Kohn-Sham potentials in density functional theory

Robert van Leeuwen

October 18, 1994

Contents

Chapter 1

General introduction

1.1 Introduction

The description of the physical properties of interacting many-particle systems has been one of the most important goals of physics during this century. The problem is to derive the properties of many-particle systems from the quantum mechanical laws of nature. This requires the solution of a partial differential equation (the Schrödinger or Dirac equation) of $3N$ spatial variables and N spin variables (for electrons) where N is the number of particles in the system. For atoms the number of electrons is in the range $N \sim 1 - 100$. Even small molecules have often more than 100 electrons and for a solid one has $N \sim 10^{23}$. It is clear that the problem cannot be solved without making approximations somewhere along the line. However 'solving' the Schrödinger or Dirac equation is only part of the many-particle problem. The most important objective is to understand and predict the properties of many-particle systems and to calculate the several measurable quantities, such as the bonding energy, polarizability, conductivity,etc.. , rather than the wave function itself.

One way to circumvent the complete solution of the many-particle problem is to construct model Hamiltonians containing only the physics one is interested in. A large number of model Hamiltonians is in use. One has for instance the Hückel method or tight-binding model, the Hubbard and Heisenberg Hamiltonians, the BCS-model etc.. to mention just a few (see for instance [1]). These approximate methods have been quite succesful in explaining a large range of physical phenomena ranging from bonding mechanisms in chemistry to ferromagnetism and superconductivity in solid state physics. However if one is looking for systematic improvements of these approximate models one still has to face the many-body problem.

A systematic way of improvement can be achieved for finite systems such as atoms and molecules by variational methods or perturbation theory. A very simple method is the configuration interaction method. One uses an approximate wave function with many parameters (sometimes millions!) which are expansion coefficients of Slater determinants, and uses the variational principle to minimize the energy with respect to these parameters. The minimizing wave function then approximates the real wave function. The method is however not applicable to infinite systems. Also several variants of perturbation theory are available. A rather elegant formalism, which has been developed in the 50's, is the Greens function technique [2, 3, 4]. Instead of the wave function one uses the one- and two-particle Greens functions to calculate the measurable quantities. The method can in some cases also be applied to infinite systems, for instance quantum liquids or the electron gas. The disadvantage of the above methods is that their application when possible to inhomogeneous systems like atoms, molecules, solids and surfaces require a large computational effort.

A different formalism which is still exact and which can treat the systems mentioned with less

computational effort is the density functional method [5, 6, 7, 8, 9] . The formalism has been developed in the 60's by Hohenberg and Kohn [10] and Kohn and Sham [11]. The Kohn-Sham approach to density functional theory (DFT) allows an exact description of the interacting many-particle systems in terms of an effective non-interacting particle system. The effective potential in this non-interacting particle system (the Kohn-Sham system) can be shown to be completely determined by the electron density of the interacting system, and is for this reason called a density functional. In particular the ground state energy of the system is a density functional. Exact expressions for this functional are, due to the complicated nature of the many-body problem, not known. However over the years, due to a great amount of thoughtful work, more accurate ,more 'physical' and more practical approximations have continued to appear.

The Kohn-Sham method has been used in solid state physics for about thirty years. By now, largely due to the development of increasingly accurate density functionals, the method has also gained a large popularity among quantum chemists, especially as it allows in many cases accurate treatments of molecular systems unattainable by the more traditional quantum mechanical methods.

The review in the following chapter gives a basic introduction to density functional theory and the Kohn-Sham method. The choice of the subjects in this review has to a large extent been influenced by the subjects of the thesis work and personal interest. It therefore treats among others exchange-correlation holes and pair-correlation functions, the OPM-model, scaling properties, and generalized gradient approximations.

Two further chapters on the more mathematical aspects of density functional theory are included. The first one is an introduction to functional calculus, the concepts of which are applied to density functional theory of Coulomb systems in the following section. This chapter contains a more rigorous discussion of some mathematical aspects of density functional theory providing a justification for the use of variational equations in DFT which assumes for instance differentiability of functionals. From a physicists point of view these points are technicalities as all electron densities to be considered in nature are smooth and well-behaved and the corresponding energy functional should also be smooth and differentiable, at least at the collection of physical densities. Although plausible, this still is an assumption. The fact that the statement is in fact true has been proved by mathematicians some 20 years after the Hohenberg-Kohn paper providing a sound basis for applications in DFT. The two chapters on functional analysis are hopefully useful for the reader interested in the functional analytical aspects of DFT.

1.2 Overview of the thesis

In this thesis we investigate the properties of the Kohn-Sham effective potential for atomic and molecular systems. We in particular investigate the exchange-correlation potential and its approximations such as the local density approximation (LDA) and the generalized gradient approximations (GGA). We further propose improved approximations for the exchange and exchange-correlation potential (shortly denoted as x-potential and xc-potential) and discuss a method to obtain exchange-correlation energy expressions from approximate potentials using line integrals.

In chapter 5 we discuss the properties of the xc-potential corresponding to the GGA's of Becke and Perdew which are compared to the exact xc-potential. The exact xc-potential is constructed from accurate atomic electron densities by a newly developed iterative method. From this comparison one can clearly see that the approximate GGA's lack some properties satisfied by the exact xc-potential, notably the long range Coulombic asymptotics which affects the highest occupied orbital energies. To satisfy the asymptotic requirement a gradient expression is developed for the xc-potential with the correct asymptotics leading to greatly improved atomic and molecular ioni-

sation energies as determined from the highest occupied Kohn-Sham orbital.

In chapter 6 we give a more extensive discussion of several gradient approximations to the exchange potential. These are Padé approximants which are constructed to satisfy the correct scaling properties, long range asymptotics and atomic shell structure of the exchange potential. The constructed Padé approximants lead to considerable improvements of the existing expressions for the GGA exchange potentials which do not exhibit the right long range asymptotics and which have an unphysical Coulombic divergence close to the atomic nucleus. The exchange energy is calculated from these Padé approximants using the Levy-Perdew relation. It turns out that to obtain good exchange energies this way a good description of the exchange potential in the core region is important.

In chapter 7 we give an analysis of electron correlation and atomic shell structure in terms of local potentials. The analysis carried out leads to a splitup of the xc-potential into two terms, a long range and a short range part, both of which are relatively easy to approximate. An important question is how exchange and correlation effects influence the shape of the xc-potential as it provides insight in the way to construct more accurate density functionals. An important feature of the exchange and exchange-correlation potential is the atomic shell structure. The x- and xc-potential can be written as the sum of two potentials, a smooth long range attractive Coulombic potential and a short range repulsive step-like potential with clear atomic shell structure. The latter potential is constant within the atomic shells but changes rapidly at the atomic shell boundaries. The exchange-correlation energy is completely determined by the long range attractive part of the xc-potential. The atomic shell structure is further clearly reflected in the properties of the exchange and exchange-correlation hole and the conditional probability amplitude. Both quantities are related to the distribution of the other electrons when one electron, called the reference electron, is known to be at a certain position in space. The exchange hole and the conditional probability amplitude undergo rapid changes when the reference electron crosses atomic shell boundaries. As several of the constituents of the Kohn-Sham and xc-potential can be expressed in terms of the conditional probability amplitudes information on the structure of the Kohn-Sham potential can directly be obtained from the study of these probability amplitudes which contain all the exchange and correlation information of the system. For instance the peak structure at the atomic shell boundaries in the so-called kinetic potential which is a part of the functional derivative of the kinetic energy functional is directly related to changes in the conditional probabililty amplitudes when the reference electron crosses atomic shell boundaries. As discussed in chapter 10 the same potential describes the left-right correlation effect in dissociating molecules.

In chapter 7 we noted the step structure in the x- and xc-potential. In chapter 8 this step structure is derived by making an approximation in the functional derivative of the pair correlation function. The result is an equation for the exchange potential which has been derived earlier by Krieger, Li and Iafrate (KLI) in a different way. This potential turns out to be a very accurate approximation to the exact exchange potential. Another important result is an approximate expression for the inverse density response function as it occurs in several places within density functional theory (e.g. in the construction of xc-potentials from electron densities and in the calculation of excitation energies in time-dependent DFT).

In chapter 9 we propose an approximation for the above discussed short range step-like and long range Coulombic part of the exchange potential. The long range part is approximated by a density gradient expression with the correct asymptotics derived from the Becke GGA for the exchange energy. The short range repulsive step-like part is written as a summation of orbital densities with coefficients that depend on the orbital energies. The analytic dependence of these coefficients on the orbital energies is determined by imposing the correct scaling relations, gauge invariance and

reduction to the correct electron gas result for homogeneous systems. The resulting potential turns out to be a practical and accurate representation of the exchange potential and it moreover yields accurate (of Becke GGA quality) values for the exchange energies.

In chapters 10 and 11 we consider molecular systems. In chapter 10 we discuss the properties of the GGA's of Becke and Perdew for the case of molecular dissociation. The Becke-Perdew GGA is widely applied in quantum chemistry as it yields accurate (with an accuracy of about 0.2 eV) bond energies and geometries for molecules. In this chapter we try to obtain more insight in the reasons for this good performance of the gradient corrected functionals. We first discuss the properties of the exact Kohn-Sham potential in the bond midpoint region. As has been noted before the Kohn-Sham potential in the bond midpoint region exhibits a peak structure related to the left-right correlation effect. This peak structure arises from the kinetic part of the exchange-correlation functional and can be expressed in terms of the conditional probability amplitude. We show that this peak structure is also present in Becke-Perdew GGA however with a wrong functional dependence on the electron density. We furthermore explain why the Becke GGA for the exchange in general yields repulsive contributions to the bond energy and why the Perdew GGA for correlation yields an attractive contribution. This is related to the fact that density gradients at the saddle point in the electron density (the bond midpoint) are zero. Numerical results are presented for the dissociating nitrogen molecule and contributions from different regions of space are analyzed.

In chapter 11 we discuss exact and approximate (the Becke-Perdew GGA) density functional results for the dissociating hydrogen molecule. The hydrogen molecule is chosen for the analysis as the exact DFT quantities needed for comparison with the approximate results can be calculated from accurate electron densities obtained from large configuration interaction calculations on this molecule. A disadvantage however is that this molecule is not a 'standard' molecule in the sense that is does not have core orbitals and therefore it lacks the Pauli repulsion contribution to the bond energy. The DFT results are further compared to the results of the Hartree-Fock approximation and the configuration interaction method which yields almost exact results. It turns out that the DFT results with GGA's included give a very accurate representation of the binding curve of the hydrogen molecule. The DFT curve deviates from the exact curve however in the dissociation limit, which can be cured by performing an unrestricted Kohn-Sham calculation. The DFT dissociation error is however much smaller than the error in the Hartree-Fock approximation. This is due to the fact that the LDA and GGA exchange-correlation hole is localized around the reference electron and resembles much more the true exchange-correlation hole than the delocalized Hartree-Fock exchange hole. The fact that there is still a dissociation error in DFT is due to the fact that the approximate density functionals are not invariant under spin rotations, which leads to problems for degenerate ground states (which are in practice solved by performing unrestricted calculations). We finally analyze the bond energy contribution of the Becke-Perdew GGA for certain density variations.

In chapter 12 we discuss the problem how to calculate the exchange-correlation energy if only the exchange-correlation potential is known. From the exchange potential one can calculate the exchange energy using the Levy-Perdew relation which can be derived using the scaling properties of the exchange potential. For the correlation potentials such scaling relations are however not known. However as we show the correlation energy can be obtained from line integrals in the space of electron densities. Path dependence of the line integrals is discussed and integrability conditions are presented. We further discuss how several invariance properties of the xc-potential can be derived by selecting certain paths corresponding to certain symmetries (rotation, translation, scaling) and derive some inequalities for the correlation potential. We then show that the line integral formalism generalizes the transition state method of Ziegler for the calculation of molecular bond energies.

