 (1980).
84. R. Colle and O. Salvetti, Theoret. Chim. Acta **37**, 329 (1975).
85. J.P. Perdew, H.Q. Tran, and E.D. Smith, Phys. Rev. B **42**, 11627 (1990).
86. S. Ivanov and M. Levy, submitted to J. Chem. Phys.
87. J.P. Perdew and K. Burke, Int. J. Quantum Chem. Symp. **27**, 309 (1996).
88. J.C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974).
89. H. Lee, C. Lee, and R.G. Parr, Phys. Rev. A **44**, 768 (1991).
90. K. Burke, J.P. Perdew, and M. Levy, in *Modern Density Functional Theory: A Tool for Chemistry*, Vol. 2 of *Theoretical and Computational Chemistry*, edited by J.M. Seminario and P. Politzer (Elsevier, Amsterdam, 1995), p. 29.
91. A.D. Becke, A. Savin, and H. Stoll, Theoret. Chim. Acta **91**, 147 (1995).
92. R. Merkle, A. Savin, and H. Preuss, J. Chem. Phys. **97**, 9216 (1992).
93. V. Tschinke and T. Ziegler, J. Chem. Phys. **93**, 8051 (1990).