Chapter 2

A review of density functional theory

2.1 Key concepts and formulas

The evolution of a system of N particles moving in a external potential $v(\mathbf{r})$ and having internal particle interactions $w(\mathbf{r}_1\mathbf{r}_2)$ is in many cases where relativistic effects are not important to a very good approximation determined by the Schrödinger equation

$$H(\mathbf{r}_1,\ldots,\mathbf{r}_N)\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N,t) = i\partial_t\Psi(\mathbf{r}_1\ldots\mathbf{r}_N,t) \tag{2.1}$$

where the Hamiltonian H is given by

$$H(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \sum_i^N -\frac{1}{2}\nabla_i^2 + v(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j} w(\mathbf{r}_i\mathbf{r}_j) \tag{2.2}$$

In the above equations we neglected the internal degrees of freedom of the particles such as spin. For many-electron systems such as atoms, molecules or solids we have to take the electron spin into account and the stationary states of the system are determined from the time-independent Schrödinger equation

$$H(\mathbf{r}_1\ldots\mathbf{r}_N)\Psi(\mathbf{r}_1\sigma_1\ldots\mathbf{r}_N\sigma_N) = E\Psi(\mathbf{r}_1\sigma_1\ldots\mathbf{r}_N\sigma_N) \tag{2.3}$$

where σ_i are the spin coordinates of the electrons and we require the wave function of the system to be antisymmetric under interchange of space and spin variables. The Hamiltonian can alternatively be written in second quantized notation as

$$\hat{H} = \hat{T} + \hat{V} + \hat{W} \tag{2.4}$$

where

$$\hat{T} = -\frac{1}{2}\sum_\sigma \int \hat{\psi}_\sigma^+(\mathbf{r})\nabla^2\hat{\psi}_\sigma(\mathbf{r})d\mathbf{r} \tag{2.5}$$

$$\hat{V} = \sum_\sigma \int \hat{\psi}_\sigma^+(\mathbf{r})v(\mathbf{r})\hat{\psi}_\sigma(\mathbf{r})d\mathbf{r} \tag{2.6}$$

$$\hat{W} = \sum_{\sigma\sigma'} \int \hat{\psi}_\sigma^+(\mathbf{r})\hat{\psi}_{\sigma'}^+(\mathbf{r}')w(\mathbf{r},\mathbf{r}')\hat{\psi}_{\sigma'}(\mathbf{r}')\hat{\psi}_\sigma(\mathbf{r})d\mathbf{r}d\mathbf{r}' \tag{2.7}$$

We further denote $x_i = (\mathbf{r}_i\sigma_i)$. The action of the operators $\hat{\psi}_\sigma^+(\mathbf{r})$ and $\hat{\psi}_\sigma(\mathbf{r})$ on a N-electron state Ψ_N in coordinate space is defined as

$$[\hat{\psi}_\sigma^+(\mathbf{r})\Psi]_{N+1}(x_1\ldots x_N) = \frac{1}{\sqrt{N+1}}\hat{A}[\delta(x-x_{N+1})\Psi_N(x_1\ldots x_N)] \tag{2.8}$$

9

where

$$\delta(x - x_i) = \delta(\mathbf{r} - \mathbf{r}_i)\delta_{\sigma\sigma_i} \tag{2.9}$$

and

$$[\hat{\psi}_\sigma(\mathbf{r})\Psi]_{N-1}(x_1 \ldots x_{N-1}) = \sqrt{N}\Psi_N(x_1 \ldots x_{N-1}, x) \tag{2.10}$$

where $\hat{A}$ is the antisymmetrization operator defined as

$$(\hat{A}\Psi)(x_1 \ldots x_N) = \sum_{P \in S_N} \text{sign}(P)\Psi(x_{P(1)} \ldots x_{P(N)}) \tag{2.11}$$

where the summation runs over all permutations P of the group of permutations S_N of N elements. The operator $\hat{\psi}_\sigma^+(\mathbf{r})$ adds a particle with spin σ at position $\mathbf{r}$ to the N-particle state Ψ_N and operator $\hat{\psi}_\sigma(\mathbf{r})$ removes a particle with spin σ from the N-particle state Ψ_N. One can readily verify using the above definitions that the Hamiltonians 2.3 and 2.4 are equivalent. One can furthermore derive the anticommutation relations

$$\left[\hat{\psi}_\sigma(\mathbf{r}), \hat{\psi}_{\sigma'}(\mathbf{r}')\right]_+ = \left[\hat{\psi}_\sigma^+(\mathbf{r}), \hat{\psi}_{\sigma'}^+(\mathbf{r}')\right]_+ = 0 \tag{2.12}$$

$$\left[\hat{\psi}_\sigma(\mathbf{r}), \hat{\psi}_{\sigma'}^+(\mathbf{r}')\right]_+ = \delta(\mathbf{r} - \mathbf{r}')\delta_{\sigma\sigma'} \tag{2.13}$$

where $[A, B]_+ = AB + BA$.

We now define the electron density as

$$\rho(\mathbf{r}) = \langle\Psi|\hat{\rho}(\mathbf{r})|\Psi\rangle = \langle\Psi|\sum_\sigma \hat{\psi}_\sigma^+(\mathbf{r})\hat{\psi}_\sigma(\mathbf{r})|\Psi\rangle = N\sum_\sigma \int |\Psi(\mathbf{r}\sigma x_2 \ldots x_N)|^2 dx_2 \ldots dx_N \tag{2.14}$$

where dx denotes integration over the spatial variable $\mathbf{r}$ and summation over the spin variable σ. The electron density $\rho(\mathbf{r})$ is proportional to the probability per unit volume of finding an electron at position $\mathbf{r}$.

2.2   Density matrices and density functionals

An important method of calculating ground state properties of many-particle systems is the variational method which is based on the variational principle which states in the case of fermion systems that for any anti-symmetric wavefunction Ψ :

$$E[\Psi] = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E_0 \tag{2.15}$$

where E_0 is the ground state energy of the system. This can be derived assuming $\hat{H}$ has a lowest eigenvalue which is almost always true for all (non-relativistic) cases of interest. This means that the ground state wavefunction is a stationary point of the energy functional $E[\Psi]$, so we have

$$0 = \frac{\delta E}{\delta\Psi(x_1 \ldots x_N)} = \frac{1}{\langle\Psi|\Psi\rangle}\left[\frac{\delta\langle\Psi|\hat{H}|\Psi\rangle}{\delta\Psi(x_1 \ldots x_N)} - E[\Psi]\frac{\delta\langle\Psi|\Psi\rangle}{\delta\Psi(x_1 \ldots x_N)}\right]$$

$$= \frac{1}{\langle\Psi|\Psi\rangle}\left[\hat{H}\Psi^*(x_1 \ldots x_N) - E\Psi^*(x_1 \ldots x_N)\right] \tag{2.16}$$

which is just Schrödingers equation as expected.

From the expression of the Hamiltonian 2.4 it follows that we can write the expectation value of $\hat{H}$ with state $|\Psi\rangle$ as

$$\langle\Psi|\hat{H}|\Psi\rangle = -\frac{1}{2}\sum_\sigma\int\nabla_r^2\gamma_\sigma(\mathbf{r},\mathbf{r}')|_{r=r'}\,d\mathbf{r} + \sum_\sigma\int\rho_\sigma(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \frac{1}{2}\sum_{\sigma\sigma'}\int\Gamma_{\sigma\sigma'}(\mathbf{r},\mathbf{r}')w(\mathbf{r},\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad (2.17)$$

In the above expression we have defined the one-particle density matrix γ_σ as

$$\gamma_\sigma(\mathbf{r},\mathbf{r}') = \langle\Psi|\hat{\psi}_\sigma^+(\mathbf{r}')\hat{\psi}_\sigma(\mathbf{r})|\Psi\rangle = N\int\Psi^*(\mathbf{r}'\sigma x_2\ldots x_N)\Psi(\mathbf{r}\sigma x_2\ldots x_N)dx_2\ldots x_N \qquad (2.18)$$

and the diagonal two-particle density matrix $\Gamma_{\sigma\sigma'}$ as

$$\begin{aligned}\Gamma_{\sigma\sigma'}(\mathbf{r},\mathbf{r}') &= \langle\Psi|\hat{\psi}_\sigma^+(\mathbf{r})\hat{\psi}_{\sigma'}^+(\mathbf{r}')\hat{\psi}_{\sigma'}(\mathbf{r}')\hat{\psi}_\sigma(\mathbf{r})|\Psi\rangle \\ &= N(N-1)\int|\Psi(\mathbf{r}\sigma,\mathbf{r}'\sigma'x_3\ldots x_N)|^2 dx_3\ldots dx_N \end{aligned} \qquad (2.19)$$

A general M-particle density matrix (for an overview of the properties of density matrices see reference [12]) is defined as:

$$D_M(\mathbf{r}_1\sigma_1\ldots\mathbf{r}_M\sigma_M;\mathbf{r}_M\sigma_1'\ldots\mathbf{r}_M'\sigma_M') = \langle\Psi|\hat{\psi}_{\sigma_1'}^+(\mathbf{r}_1')\ldots\hat{\psi}_{\sigma_M'}^+(\mathbf{r}_M')\hat{\psi}_{\sigma_1}(\mathbf{r}_1)\ldots\hat{\psi}_{\sigma_M}(\mathbf{r}_M)|\Psi\rangle$$

$$= \frac{N!}{(N-M)!}\int\Psi^*(\mathbf{r}_1'\sigma_1'\ldots\mathbf{r}_M'\sigma_M' x_{M+1}\ldots x_N)\Psi(\mathbf{r}_1\sigma_1\ldots\mathbf{r}_M\sigma_M x_{M+1}\ldots x_N)dx_{M+1}\ldots dx_N \quad (2.20)$$

In systems were there are only two-particle interactions present we can restrict ourselves to the two-particle density matrix. To calculate the energy of the system we then only need the one-particle density matrix and the diagonal two-particle density matrix which are related to D_2 by

$$\Gamma_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2) = D_2(\mathbf{r}_1\sigma_1,\mathbf{r}_2\sigma_2;\mathbf{r}_1\sigma_1,\mathbf{r}_2\sigma_2) \qquad (2.21)$$

$$\gamma_\sigma(\mathbf{r}_1,\mathbf{r}_1') = D_1(\mathbf{r}_1\sigma,\mathbf{r}_1'\sigma) = \frac{1}{N-1}\sum_{\sigma_2}\int D_2(\mathbf{r}_1\sigma\mathbf{r}_2\sigma_2;\mathbf{r}_1'\sigma_1\mathbf{r}_2\sigma_2)d\mathbf{r}_2 \qquad (2.22)$$

One could therefore wonder if one could derive a variational principle for density matrices instead of wavefunctions. This is readliy done by defining

$$E[D_M] = \inf_{\Psi\to D_M}\langle\Psi|\hat{H}|\Psi\rangle \qquad (2.23)$$

where the infimum of the expectation value is searched over all normalized anti-symmetric wavefunctions Ψ yielding D_M (denoted $\Psi\to D_M$). This means that $E[D_M]$ is defined on the set $\mathcal{D}_M$ of M-particle density matrices obtainable from normalized antisymmetric wavefunctions, i.e

$$\mathcal{D}_M = \{D_M|\exists\Psi:\Psi\to D_M,\langle\Psi|\Psi\rangle = 1,\Psi\text{anti-symmetric}\} \qquad (2.24)$$

One has the variational property:

$$E[D_M]\geq E_0 \text{ if } D_M\in\mathcal{D}_M \qquad (2.25)$$

An arbitrary function $D_M(x_1\ldots x_M;x_1'\ldots x_M')$ will in general not belong to the set $\mathcal{D}_M$. One might therefore ask what the necessary constraints are for a function D_M to belong to $\mathcal{D}_M$. This problem, which is known as the N-representability problem for M-particle density matrices, has received considerable attention for the cases $M=1$ and $M=2$ [13]. The reason for this is that

equation 2.17 gives an explicit expression for the total energy in terms of D_2 and can therefore be applied for variational calculations. However to apply this formula one needs the constraints for D_2 to belong to $\mathcal{D}_2$ which can then be incorporated in the calculation by means of Lagrange multipliers. If one minimizes without these constraints the energy expression one might get lower than E_0. However the N-representability problem for two-particle density matrices has not been solved until now.

The N-representability problem for one-particle density matrices on the other hand has been solved [14] and we will therefore have a look at one-particle density matrix functionals. The energy for the one-particle density matrix can be written as

$$E[\gamma] = -\frac{1}{2} \sum_\sigma \int \nabla_r^2 \gamma_\sigma(\mathbf{r}, \mathbf{r}')|_{r=r'} d\mathbf{r} + \sum_\sigma \int \rho_\sigma(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + W[\gamma] \tag{2.26}$$

where $W[\gamma]$ is defined as

$$W[\gamma] = \inf_{\Psi \to \gamma} \langle \Psi | \hat{W} | \Psi \rangle \tag{2.27}$$

A very nice property of the functional $W[\gamma]$ is that it is defined independent of the external potential v. It can therefore in principle (but not in practice) be calculated once and for all independent of v. For practical applications one needs to know the necessary conditions for γ to belong to $\mathcal{D}_1$. These conditions can be formulated as constraints on the eigenvalues $\mu_{i\sigma}$ of γ_σ. If we have

$$\int \gamma_\sigma(\mathbf{r}_1, \mathbf{r}_2) \phi_{i\sigma}(\mathbf{r}_2) d\mathbf{r}_2 = \mu_{i\sigma} \phi_{i\sigma}(\mathbf{r}_2) \tag{2.28}$$

then the set $\mathcal{D}_1$ is equal to

$$\mathcal{D}_1 = \{\gamma | \int \gamma_\sigma(\mathbf{r}, \mathbf{r}) d\mathbf{r} = N_\sigma, 0 \le \mu_{i\sigma} \le 1\} \tag{2.29}$$

The $\mu_{i\sigma}$ are called the natural occupation numbers and the orbitals $\phi_{i\sigma}$ are called the natural orbitals. The fact that $0 \le \mu_{i\sigma} \le 1$ is a consequence of the Pauli-principle, any orbital ϕ_σ can be occupied with maximally one electron with spin σ. For the application of the variation theorem one needs approximate expressions for $W[\gamma]$. One of the simplest expressions follows by considering a γ obtainable from a Slater determinant which yields

$$W[\gamma] = \frac{1}{2} \sum_\sigma \int w(\mathbf{r}_1, \mathbf{r}_2)[\gamma_\sigma(\mathbf{r}_1, \mathbf{r}_1) \gamma_\sigma(\mathbf{r}_2, \mathbf{r}_2) - \gamma_\sigma(\mathbf{r}_1, \mathbf{r}_2) \gamma_\sigma(\mathbf{r}_2, \mathbf{r}_1)] d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.30}$$

As one can prove [15] that for this approximate $W[\gamma]$ the energy is always minimized by a γ_σ obtainable from a Slater determinant wave function the variational equations corresponding to this $W[\gamma]$ will be equal to the Hartree-Fock equations to be discussed in a later section. Some approximations which go beyond the one proposed here are known [16, 17] but the variational equations are not solved.

An other approach for which approximations to the functionals are easier to obtain is the density functional approach. In the spirit of the lines described above one can define an energy density functional by

$$E[\rho] = \int \rho(\mathbf{r}) v(\mathbf{r}) + F_{LL}[\rho] \tag{2.31}$$

where the functional F_{LL} introduced by Levy and Lieb is [18, 19] defined as

$$F_{LL}[\rho] = \inf_{\Psi \to \rho} \langle \Psi | \hat{T} + \hat{W} | \Psi \rangle \tag{2.32}$$

where the infimum is searched over all normalized antisymmetric wavefunctions giving electron density ρ. The infimum can actually be shown to be a minimum [19]. As for any density integrating to N electrons and which satisfies $\rho \geq 0$ one can easily construct a wavefunction (even a Slater determinant) which yields ρ the densities in the set

$$\mathcal{N} = \{\rho | \rho \geq 0, \int \rho(\mathbf{r}) d\mathbf{r} = N\} \tag{2.33}$$

satisfy the N-representability conditions. With the introduction of a Kohn-Sham system accurate approximate expressions for the functional F_{LL} can be found. This leads in practice to the solution of a system of equations for noninteracting particles moving in an effective external potential.

2.3 The pair-correlation function and the exchange-correlation hole

Important insight into the behaviour of interacting many-particle systems can be obtained by the analysis of correlation functions. Especially insightful are the pair-correlation function $g_\sigma(\mathbf{r}_1, \mathbf{r}_2)$ and the exchange-correlation hole function $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ Extensive discussions of the exchange-correlation hole function have been given by Slater [20] and McWeeny [21]. We define the function

$$P_{\sigma_1 \sigma_2}(\mathbf{r}_2 | \mathbf{r}_1) = \frac{\Gamma_{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2)}{\rho_{\sigma_1}(\mathbf{r}_1)} \tag{2.34}$$

The function $P_{\sigma_1 \sigma_2}(\mathbf{r}_2 | \mathbf{r}_1)$ gives the probability to find an electron with spin σ_2 at $\mathbf{r}_2$ if we know that there is an electron at $\mathbf{r}_1$ with spin σ_1. We define the exchange-correlation hole function $\rho_{xc}^{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2)$ by the equation

$$P_{\sigma_1 \sigma_2}(\mathbf{r}_2 | \mathbf{r}_1) = \rho_{\sigma_2}(\mathbf{r}_2) + \rho_{xc}^{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2) \tag{2.35}$$

As the presence of the reference electron at $\mathbf{r}_1$ reduces the probability to find an electron near $\mathbf{r}_1$ the function ρ_{xc} represents a hole in the electron density $\rho_\sigma(\mathbf{r}_2)$. It has the property

$$\sum_{\sigma_2} \int \rho_{xc}^{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = -1 \tag{2.36}$$

The hole therefore contains one electron. The electron repulsion energy between the electrons may now be written

$$\langle \Psi | \hat{W} | \Psi \rangle = \frac{1}{2} \int \frac{\rho(\mathbf{r}_1) \rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + W_{xc} \tag{2.37}$$

where $\hat{W}$ is the Coulombic repulsive potential between the electrons and the exchange-correlation energy is given by

$$W_{xc} = \frac{1}{2} \sum_{\sigma_1 \sigma_2} \int \frac{\rho_{\sigma_1}(\mathbf{r}_1) \rho_{xc}^{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 = \frac{1}{2} \sum_{\sigma_1} \int \rho_{\sigma_1}(\mathbf{r}_1) v_{xc,\sigma_1}^h(\mathbf{r}_1) d\mathbf{r}_1 \tag{2.38}$$

where $v_{xc,\sigma}^h(\mathbf{r}_1)$ is the potential of the exchange-correlation hole of the reference electron at $\mathbf{r}_1$ defined as

$$v_{xc,\sigma_1}^h(\mathbf{r}_1) = \sum_{\sigma_2} \int \frac{\rho_{xc}^{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.39}$$

For neutral finite systems such as atoms the asymptotic behaviour of $v_{xc,\sigma}^h$ at large distance r from the system decays Coulombically as $-1/r$. If we move the reference electron to infinity then, for instance for an atom, the probability to find an electron near the atomic nucleus is reduced. Another way of putting this is to say that the reference electron experiences the potential of the positive ion it leaves behind. Let us discuss the exchange-correlation hole function a bit further. The probability to find another electron near the reference electron is reduced for two different reasons, one reason is the Pauli-principle which only affects electrons with like spin, and the other reason is the Coulomb repulsion which is spin-independent. The Pauli principle forbids, even if there where no Coulomb interactions between the electrons, electrons with the same spin to occupy the same position in space. This is called the exchange effect. The corresponding hole function $\rho_x^{\sigma_1\sigma_2}$ is called the exchange-hole function. For a non-interacting system of electrons the exchange-hole function is readily calculated. The wave function of such a system is just a Slater determinant

$$\Psi_D(\mathbf{r}_1\sigma_1\dots\mathbf{r}_N\sigma_N) = \frac{1}{\sqrt{N!}}|\phi_1(\mathbf{r}_1\sigma_1)\dots\phi_N(\mathbf{r}_N\sigma_N)| \tag{2.40}$$

We find

$$\rho_x^{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2) = -\frac{|\gamma_{\sigma_1}(\mathbf{r}_1,\mathbf{r}_2)|^2}{\rho_{\sigma_1}(\mathbf{r}_1)}\delta_{\sigma_1\sigma_2} \tag{2.41}$$

where

$$\gamma_{\sigma_1} = \sum_i^{N_{\sigma_1}} \phi_i(\mathbf{r}_1\sigma_1)\phi_i^*(\mathbf{r}_2\sigma_1) \tag{2.42}$$

So we see that the hole for unlike spin electrons is zero as expected. If the reference electron at $\mathbf{r}_1$ is well-located within orbital ϕ_i and the other orbital densities are small at $\mathbf{r}_1$ then $\rho_{\sigma_1}(\mathbf{r}_1) \approx |\phi_i(\mathbf{r}_1)|^2$ and we find approximately

$$\rho_x^{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2) \approx -|\phi_i(\mathbf{r}_2\sigma_2)|^2\delta_{\sigma_1\sigma_2} \tag{2.43}$$

The exchange-correlation hole for an interacting system of electrons is much more difficult to calculate. Note however that to calculate W_{xc} or the hole potential v_{xc}^h we only need the spherical averaged part $\bar{\rho}_{xc}$ of the exchange-correlation hole, defined as

$$\bar{\rho}_{xc}^{\sigma_1\sigma_2}(\mathbf{r}_1,s) = \int \rho_{xc}^{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_1+\mathbf{s})d\Omega_s \tag{2.44}$$

where $d\Omega_s$ is the integration volume of the angular variables for $\mathbf{s}$. The exchange-correlation hole potential becomes

$$v_{xc,\sigma_1}^h(\mathbf{r}_1) = \sum_{\sigma_2} \int_0^\infty 4\pi s^2 \frac{\bar{\rho}_{xc}^{\sigma_1\sigma_2}(\mathbf{r}_1,s)}{s}ds \tag{2.45}$$

Another useful function for the description of electron correlation in many-electron systems is the pair-correlation function $g_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2)$ defined as

$$g_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\Gamma_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2)}{\rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)} \tag{2.46}$$

Then we can write

$$\langle\Psi|\hat{W}|\Psi\rangle = \frac{1}{2}\sum_{\sigma_1\sigma_2}\int \rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)\frac{g_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.47}$$

The function describes the screening of the interparticle interactions due to exchange- and correlation effects, i.e. the $1/r_{12}$ potential is replaced by the screened potential g/r_{12}. The expression for W_{xc} then becomes

$$W_{xc} = \frac{1}{2} \sum_{\sigma_1\sigma_2} \int \rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2) \frac{(g_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2)-1)}{|\mathbf{r}_1-\mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.48}$$

The pair-correlation functions satisfies the sum rule

$$\sum_{\sigma_2} \int \rho_{\sigma_2}(\mathbf{r}_2)(g_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2)-1)d\mathbf{r}_2 = -1 \tag{2.49}$$

And for the exchange-correlation hole potential we have

$$v^h_{xc,\sigma_1}(\mathbf{r}_1) = \sum_{\sigma_2} \int \frac{\rho_{\sigma_2}(\mathbf{r}_2)(g_{\sigma_1\sigma_2}(\mathbf{r}_1,\mathbf{r}_2)-1)}{|\mathbf{r}_1-\mathbf{r}_2|} d\mathbf{r}_2 \tag{2.50}$$

From the sum rule we again find that v^h_{xc} has a Coulombic asymptotic behaviour.

2.4 Effective non-interacting particle models

Important insight in the properties of many-particle systems can be obtained from the study of effective non-interacting particle models. In these models the interacting system of particles is replaced by a non-interacting system of particles in which the external field is replaced by an effective external field which incorporates to some extent the interparticle interactions in an average way. Within the Kohn-Sham approach to DFT one can do this in a formally exact manner. In this section we will however discuss some approximate models, which by their relative ease in computation and by their physically appealling interpretation, have gained a considerable popularity. It turns out that from quite simple models such as the Hückel method in chemistry one can already understand many properties of chemical bonding. With the appearance of more powerful computers more sophisticated approximations, such as the Hartree-Fock approximation, have become quite popular in chemistry.

In the Hartree-Fock approximation the ground state wave function is approximated by a Slater determinant

$$\Psi_D(\mathbf{r}_1\sigma_1\ldots\mathbf{r}_N\sigma_N) = \frac{1}{\sqrt{N!}}|\phi_1(\mathbf{r}_1\sigma_1)\ldots\phi_N(\mathbf{r}_N\sigma_N)| \tag{2.51}$$

The corresponding energy expression is defined as the expectation value of the Hamiltonian with this Slater determinant wave function. The interparticle interaction $\hat{W}$ is taken to be Coulombic.

$$E[\{\phi_i\}] = -\frac{1}{2}\sum_\sigma \int \phi_i^*(\mathbf{r}\sigma)\nabla^2\phi_i(\mathbf{r}\sigma)d\mathbf{r} + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_x[\{\phi_i\}] \tag{2.52}$$

where the electron density is given by

$$\rho(\mathbf{r}) = \sum_\sigma \sum_i |\phi_i(\mathbf{r}\sigma)|^2 \tag{2.53}$$

and the exchange energy by

$$E_x[\{\phi_i\}] = -\frac{1}{2}\int \frac{|\gamma_\sigma(\mathbf{r},\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \tag{2.54}$$

where

$$\gamma_\sigma(\mathbf{r}, \mathbf{r}') = \sum_i^{N_\sigma} \phi_i^*(\mathbf{r}\sigma)\phi_i(\mathbf{r}'\sigma) \tag{2.55}$$

In order to obtain the equations for the effective one-particle orbitals ϕ_i we use the variational theorem and we minimize the functional

$$\Omega[\{\phi_i\}] = E[\{\phi_i\}] - \sum_{ij} \lambda_{ij} \left[\int \phi_i^*(\mathbf{r}\sigma)\phi_j(\mathbf{r}\sigma)d\mathbf{r} - \delta_{ij} \right] \tag{2.56}$$

where the Lagrange multipliers λ_{ij} ensure orthonormality of the orbitals. The Euler-Lagrange equations are then obtained from

$$\frac{\delta\Omega}{\delta\phi_i(\mathbf{r}\sigma)} = 0 \qquad \frac{\delta\Omega}{\delta\phi_i^*(\mathbf{r}\sigma)} = 0 \qquad \frac{\delta\Omega}{\delta\lambda_{ij}} = 0 \tag{2.57}$$

yielding

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}d\mathbf{r}' \right] \phi_i(\mathbf{r}\sigma) - \int \Sigma_{x\sigma}^{HF}(\mathbf{r}, \mathbf{r}')\phi_i(\mathbf{r}'\sigma)d\mathbf{r}' = \epsilon_i\phi_i(\mathbf{r}\sigma) \tag{2.58}$$

where the one-particle energies ϵ_i are obtained from a diagonalization of the matrix of Lagrange multipliers λ_{ij}. The integral kernel $\Sigma_{x\sigma}^{HF}$ is given by

$$\Sigma_{x\sigma}^{HF}(\mathbf{r}, \mathbf{r}') = \frac{\gamma_\sigma(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.59}$$

The above equations constitute the Hartree-Fock equations which is a set of integro-differential equations for the orbitals ϕ_i. Some of the features of the Hartree-Fock model become more clear if we rewrite these equations in a slightly modified form where we replace the integral kernel with a set of orbital dependent potentials

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + v_i^{HF}(\mathbf{r}\sigma) \right] \phi_i(\mathbf{r}\sigma) = \epsilon_i\phi_i(\mathbf{r}\sigma) \tag{2.60}$$

where

$$v_i^{HF}(\mathbf{r}\sigma) = - \int \frac{n_i(\mathbf{r}\sigma, \mathbf{r}'\sigma)}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' \tag{2.61}$$

and

$$n_i(\mathbf{r}\sigma, \mathbf{r}'\sigma) = \frac{\phi_i^*(\mathbf{r}\sigma)\gamma_\sigma(\mathbf{r}, \mathbf{r}')\phi_i(\mathbf{r}'\sigma)}{|\phi_i(\mathbf{r}\sigma)|^2} \tag{2.62}$$

The most important contribution of γ_σ to n_i when ϕ_i is an occupied orbital is the term $\phi_i(\mathbf{r}\sigma)\phi_i^*(\mathbf{r}'\sigma)$ so we find that approximately

$$n_i(\mathbf{r}\sigma, \mathbf{r}'\sigma) \approx |\phi_i(\mathbf{r}'\sigma)|^2 \tag{2.63}$$

so we have

$$v_i^{HF}(\mathbf{r}\sigma) \approx \int \frac{|\phi_i(\mathbf{r}'\sigma)|^2}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' \tag{2.64}$$

So we see that the term n_i for occupied orbitals substracts out the density of orbital ϕ_i. This is something which is intuitively clear. The effective electron should move in the field of the charge cloud of the other electrons, excluding the field of its own charge cloud. For atomic systems with nuclear charge Z this implies that the asymptotic form of the total effective Hartree-Fock potential for occupied orbitals is given by

$$v_{TOT,i}^{HF}(\mathbf{r}\sigma) \sim -\frac{Z}{r} + \int \frac{\rho(\mathbf{r}') - |\phi_i(\mathbf{r}'\sigma)|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \sim -\frac{(Z - N + 1)}{r} \quad (r \to \infty) \tag{2.65}$$

which decays like $-1/r$ for neutral atoms. For unoccupied orbitals we have

$$v_{TOT,i}^{HF}(\mathbf{r}\sigma) \sim -\frac{Z}{r} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \sim -\frac{(Z - N)}{r} \quad (r \to \infty) \tag{2.66}$$

which decays faster than Coulombic for neutral atoms. This last result implies that the virtual spectrum of a neutral atom within the Hartree-Fock approximation does not resemble the Rydberg-like series characteristic for particles bounded by a Coulombic potential. The virtual orbitals (when bounded) of a neutral atom will be quite diffuse. This is of some importance when the virtual Hartree-Fock orbitals are used in more elaborate correlated methods such as configuration interaction expansions.

The Hartree-Fock equations have the computational disadvantage that every effective electron moves in a different potential. To simplify the equations one could therefore approximate the the HF-exchange energy locally by the HF-exchange energy of the homogeneous electron gas yielding

$$E_x = -\alpha \left(\frac{3}{8\pi}\right)^{\frac{1}{3}} \int \rho(\mathbf{r})^{\frac{4}{3}} d\mathbf{r}_1 \tag{2.67}$$

where $\alpha = 2/3$. The variational equations then become

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' - \frac{4}{3}\alpha \left(\frac{3}{8\pi}\right)^{\frac{1}{3}} \rho(\mathbf{r})^{\frac{1}{3}} \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.68}$$

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{2.69}$$

This is a considerable simplification of the Hartree-Fock equations. The main advantage from a computational point of view is that one now does not have to calculate a nonlocal potential which requires large computational effort. In actual applications for atoms, molecules and solids it turns out that the accuracy can be improved by replacing the $\alpha = 2/3$ of the electron gas by $\alpha = 0.7$. The method, originally developed by Slater [20], is known as the $X\alpha$-method and has extensively been used in solid state physics and quantum chemistry due to its simplicity and surprising accuracy which in many cases exceeds the accuracy of the Hartree-Fock approximation. The reason for this is among others, that the well-known molecular dissociation error in Hartree-Fock is considerably reduced in the $X\alpha$-approximation due to the localized character of the $X\alpha$-exchange hole which actually resembles more the true exchange-correlation hole than the exchange hole. A more extensive discussion of the feature in connection with the Local Density Approximation (LDA) is given in section of this chapter.

The $X\alpha$-method shows that it is possible to obtain fairly accurate results within an effective non-interacting particle model with a local potential depending only on the electron density ρ. In the next two sections we will show that the $X\alpha$-model can be regarded as an approximation to an

exact non-interacting particle model, known as the Kohn-Sham system, containing a density dependent local potential. From the solution of the Kohn-Sham equations it is in principle possible to obtain the exact ground state density and energy. In practice of course we have to resort to approximations.

2.5   The Hohenberg-Kohn theorem

In the following we will demonstrate that the ground state energy of any interacting many-particle system with a given fixed interparticle interaction $\hat{W}$ is uniquely determined by its ground state density. Furthermore the knowledge of the ground state density ρ determines the external potential v of the system (to within a trivial constant).

An essential part of the proof of the latter $\rho - v$ relation is the fact that when calculating the expectation value of the Hamiltonian $\hat{H}$ with a certain state $|\Psi\rangle$ a term containing the electron density and the external potential explicitly separates out as

$$\langle\Psi|\hat{V}|\Psi\rangle = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \tag{2.70}$$

For the same reason one can prove from

$$\langle\Psi|\hat{W}|\Psi\rangle = \sum_{\sigma\sigma'} \int \Gamma_{\sigma\sigma'}(\mathbf{r},\mathbf{r}')w(\mathbf{r},\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{2.71}$$

that the two-particle interaction $\hat{W}$ is (to within a constant) a unique functional of the diagonal two-particle density matrix $\Gamma_{\sigma\sigma'}$.

Another essential ingredient of the proof of the HK-theorem is the use of the variational theorem. We will derive the theorem for systems with non-degenerate ground states. The theorem is quite easily extended to the case of degenerate ground states. However to avoid unnecessary complication we will discuss the degenerate case in a later instant.

An external potential $\hat{V}$ leads to a ground state density in an obvious way. First of all $\hat{V}$ leads to a ground state $|\Psi\rangle$ by solving the Schrödinger equation

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{2.72}$$

This provides us with a map $C : \mathcal{V} \mapsto \Phi$ from the set of external potentials $\mathcal{V}$ to the set of ground state wave functions Φ. Secondly, for any ground state wave function one can calculate the corresponding electron density from

$$\rho(\mathbf{r}) = \langle\Psi|\hat{\rho}(\mathbf{r})|\Psi\rangle = N \sum_{\sigma} \int |\Psi(\mathbf{r}\sigma x_2 \ldots x_N)|^2 dx_2 \ldots dx_N \tag{2.73}$$

which gives a second map $D : \Phi \mapsto \mathcal{N}$ from the set of ground state wave functions to the set of ground state densities $\mathcal{N}$. The statement of the Hohenberg-Kohn theorem is then: The map $DC : \mathcal{V} \mapsto \mathcal{N}$ is invertible. To prove this we first show that two different external potentials $\hat{V} \neq \hat{V}' + C$ (with C an arbitrary constant) lead to two different wave functions (that is differing more than a phase factor).

The proof is by contradiction. Suppose we have

$$(\hat{T} + \hat{V}_1 + \hat{W})|\psi_1\rangle = E_1|\psi_1\rangle \tag{2.74}$$

$$(\hat{T} + \hat{V}_2 + \hat{W})|\psi_2\rangle = E_2|\psi_2\rangle \tag{2.75}$$

where $\hat{V}_1 \neq \hat{V}_2 + C$ where C is a constant and let us assume that $|\psi_1\rangle = |\psi_2\rangle = |\psi\rangle$, then by subtraction of these equations we have

$$(\hat{V}_1 - \hat{V}_2)|\psi\rangle = (E_1 - E_2)|\psi\rangle \tag{2.76}$$

As $\hat{V}_1$ and $\hat{V}_2$ are multiplicative operators we must have $\hat{V}_1 - \hat{V}_2 = E_1 - E_2$ which is in contradiction with $V_1 \neq \hat{V}_2 + C$ unless $|\psi\rangle$ vanishes in some region in space. This is however impossible by the unique continuation theorem valid for all reasonably well-behaved potentials (for instance for potentials not containing infinite barriers etc.). From the above contradiction we then conclude that our assumption was wrong so two potentials differing more than a constant yield different ground state wave functions and hence the map C is invertable. A more rigorous discussion of the set of allowable potentials in DFT is presented in chapter 4 of this thesis.

The second part of the proof consists of showing that two different ground state wave functions yield different ground state densities. The proof proceeds again by contradiction. Let us assume that $|\psi_1\rangle$ and $|\psi_2\rangle$ yield the same ground state density ρ. Then we have

$$E_1 = \langle\psi|\hat{H}_1|\psi_1\rangle < \langle\psi_2|\hat{H}_1|\psi_2\rangle = \langle\psi_2|\hat{H}_2 + \hat{V}_1 - \hat{V}_2|\psi_2\rangle = E_2 + \int \rho(\mathbf{r})(v_1(\mathbf{r}) - v_2(\mathbf{r}))d\mathbf{r} \tag{2.77}$$

similarly we find

$$E_2 < E_1 + \int \rho(\mathbf{r})(v_2(\mathbf{r}) - v_1(\mathbf{r}))d\mathbf{r} \tag{2.78}$$

Adding both inequalities then leads to the contradiction

$$E_1 + E_2 < E_1 + E_2 \tag{2.79}$$

and therefore our assumption was wrong, $|\psi_1\rangle$ and $|\psi_2\rangle$ must yield different densities and the map D is invertible. Therefore by our previous result the map DC is also invertable and the ground state density uniquely determines the external potential.

Due to the invertability of the map D every ground state expectation value of an operator $\hat{A}$ is determined by the ground state density

$$A[\rho] = \langle\psi[\rho]|\hat{A}|\psi[\rho]\rangle \tag{2.80}$$

In particular for the ground state energy we find

$$E_v[\rho] = \langle\psi[\rho]|\hat{H}|\psi[\rho]\rangle = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \tag{2.81}$$

where the functional $F_{HK}[\rho]$ is defined as

$$F_{HK}[\rho] = \langle\psi[\rho]|\hat{T} + \hat{W}|\psi[\rho]\rangle \tag{2.82}$$

The functional F_{HK} is universal in the sense that it does not depend on the external potential $\hat{V}$. It can therefore in principle (but not in practice) be calculated for all ground state densities ρ. This is the most important conclusion of the HK-theorem, once approximations for F_{HK} are known we can apply it to all electronic systems ranging from atoms and molecules to solids.

The energy functional E_v satisfies a variational property

$$E_0 = E_{v_0}[\rho_0] = \langle\psi[\rho_0]|\hat{T} + \hat{V}_0 + \hat{W}|\psi[\rho_0]\rangle < \langle\psi[\rho]|\hat{T} + \hat{V}_0 + \hat{W}|\psi[\rho]\rangle = E_{v_0}[\rho] \tag{2.83}$$

if $\rho \neq \rho_0$. We therefore find

$$E_0 = \min_{\rho \in \mathcal{N}} E_{v_0}[\rho] = \min_{\rho \in \mathcal{N}} \left[\int \rho(\mathbf{r})v_0(\mathbf{r})d\mathbf{r} + F_{HK}[\rho]\right] \tag{2.84}$$

A more thorough discussion of the properties of F_{HK} including a discussion of the extension of F_{HK} to a functional F_{LL} defined on a larger set of densities is given in chapter 4 of this thesis.

2.6   The Kohn-Sham equations

To apply the DFT formalism one obviously needs good approximations for the functional F_{HK} or F_{LL}. Accurate approximations have been obtained by means of the Kohn-Sham method [11]. A main advantage of this scheme is that it allows a straightforward determination of a large part of the kinetic energy in a simple way. Another advantage, from a more physical point of view, is that it provides an exact one-particle picture of interacting electronic systems. This then provides a rigorous basis for the one-particle arguments used in solid state physics and chemistry to explain and predict certain features of chemical bonding.

We introduce a non-interacting particle system with Hamiltonian $\hat{H}_s$, ground state density ρ and external potential v_s. The ground state (assumed to be non-degenerate for the moment) is a Slater determinant of orbitals ϕ_i which satisfy the equations (we neglect spin for the moment)

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.85}$$

with

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{2.86}$$

If we apply the Hohenberg-Kohn theorem to this non-interacting system we find that there is at most one external potential v_s to within a constant which generates ρ (one usually chooses this constant in such a way that $v_s \to 0$ for $|\mathbf{r}| \to \infty$.) Therefore, for a given ground state density ρ, all the properties of the system are determined. This is in particular true for the kinetic energy $T_s[\rho]$ and the total energy $E_{v_s}[\rho]$ given by

$$T_s[\rho] = \sum_i^N -\frac{1}{2} \int \phi_i^*(\mathbf{r}_1)\nabla^2\phi_i(\mathbf{r})d\mathbf{r} \tag{2.87}$$

and

$$E_{v_s}[\rho] = T_s[\rho] + \int \rho(\mathbf{r})v_s(\mathbf{r})d\mathbf{r} \tag{2.88}$$

The functional $T_s[\rho]$, which is called the Kohn-Sham kinetic energy, is just a particular case of the functional F_{HK} for the case $\hat{W} = 0$. The ground state density of the system can now equivalently be obtained by the solution of the Euler-Lagrange equations

$$0 = \frac{\delta}{\delta\rho(\mathbf{r})}\left[E_{v_s}[\rho] - \mu_s \int \rho(\mathbf{r})d\mathbf{r} \right] = \frac{\delta T_s}{\delta\rho(\mathbf{r})}[\rho] + v_s(\mathbf{r}) - \mu_s \tag{2.89}$$

where we introduced the Lagrange multiplier μ_s to satisfy the constraint that the density integrates to the correct number of electrons. Let us now go back to the interacting system where the energy functional reads

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \tag{2.90}$$

and let us define the exchange-correlation energy functional $E_{xc}[\rho]$ as

$$E_{xc}[\rho] = F_{HK}[\rho] - \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 - T_s[\rho] \tag{2.91}$$

As F_{HK} is only defined for ground state densities of interacting systems and T_s is only definend for ground state densities of non-interacting systems, we have implicitly assumed that for any ground state density ρ of an interacting system there exists a non-interacting system with the same ground state density. This is the central assertion of the Kohn-Sham scheme. If we assume this assertion to be true then the Euler-Lagrange equation of the interacting system is given by

$$0 = \frac{\delta}{\delta\rho(\mathbf{r})}\left[E_v[\rho] - \mu \int \rho(\mathbf{r})d\mathbf{r}\right] = \frac{\delta T_s}{\delta\rho(\mathbf{r})}[\rho] + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{xc}}{\delta\rho(\mathbf{r})}[\rho] - \mu \qquad (2.92)$$

If we define the exchange-correlation potential v_{xc} by

$$v_{xc}([\rho]; \mathbf{r}) = \frac{\delta E_{xc}}{\delta\rho(\mathbf{r})}[\rho] \qquad (2.93)$$

we find using the fact that the equations for the interacting and the non-interacting system are solved for the same density that

$$v_s(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + v_{xc}([\rho]; \mathbf{r}) \qquad (2.94)$$

to within a constant (we can always choose the arbitrary constant in v_{xc} in such a way that $v_{xc} \to 0$ for $|\mathbf{r}| \to \infty$ and then $\mu - \mu_s = 0$). We therefore find that we can find the ground state density of the interacting system by solving the Kohn-Sham equations

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + v_{xc}([\rho]; \mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \qquad (2.95)$$

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \qquad (2.96)$$

The advantage of the above approach is that we now have reduced the problem of finding good approximations to the functional F_{HK} to finding good approximations for the much smaller quantity E_{xc}.

Note that the kinetic energy of the interacting system $T[\rho] = \langle\psi[\rho]|\hat{T}|\psi[\rho]\rangle$ with $\psi[\rho]$ the ground state wave function of the interacting system with density ρ, is not equal to the kinetic energy of the non-interacting system $T_s[\rho]$ with the same ground state density. It follows from

$$\langle\psi[\rho]|\hat{T}|\psi[\rho]\rangle + \int \rho(\mathbf{r})v_s(\mathbf{r})d\mathbf{r} = \langle\psi[\rho]|\hat{H}_s|\psi[\rho]\rangle$$

$$\geq \langle\psi_s[\rho]|\hat{H}_s|\psi_s[\rho]\rangle = T_s[\rho] + \int \rho(\mathbf{r})v_s(\mathbf{r})d\mathbf{r} \qquad (2.97)$$

where $\psi_s[\rho]$ is the Kohn-Sham determinant and $\hat{H}_s$ is the Kohn-Sham Hamiltonian, that

$$T[\rho] \geq T_s[\rho] \qquad (2.98)$$

and

$$T_s[\rho] = \min_{\psi \to \rho}\langle\Psi|\hat{T}|\Psi\rangle \qquad (2.99)$$

where we minimize over all anti-symmetric and normalized wave functions yielding density ρ. If we define the exchange-correlation part of the kinetic energy as

$$T_{xc}[\rho] = T[\rho] - T_s[\rho] \qquad (2.100)$$

we therefore find

$$T_{xc}[\rho] \geq 0 \tag{2.101}$$

So we have

$$E_{xc}[\rho] = T_{xc}[\rho] + W_{xc}[\rho] \geq W_{xc}[\rho] \tag{2.102}$$

where

$$W_{xc}[\rho] = \langle \psi[\rho] | \hat{W} | \psi[\rho] \rangle - \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.103}$$

Before we will look into approximations for the exchange-correlation energy functional we will first discuss some exact scaling properties of density functionals which give useful constraints on approximate functionals.

2.7 Scaling properties

We will now derive some scaling properties [22] of some density functionals which can be defined in terms of the Kohn-Sham orbitals. Consider the Kohn-Sham equations

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.104}$$

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = 1 \tag{2.105}$$

and define the scaled coordinate $\mathbf{r} = \lambda \mathbf{r}'$ then

$$\nabla_{\mathbf{r}} = \frac{1}{\lambda}\nabla_{\mathbf{r}'} \tag{2.106}$$

and we obtain

$$\left[-\frac{1}{2}\nabla^2_{\mathbf{r}'} + \lambda^2 v_s(\lambda\mathbf{r}) \right] \phi_i(\lambda\mathbf{r}') = \lambda^2 \epsilon_i \phi_i(\lambda\mathbf{r}') \tag{2.107}$$

If we define

$$\phi_{i,\lambda}(\mathbf{r}) = \lambda^{\frac{3}{2}}\phi_i(\lambda\mathbf{r}) \tag{2.108}$$

$$\epsilon_{i,\lambda} = \lambda^2 \epsilon_i \tag{2.109}$$

$$\rho_\lambda(\mathbf{r}) = \lambda^3 \rho(\lambda\mathbf{r}) \tag{2.110}$$

then we obtain

$$\left[-\frac{1}{2}\nabla^2_{\mathbf{r}'} + \lambda^2 v_s(\lambda\mathbf{r}) \right] \phi_{i,\lambda}(\mathbf{r}') = \epsilon_{i,\lambda}\phi_{i,\lambda}(\mathbf{r}') \tag{2.111}$$

$$\sum_i^N |\phi_{i,\lambda}(\mathbf{r})|^2 = \rho_\lambda(\mathbf{r}) \tag{2.112}$$

We therefore conclude that if $v_s(\mathbf{r})$ generates ground state density $\rho(\mathbf{r})$, which we will denote with $v_s([\rho];\mathbf{r})$, then $v_{s,\lambda}(\mathbf{r}) = \lambda^2 v_s(\lambda\mathbf{r})$ generates ground state density $\rho_\lambda(\mathbf{r})$. However as there is by the Hohenberg-Kohn theorem at most one potential which generates ρ_λ we must conclude that

$$v_s([\rho_\lambda];\mathbf{r}) = \lambda^2 v_s([\rho];\lambda\mathbf{r}) \tag{2.113}$$

to within a constant. We further obtain

$$\phi_i([\rho_\lambda];\mathbf{r}) = \lambda^{\frac{3}{2}} \phi_i([\rho];\lambda\mathbf{r}) \tag{2.114}$$

$$\epsilon_i[\rho_\lambda] = \lambda^2 \epsilon_i[\rho] \tag{2.115}$$

This implies that the Kohn-Sham kinetic energy functional satisfies

$$T_s[\rho_\lambda] = \lambda^2 T_s[\rho] \tag{2.116}$$

Let us now define the exchange energy functional as

$$E_x[\rho] = -\frac{1}{4} \int \frac{|\gamma_s(\mathbf{r}_1,\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.117}$$

where

$$\gamma_s(\mathbf{r}_1,\mathbf{r}_2) = \sum_i^N \phi_i(\mathbf{r}_1)\phi_i^*(\mathbf{r}_2) \tag{2.118}$$

is the one-particle density matrix for the Kohn-Sham system. Using the above derived equations we find

$$\gamma_s([\rho_\lambda];\mathbf{r}_1,\mathbf{r}_2) = \sum_i^N \phi_i([\rho_\lambda];\mathbf{r}_1)\phi_i^*([\rho_\lambda];\mathbf{r}_2) =$$

$$= \lambda^3 \sum_i^N \phi_i([\rho];\lambda\mathbf{r}_1)\phi_i^*([\rho];\lambda\mathbf{r}_2) = \lambda^3 \gamma_s([\rho];\lambda\mathbf{r}_1,\lambda\mathbf{r}_2) \tag{2.119}$$

This then yields

$$E_x[\rho_\lambda] = \lambda E_x[\rho] \tag{2.120}$$

The functionals T_s and E_x are special cases of homogeneously scaling functionals of the form

$$A[\rho_\lambda] = \lambda^k A[\rho] \tag{2.121}$$

For these functionals we obviously have the property

$$\frac{d}{d\lambda} A[\rho_\lambda]|_{\lambda=1} = k A[\rho] \tag{2.122}$$

If we further denote

$$a([\rho];\mathbf{r}) = \frac{\delta A}{\delta\rho(\mathbf{r})}[\rho] \tag{2.123}$$

then on the other hand

$$\frac{d}{d\lambda} A[\rho_\lambda]|_{\lambda=1} = \int \frac{\delta A}{\delta\rho_\lambda(\mathbf{r})}[\rho_\lambda]|_{\lambda=1} \frac{d\rho_\lambda(\mathbf{r})}{d\lambda}|_{\lambda=1} d\mathbf{r}$$

$$= \int a([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r}))d\mathbf{r} = -\int \rho(\mathbf{r})\mathbf{r} \cdot \nabla a([\rho]; \mathbf{r})d\mathbf{r} \tag{2.124}$$

where in the last step we performed partial integration. We therefore find

$$A[\rho] = \frac{1}{k} \int a([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r}))d\mathbf{r} \tag{2.125}$$

In particular we find for T_s and E_x that

$$T_s[\rho] = \frac{1}{2} \int v_s([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r}))d\mathbf{r} \tag{2.126}$$

$$E_x[\rho] = \int v_x([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r}))d\mathbf{r} \tag{2.127}$$

where the exchange potential v_x is defined as the functional derivative of E_x

$$v_x([\rho]; \mathbf{r}) = \frac{\delta E_x}{\delta\rho(\mathbf{r})}[\rho] \tag{2.128}$$

This last relation between the exchange potential and exchange energy is often denoted as the Levy-Perdew relation [22]. The scaling property of the functional $A[\rho]$ further implies that

$$a([\rho_\lambda]; \mathbf{r}) = \lambda^k a([\rho]; \lambda\mathbf{r}) \tag{2.129}$$

This we will derive as follows. From

$$\delta A[\rho] = \int a([\rho]; \mathbf{r})\delta\rho(\mathbf{r})d\mathbf{r} \tag{2.130}$$

$$\delta A[\rho_\lambda] = \int a([\rho_\lambda]; \mathbf{r})\delta\rho_\lambda(\mathbf{r})d\mathbf{r} \tag{2.131}$$

$$\delta\rho_\lambda(\mathbf{r}) = \lambda^3 \delta\rho(\lambda\mathbf{r}) \tag{2.132}$$

$$\delta A[\rho_\lambda] = \lambda^k \delta A[\rho] \tag{2.133}$$

it follows that

$$\int a([\rho_\lambda]; \mathbf{r})\delta\rho_\lambda(\mathbf{r})d\mathbf{r} = \delta A[\rho_\lambda] = \lambda^k \delta A[\rho] = \lambda^k \int a([\rho]; \mathbf{r})\delta\rho(\mathbf{r})d\mathbf{r} =$$

$$= \lambda^k \int a([\rho]; \lambda\mathbf{r})\delta\rho(\lambda\mathbf{r})d\lambda\mathbf{r} = \lambda^k \int a([\rho]; \lambda\mathbf{r})\lambda^3 \delta\rho(\lambda\mathbf{r})d\mathbf{r} = \lambda^k \int a([\rho]; \lambda\mathbf{r})\rho_\lambda(\mathbf{r})d\mathbf{r} \tag{2.134}$$

which yields

$$0 = \int \left[a([\rho_\lambda]; \mathbf{r}) - \lambda^k a([\rho]; \lambda\mathbf{r}) \right] \delta\rho_\lambda(\mathbf{r})d\mathbf{r} \tag{2.135}$$

This relation is true for all λ and all variations

$$\int \delta\rho_\lambda(\mathbf{r})d\mathbf{r} = 0 \tag{2.136}$$

This then implies that

$$a([\rho_\lambda]; \mathbf{r}) = \lambda^k a([\rho]; \lambda\mathbf{r}) \tag{2.137}$$

Application of this relation to the functionals T_s and E_x yields [23]

$$v_s([\rho_\lambda]; \mathbf{r}) = \lambda^2 v_s([\rho]; \lambda \mathbf{r}) \tag{2.138}$$

$$v_x([\rho_\lambda]; \mathbf{r}) = \lambda v_x([\rho]; \lambda \mathbf{r}) \tag{2.139}$$

The first equation of the above we alreday derived above. For the correlation energy functional defined as

$$E_c[\rho] = E_{xc}[\rho] - E_x[\rho] \tag{2.140}$$

and its corresponding correlation potential

$$v_c([\rho]; \mathbf{r}) = \frac{\delta E_c}{\rho(\mathbf{r})}[\rho] \tag{2.141}$$

no scaling relations are known. One can however prove the scaling inequalities

$$E_c[\rho_\lambda] < \lambda E_c[\rho] \qquad \lambda < 1 \tag{2.142}$$

$$E_c[\rho_\lambda] > \lambda E_c[\rho] \qquad \lambda > 1 \tag{2.143}$$

Proofs of the above relations and other scaling relations can be found in reference [22].

2.8 The coupling constant integration

The coupling constant integration technique [24, 25, 26] is a useful way of deriving relations among density functionals. It provides a way of expressing the exchange-correlation energy in terms of a coupling constant integrated pair correlation function $\bar{g}$.
We define the Hamiltonian $\hat{H}_\lambda$ by

$$\hat{H}_\lambda = \hat{T} + \hat{V}_\lambda + \lambda \hat{W} \tag{2.144}$$

For $\lambda = 1$ we have the fully interacting system with external potential $\hat{V} = \hat{V}_{\lambda=1}$, having ground state density ρ. For $0 \leq \lambda \leq 1$ we define $\hat{V}_\lambda$ in such a way that the ground state density ρ remains unchanged. For $\lambda = 0$ we then obviously have $\hat{V} = \hat{V}_s$ which is equal to the Kohn-Sham potential corresponding to density ρ. If we denote the ground state for coupling constant λ by $|\psi_\lambda\rangle$ then by the Hellmann-Feynman theorem we obtain

$$\frac{dE}{d\lambda}(\lambda) = \langle \psi_\lambda | \frac{d\hat{H}_\lambda}{d\lambda} | \psi_\lambda \rangle = \langle \psi_\lambda | \hat{W} | \psi_\lambda \rangle + \frac{d}{d\lambda} \int v_\lambda(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \tag{2.145}$$

Therefore

$$E(1) = E(0) + \int_0^1 \langle \psi_\lambda | \hat{W} | \psi_\lambda \rangle d\lambda + \int \rho(\mathbf{r})(v(\mathbf{r}) - v_s(\mathbf{r})) d\mathbf{r}$$

$$= T_s[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \int_0^1 \langle \psi_\lambda | \hat{W} | \psi_\lambda \rangle d\lambda =$$

$$= T_s[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\rho] \tag{2.146}$$

where

$$E_{xc}[\rho] = \frac{1}{2} \sum_{\sigma_1 \sigma_2} \int \frac{\rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} (\bar{g}_{\sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2) - 1)d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.147}$$

Here we defined the coupling constant integrated pair-correlation function $\bar{g}$ as

$$\bar{g}_{\sigma_1 \sigma_2} = \int_0^1 g_{\lambda, \sigma_1 \sigma_2}(\mathbf{r}_1, \mathbf{r}_2) \tag{2.148}$$

where g_λ is the pair-correlation function corresponding to $|\psi_\lambda\rangle$. The relation between E_{xc} and the coupling constant integrated pair-correlation function is useful for constructing approximations to E_{xc} by means of constructing approximations to $\bar{g}$. Useful in this respect are relations between density scaling and coupling constant integration. Suppose that Ψ_λ is the ground state eigenfunction of $\hat{H}_\lambda$ with a ground state density ρ, so we can denote $\Psi_\lambda = \Psi_\lambda[\rho]$. We have (we leave the spin indices out for notational convenience)

$$\left[\sum_i -\frac{1}{2}\nabla_i^2 + v_\lambda(\mathbf{r}_i) + \frac{1}{2}\sum_{i \neq j} \frac{\lambda}{|\mathbf{r}_i - \mathbf{r}_j|} \right] \Psi_\lambda(\mathbf{r}_1 \ldots \mathbf{r}_N) = E\Psi_\lambda(\mathbf{r}_1 \ldots \mathbf{r}_N) \tag{2.149}$$

If we define the scaled coordinate $\mathbf{r} = \alpha\mathbf{r}'$ we obtain

$$\left[\sum_i -\frac{1}{2}\nabla_{\mathbf{r}'_i}^2 + \alpha^2 v_\lambda(\alpha\mathbf{r}'_i) + \frac{1}{2}\sum_{i \neq j} \frac{\alpha\lambda}{|\mathbf{r}'_i - \mathbf{r}'_j|} \right] \Psi_\lambda(\alpha\mathbf{r}'_1 \ldots \alpha\mathbf{r}'_N) = \alpha^2 E\Psi_\lambda(\alpha\mathbf{r}'_1 \ldots \alpha\mathbf{r}'_N) \tag{2.150}$$

If we take $\alpha = 1/\lambda$ and define Φ by

$$\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N) = \lambda^{\frac{3N}{2}} \Phi(\lambda\mathbf{r}_1 \ldots \lambda\mathbf{r}_N) \tag{2.151}$$

Then Φ satisfies

$$\left[\sum_i -\frac{1}{2}\nabla_{\mathbf{r}'_i}^2 + \lambda^{-2}v_\lambda(\lambda^{-1}\mathbf{r}'_i) + \frac{1}{2}\sum_{i \neq j} \frac{1}{|\mathbf{r}'_i - \mathbf{r}'_j|} \right] \Phi(\mathbf{r}'_1 \ldots \mathbf{r}'_N) = \lambda^{-2}E\Phi(\mathbf{r}'_1 \ldots \mathbf{r}'_N) \tag{2.152}$$

So Φ is a ground state wave function at full coupling strength yielding density $\rho_{1/\lambda}(\mathbf{r}) = \lambda^{-3}\rho(\lambda^{-1}\mathbf{r})$. We therefore denote $\Phi = \Psi_{\lambda=1}[\rho_{1/\lambda}]$). So we find that

$$\Psi_\lambda[\rho](\mathbf{r}_1 \ldots \mathbf{r}_N) = \lambda^{\frac{3N}{2}} \Phi(\lambda\mathbf{r}_1 \ldots \lambda\mathbf{r}_N) = \lambda^{\frac{3N}{2}} \Psi_{\lambda=1}[\rho_{1/\lambda}](\lambda\mathbf{r}_1 \ldots \lambda\mathbf{r}_N) \tag{2.153}$$

The diagonal two-particle density matrix at coupling strength λ is then [27]

$$\Gamma_{\lambda, \sigma_1 \sigma_2}([\rho]; \mathbf{r}_1, \mathbf{r}_2) = \sum_{\sigma_3 \ldots \sigma_N} \int |\Psi_\lambda[\rho](\mathbf{r}_1\sigma_1 \ldots \mathbf{r}_N\sigma_N)|^2 d\mathbf{r}_3 \ldots d\mathbf{r}_N =$$

$$= \sum_{\sigma_3 \ldots \sigma_N} \int \lambda^{3N} |\Psi_{\lambda=1}[\rho_{1/\lambda}](\lambda\mathbf{r}_1\sigma_1 \ldots \lambda\mathbf{r}_N\sigma_N)|^2 d\mathbf{r}_3 \ldots d\mathbf{r}_N =$$

$$= \sum_{\sigma_3 \ldots \sigma_N} \lambda^6 \int |\Psi_{\lambda=1}[\rho_{1/\lambda}](\lambda\mathbf{r}_1\sigma_1 \ldots \lambda\mathbf{r}_N\sigma_N)|^2 d\lambda\mathbf{r}_3 \ldots d\lambda\mathbf{r}_N = \lambda^6 \Gamma_{\sigma_1 \sigma_2}([\rho_{1/\lambda}]; \lambda\mathbf{r}_1, \lambda\mathbf{r}_2) \tag{2.154}$$

In particular we find for the pair-correlation function at coupling constant λ that

$$g_{\lambda, \sigma_1 \sigma_2}([\rho]; \mathbf{r}_1, \mathbf{r}_2) = \frac{\Gamma_{\lambda, \sigma_1 \sigma_2}([\rho]; \mathbf{r}_1, \mathbf{r}_2)}{\rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)} = \lambda^6 \frac{\Gamma_{\sigma_1 \sigma_2}([\rho_{1/\lambda}]; \lambda\mathbf{r}_1, \lambda\mathbf{r}_2)}{\rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)} =$$

$$= \frac{\Gamma_{\sigma_1 \sigma_2}([\rho_{1/\lambda}]; \lambda \mathbf{r}_1, \lambda \mathbf{r}_2)}{\rho_{1/\lambda, \sigma_1}(\lambda \mathbf{r}_1) \rho_{1/\lambda, \sigma_2}(\lambda \mathbf{r}_2)} = g_{\sigma_1 \sigma_2}([\rho_{1/\lambda}]; \lambda \mathbf{r}_1, \lambda \mathbf{r}_2) \tag{2.155}$$

So we can calculate the coupling constant integrated pair-correlation function $\bar{g}$ as

$$\bar{g}_{\sigma_1 \sigma_2}([\rho]; \mathbf{r}_1, \mathbf{r}_2) = \int_0^1 g_{\sigma_1 \sigma_2}([\rho_{1/\lambda}]; \lambda \mathbf{r}_1, \lambda \mathbf{r}_2) d\lambda \tag{2.156}$$

This is a useful relation as approximations for $g[\rho]$ are often based on models for full coupling strength. The coupling constant average can then be obtained from the above formula.

We will further derive an expression for v_λ in terms of potentials at full coupling strength. We will first show that if we define (with $\hat{W}$ being Coulombic)

$$F_{LL}^\lambda[\rho] = \min_{\psi \to \rho} \langle \psi | \hat{T} + \lambda \hat{W} | \psi \rangle \tag{2.157}$$

where we search over all normalized anti-symmetric wave functions yielding density ρ that

$$F_{LL}[\rho_\lambda] = \lambda^2 F_{LL}^{1/\lambda}[\rho_{1/\lambda}] \tag{2.158}$$

where $\rho_\lambda(\mathbf{r}) = \lambda^3 \rho(\lambda \mathbf{r})$ and $F_{LL} = F_{LL}^{\lambda=1}$. Equivalently one has

$$F_{LL}^\lambda[\rho] = \lambda^2 F_{LL}[\rho_{1/\lambda}] \tag{2.159}$$

The above equations are readily proved (see Y.Wang in ref [13]). Every wave function ψ yielding density ρ_λ can be written as

$$\psi(\mathbf{r}_1 \ldots \mathbf{r}_N) = \lambda^{\frac{3N}{2}} \Phi(\lambda \mathbf{r}_1 \ldots \lambda \mathbf{r}_N) = \Phi_\lambda(\mathbf{r}_1 \ldots \mathbf{r}_N) \tag{2.160}$$

This equation actually defines Φ and Φ_λ. Then Φ yields density ρ. Therefore we have

$$F_{LL}[\rho_\lambda] = \min_{\psi \to \rho_\lambda} \langle \psi \hat{T} + \hat{W} | \psi \rangle = \min_{\Phi \to \rho} \langle \Phi_\lambda | \hat{T} + \hat{W} | \Phi_\lambda \rangle =$$

$$= \min_{\Phi \to \rho} \langle \Phi | \lambda^2 \hat{T} + \lambda \hat{W} | \Phi \rangle = \lambda^2 \min_{\Phi \to \rho} \langle \Phi | \hat{T} + \frac{1}{\lambda} \hat{W} | \Phi \rangle = \lambda^2 F_{LL}^{1/\lambda}[\rho] \tag{2.161}$$

which proves the statement. The total energy functional for the system at coupling strength λ is given by

$$E_{v_\lambda}^\lambda[\rho] = \int v_\lambda(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} + F_{LL}^\lambda[\rho] \tag{2.162}$$

For $\lambda = 1$ we recover our usual energy functional for the fully interacting system. And for $\lambda = 0$ we have $F_{LL}^0[\rho] = T_s[\rho]$ and we obtain the energy functional of the non-interacting system which is the Kohn-Sham system. The variational equation for the density is given by

$$0 = \frac{\delta}{\delta \rho(\mathbf{r})} \left[E_{v_\lambda}^\lambda[\rho] - \mu_\lambda \int \rho(\mathbf{r}) d\mathbf{r} \right] = \frac{\delta F_{LL}^\lambda}{\delta \rho(\mathbf{r})}[\rho] + v_\lambda(\mathbf{r}) - \mu_\lambda \tag{2.163}$$

In particular for $\lambda = 1$ we find

$$0 = \frac{\delta F_{LL}}{\delta \rho(\mathbf{r})}[\rho] + v(\mathbf{r}) - \mu \tag{2.164}$$

By the definition of v_λ the variational equations are for all λ solved for the same density ρ From the last two equations we then find (to within a constant) that

$$v_\lambda(\mathbf{r}) = \frac{\delta F_{LL}}{\delta\rho(\mathbf{r})}[\rho] - \frac{\delta F_{LL}^\lambda}{\delta\rho(\mathbf{r})}[\rho] + v(\mathbf{r}) \tag{2.165}$$

We now split up F_{LL}^λ as

$$F_{LL}^\lambda[\rho] = T_s[\rho] + \frac{\lambda}{2}\int\frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}^\lambda[\rho] \tag{2.166}$$

where the above equation defines $E_{xc}^\lambda[\rho]$. Then we can write

$$v_\lambda(\mathbf{r}) = v(\mathbf{r}) + (1-\lambda)v_H(\mathbf{r}) + \frac{\delta E_{xc}}{\delta\rho(\mathbf{r})} - \frac{\delta E_{xc}^\lambda}{\delta\rho(\mathbf{r})} \tag{2.167}$$

where we defined the Hartree potential v_H as

$$v_H(\mathbf{r}) = \int\frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' \tag{2.168}$$

From the known scaling property of T_s and the scaling property 2.159 it further follows that

$$E_{xc}^\lambda[\rho] = \lambda^2 E_{xc}[\rho_{1/\lambda}] \tag{2.169}$$

If we define the exchange energy for the system with coupling constant λ as

$$E_x^\lambda[\rho] = -\lambda\frac{1}{4}\int\frac{|\gamma_s([\rho];\mathbf{r}_1,\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2\|}d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.170}$$

where γ_s is the Kohn-Sham one particle density matrix then it follows directly from the above definition that

$$E_x^\lambda[\rho] = \lambda E_x[\rho] \tag{2.171}$$

If we furthermore define the correlation energy functional at coupling constant λ as

$$E_c^\lambda[\rho] = E_{xc}^\lambda[\rho] - E_x^\lambda[\rho] \tag{2.172}$$

we find

$$E_c^\lambda[\rho] = \lambda^2 E_c[\rho_{1/\lambda}] \tag{2.173}$$

So we obtain

$$v_{xc}^\lambda([\rho];\mathbf{r}) = \frac{\delta E_{xc}^\lambda}{\delta\rho(\mathbf{r})}[\rho] = \lambda\frac{\delta E_x}{\delta\rho(\mathbf{r})}[\rho] + \lambda^2\frac{\delta E_c}{\delta\rho(\mathbf{r})}[\rho_{1/\lambda}] = \lambda v_x([\rho];\mathbf{r}) + \lambda^2 v_c([\rho_{1/\lambda}];\mathbf{r}) \tag{2.174}$$

and we finally get the following expression for v_λ [28]

$$v_\lambda(\mathbf{r}) = v(\mathbf{r}) + (1-\lambda)(v_x([\rho];\mathbf{r}) + v_H(\mathbf{r})) + v_c([\rho];\mathbf{r}) - \lambda^2 v_c([\rho_{1/\lambda}];\mathbf{r}) \tag{2.175}$$

For $\lambda = 1$ we find $v_\lambda = v$ and for $\lambda = 0$ we have $v_\lambda = v_s$. This equation has recently been used to derive exact perturbative expressions for the correlation functional in terms of the Kohn-Sham orbitals and eigenvalues, by performing an expansion in the coupling constant parameter λ [29]. How these expressions can be used in a Kohn-Sham calculation is discussed in the section on the Optimized Potential Model.

2.9 Virial relations and the kinetic part of E_{xc}

Exact relations among density functionals can be derived from the general quantum mechanical virial theorem

$$2\langle \hat{T} \rangle = \langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \nabla_i U(\mathbf{r}_1 \ldots \mathbf{r}_N) \rangle \tag{2.176}$$

where the brackets stand for expectation values with the ground state wave function where U represents the sum of all potentials

$$\hat{U} = \hat{V} + \hat{W} = \sum_{i}^{N} v(\mathbf{r}_i) + \frac{1}{2} \sum_{i \neq j}^{N} w(\mathbf{r}_i, \mathbf{r}_j) \tag{2.177}$$

If $\hat{W}$ is Coulombic this yields

$$2\langle \hat{T} \rangle + \langle \hat{W} \rangle = \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla v(\mathbf{r}) d\mathbf{r} \tag{2.178}$$

If we define

$$T[\rho] = \langle \psi[\rho] | \hat{T} | \psi[\rho] \rangle \tag{2.179}$$

$$W[\rho] = \langle \psi[\rho] | \hat{W} | \psi[\rho] \rangle \tag{2.180}$$

Then we have

$$2T[\rho] + W[\rho] = \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla v(\mathbf{r}) d\mathbf{r} \tag{2.181}$$

One should note however that the above relation is not universal. For a given density ρ the above relation is only satisfied for the particular external potential which generates the prescribed density ρ. Universal relations can however be derived on the basis of the virial relation. From the variational relations of the energy functional

$$0 = \frac{\delta}{\delta \rho(\mathbf{r})} \left[E_v[\rho] - \mu \int \rho(\mathbf{r}) d\mathbf{r} \right] = \frac{\delta T}{\delta \rho(\mathbf{r})} + \frac{\delta W}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) - \mu \tag{2.182}$$

we find by acting with $\rho \mathbf{r} \cdot \nabla$ on the above equation that

$$\int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla \frac{\delta T}{\delta \rho(\mathbf{r})} d\mathbf{r} + \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla \frac{\delta W}{\delta \rho(\mathbf{r})} d\mathbf{r} = -\int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla v(\mathbf{r}) d\mathbf{r} \tag{2.183}$$

which by addition to the virial relation yields [22]

$$2T[\rho] + \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla \frac{\delta T}{\delta \rho(\mathbf{r})} d\mathbf{r} d\mathbf{r} = -W[\rho] - \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla \frac{\delta W}{\delta \rho(\mathbf{r})} d\mathbf{r} \tag{2.184}$$

The above equation does not involve the external potential anymore and is universal. For the special case of non-interacting particles we have $\hat{W} = 0$ and $T[\rho] = T_s[\rho]$. In that case the above relation yields

$$2T_s[\rho] + \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla \frac{\delta T_s}{\delta \rho(\mathbf{r})} d\mathbf{r} = 2T_s[\rho] + \int \rho(\mathbf{r}) \mathbf{r} \cdot \nabla v_s([\rho]; \mathbf{r}) d\mathbf{r} = 0 \tag{2.185}$$

which we already proved by scaling. The last two equations in combination with

$$T[\rho] = T_s[\rho] + T_{xc}[\rho] \tag{2.186}$$

$$W[\rho] = E_{xc}[\rho] + T_{xc}[\rho] + \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.187}$$

then yields [30, 22]

$$T_{xc}[\rho] = \int (3\rho(\mathbf{r}) + r \cdot \nabla\rho(\mathbf{r})) v_{xc}([\rho]; \mathbf{r}) d\mathbf{r} - E_{xc}[\rho] \tag{2.188}$$

So we have now established a relation between the kinetic part of the exchange-correlation energy functional and the exchange-correlation energy and potential. If we further split up the exchange-correlation potential $v_{xc} = v_x + v_c$ and use the Levy-Perdew relation for the exchange functional we obtain

$$T_{xc}[\rho] = \int (3\rho(\mathbf{r}) + r \cdot \nabla\rho(\mathbf{r})) v_c([\rho]; \mathbf{r}) d\mathbf{r} - E_c[\rho] \tag{2.189}$$

The above relation provides a way to calculate approximate expressions of the kinetic contribution of the exchange-correlation energy from approximate correlation functionals. The equation can be rewritten as

$$T_{xc}[\rho] = \frac{dE_c}{d\lambda}[\rho_\lambda]|_{\lambda=1} - E_c[\rho] \tag{2.190}$$

as can be checked by differentiation. Further relations for T_{xc} in connection with functional integration can be found in chapter 12 of this thesis.

2.10   Asymptotic properties of the exchange-correlation potential

In this section we will discuss the long-range properties of the exchange-correlation potential in finite systems such as atoms and molecules. For simplicity we will discuss the case of spin unpolarized systems. We can write the exchange-correlation energy in terms of the coupling constant integrated pair-correlation function $\bar{g}$ as

$$E_{xc}[\rho] = \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} (\bar{g}([\rho]; \mathbf{r}_1, \mathbf{r}_2) - 1) d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.191}$$

By functional differentiation the expression for the exchange-correlation potential becomes

$$v_{xc}(\mathbf{r}) = v_{xc,scr}(\mathbf{r}) + v_{xc,scr}^{resp}(\mathbf{r}) \tag{2.192}$$

where the exchange-correlation screening part (where $\bar{g}$ represents the sreening of the interparticle interactions due to exchange-correlation effects) of v_{xc} is defined as

$$v_{xc,scr}(\mathbf{r}_3) = \int \frac{\rho(\mathbf{r}_1)(\bar{g}([\rho]; \mathbf{r}_1, \mathbf{r}_3) - 1)}{|\mathbf{r}_1 - \mathbf{r}_3|} d\mathbf{r}_1 = \int_0^\infty \frac{\rho_{xc}(\mathbf{r}_3, s)}{s} 4\pi s^2 ds \tag{2.193}$$

and the screening response potential as

$$v_{xc,scr}^{resp}(\mathbf{r}_3) = \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \frac{\delta\bar{g}([\rho]; \mathbf{r}_1, \mathbf{r}_2)}{\delta\rho(\mathbf{r}_3)} d\mathbf{r}_1 d\mathbf{r}_2 \tag{2.194}$$

The potential $v_{xc,scr}$ is just the potential of the coupling constant integrated exchange-correlation hole. Due to the fact that this hole integrates to one electron (or the sumrule property of $\bar{g}$) we find that for neutral finite systems

$$v_{xc,scr}(\mathbf{r}) \sim -\frac{1}{r} \qquad (|\mathbf{r}| \to \infty) \tag{2.195}$$

This is therefore a long range Coulombic potential. If we pull an electron to infinity then it will just experience the Coulombic potential of the positive ion it leaves behind. For the calculation of the exchange-correlation energy we only need a good approximation for the screening potential $v_{xc,scr}$ as we have

$$E_{xc}[\rho] = \frac{1}{2} \int \rho(\mathbf{r}) v_{xc,scr}(\mathbf{r}) d\mathbf{r} \tag{2.196}$$

The other part $v_{xc,scr}^{resp}$ of the exchange-correlation potential is short ranged and for atoms it shows a clear atomic shell structure. It is constant within the atomic shells but changes rapidly at the atomic shell boundaries. The function $\delta\bar{g}/\delta\rho$ which describes the sensitivity of the exchange-correlation screening due to density variations shows a clear dependence on the atomic shell structure. An extensive discussion of this function is given in chapter of this thesis. As $v_{xc,scr}$ is long-ranged and $v_{xc,scr}^{resp}$ is short-ranged (decaying faster than Coulombic) we find

$$v_{xc} \sim -\frac{1}{r} \qquad (r \to \infty) \tag{2.197}$$

A more rigorous derivation of the above relation can be found in reference [31]. The asymptotic behaviour of v_{xc} is important for density functional theories for excitation energies which have been developed, as one then needs a correct description of the virtual spectrum, such as the Rydberg-like series in atoms. Also density dependent properties such as polarizabilities are sensitive to a correct behaviour of the exchange-correlation potential. This is an important observation as most density functionals in use, such as the current LDA+GGA exchange-correlation functionals, do not satisfy this asymptotic relation for the exchange-correlation potential [32].

2.11 The Optimized Potential Model

Constructing approximate density functionals by using direct expressions in terms of the density itself can be quite difficult in some cases. This is for instance true for the kinetic energy functional T_s, for which direct Thomas-Fermi like approximations in terms of the density are much more inaccurate. It is moreover more difficult to represent properties such as the atomic shell structure in the exchange-correlation potential correctly be means of direct expressions in the density than expressions in terms of Kohn-Sham orbitals. It is therefore interesting to study energy expressions which depend on the density through the intermediate use of the Kohn-Sham orbitals and energies, like T_s and E_x. Such expressions have recently been proposed by Görling and Levy [29]. For E_x we already had

$$E_x[\rho] = E_x[\{\phi_i\}] \tag{2.198}$$

For E_c we can write

$$E_c[\rho] = E_c[\{\phi_i, \epsilon_i\}] \tag{2.199}$$

For practical applications we need

$$v_x([\rho]; \mathbf{r}) = \frac{\delta E_x[\{\phi_i\}]}{\delta \rho(\mathbf{r})} \tag{2.200}$$

$$v_c([\rho]; \mathbf{r}) = \frac{\delta E_c[\{\phi_i, \epsilon_i\}]}{\delta \rho(\mathbf{r})} \tag{2.201}$$

This can by means of partial differentiation be written as

$$v_x([\rho]; \mathbf{r}) = \sum_i^N \int \frac{\delta E_x}{\delta \phi_i(\mathbf{r}')} \frac{\delta \phi_i(\mathbf{r}')}{\delta \rho(\mathbf{r})} d\mathbf{r}' + c.c. \tag{2.202}$$

$$v_c([\rho]; \mathbf{r}) = \sum_{i=1}^\infty \int \frac{\delta E_c}{\delta \phi_i(\mathbf{r}')} \frac{\delta \phi_i(\mathbf{r}')}{\delta \rho(\mathbf{r})} d\mathbf{r}' + c.c. + \sum_i^\infty \frac{\partial E_c}{\partial \epsilon_i} \frac{\delta \epsilon_i}{\delta \rho(\mathbf{r})} \tag{2.203}$$

From the above expressions we can see that we need to know the functional derivatives $\delta \phi_i / \delta \rho$ and $\delta \epsilon_i / \delta \rho$. We will determine these in the following. The orbitals ϕ_i satisfy the Kohn-Sham equations

$$\left[-\frac{1}{2}\nabla^2 + v_s([\rho]; \mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.204}$$

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{2.205}$$

From response theory it is then not difficult to derive [33]

$$\frac{\delta \phi_i(\mathbf{r})}{\delta v_s(\mathbf{r}')} = -G_i(\mathbf{r}, \mathbf{r}')\phi_i(\mathbf{r}') \tag{2.206}$$

$$\frac{\delta \epsilon_i}{\delta v_s(\mathbf{r}')} = |\phi_i(\mathbf{r}')|^2 \tag{2.207}$$

Where G_i is the Greens function of the Kohn-Sham system defined as

$$G_i(\mathbf{r}, \mathbf{r}') = \sum_{j \neq i} \frac{\phi_j(\mathbf{r})\phi_j^*(\mathbf{r}')}{\epsilon_i - \epsilon_j} \tag{2.208}$$

We can therefore calculate the density response function χ_s of the Kohn-Sham system as

$$\chi_s(\mathbf{r}, \mathbf{r}') = \frac{\delta \rho(\mathbf{r})}{\delta v_s(\mathbf{r}')} = \sum_i^N \frac{\delta}{\delta v_s(\mathbf{r}')} |\phi_i(\mathbf{r})|^2 = -2 \sum_i^N \phi_i(\mathbf{r}) G_i(\mathbf{r}, \mathbf{r}')\phi_i^*(\mathbf{r}') \tag{2.209}$$

The expressions for v_x and v_c can now be written as

$$v_x([\rho]; \mathbf{r}) = \sum_i^N \int \frac{\delta E_x}{\delta \phi_i(\mathbf{r}_1)} G_i(\mathbf{r}_1, \mathbf{r}_2)\phi_i(\mathbf{r}_2)\chi_s^{-1}(\mathbf{r}_2, \mathbf{r}) d\mathbf{r}_1 d\mathbf{r}_2 + c.c. \tag{2.210}$$

and similarly for the correlation potential

$$v_c([\rho]; \mathbf{r}) = \sum_i^\infty \int \frac{\delta E_c}{\delta \phi_i(\mathbf{r}_1)} G_i(\mathbf{r}_1, \mathbf{r}_2)\phi_i(\mathbf{r}_2)\chi_s^{-1}(\mathbf{r}_2, \mathbf{r}) d\mathbf{r}_1 d\mathbf{r}_2 + c.c.+$$

$$+ \sum_i^\infty \frac{\partial E_c}{\partial \epsilon_i} \int |\phi_i(\mathbf{r}_1)|^2 \chi_s^{-1}(\mathbf{r}_1, \mathbf{r}) d\mathbf{r}_1 \tag{2.211}$$

In these equations χ_s^{-1} is the inverse density response function determined by the equation

$$\int \chi_s^{-1}(\mathbf{r}_1, \mathbf{r}_3) \chi_s(\mathbf{r}_3, \mathbf{r}_2) d\mathbf{r}_3 = \delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{2.212}$$

It must be emphasized however that χ_s^{-1} is only determined to within a constant. This is because a change of v_s by a constant does not affect the density. We have

$$\delta\rho(\mathbf{r}_1) = \int \chi_s(\mathbf{r}_1, \mathbf{r}_2) \delta v_s(\mathbf{r}_2) d\mathbf{r}_2 \tag{2.213}$$

If $\delta v_s(\mathbf{r}) = C$ is a constant then $\delta\rho(\mathbf{r}) = 0$ and we find

$$\int \chi_s(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = 0 \tag{2.214}$$

The constant function is therefore an eigenfunction of χ_s with eigenvalue 0. So if a certain χ_s^{-1} satisfies equation 2.212 then $\chi_s^{-1} + C$ will satisfy this equation as well. The constant can however be fixed by choosing a particular gauge for the potential v_s , for instance $v_s(\mathbf{r}) \to 0$ $(r \to \infty)$.

As χ_s is explicitly known in terms of ϕ_i and ϵ_i the above equations for v_x and v_c together with the Kohn-Sham equations constitute a self-consistent set of equations which can be solved once approximate expressions for $E_c[\{\phi_i, \epsilon_i\}]$ are known.

Note that the problem is equivalent to minimizing a total energy expression $E_v[\{\phi_i, \epsilon_i\}]$ under the constraint that the orbitals ϕ_i obey an independent particle equation with a local potential, i.e. one has to find the local potential that solves the Euler-Lagrange equation

$$\frac{\delta E_v}{\delta v_s(\mathbf{r})}[\{\phi_i, \epsilon_i\}] = 0 \tag{2.215}$$

For this reason these equation are called the Optimized Potential Model (OPM) equations [34, 33, 35, 36]. The simplest case is the exchange-only OPM, for which we can put $E_c = 0$, yielding the x-only OPM equations

$$\left[-\frac{1}{2}\nabla^2 + v_s([\rho]; \mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.216}$$

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{2.217}$$

$$v_s([\rho]; \mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_x([\rho]; \mathbf{r}) \tag{2.218}$$

$$\int \chi_s(\mathbf{r}, \mathbf{r}') v_x(\mathbf{r}') d\mathbf{r}' = \frac{\delta E_x}{\delta v_s(\mathbf{r})} \tag{2.219}$$

For solving these equations one needs to solve an integral equation (to be more precise a Fredholm integral equation of the first kind) for the exchange potential v_x. This requires inversion of the integral kernel χ_s (after fixing a gauge) [37, 36]. One can however avoid this problem if one makes explicit orbital dependent approximations for the inverse function χ_s^{-1}. An explicit expression can be derived on the basis of an approximation for the Kohn-Sham Greens function G_i proposed by

Sharp and Horton [34] and Krieger et al. [35]. One finds (see chapter 8 of this thesis) in the natural gauge where $v_s \to 0$ $(r \to \infty)$ for χ_s^- (for the spin unpolarized non-degenerate case)

$$\chi_s^{-1}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho(\mathbf{r}_1)} - \sum_{ik}^{N-1} \eta_{ik} \frac{|\phi_i(\mathbf{r}_1)|^2 |\phi_k(\mathbf{r}_2)|^2}{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)} \tag{2.220}$$

where the coefficients η_{ik} are given by

$$\eta_{ik} = (\bar{I} - \bar{N})_{ik}^{-1} \tag{2.221}$$

where $\bar{I}$ is the unity matrix and N is the matrix

$$N_{ik} = \int \frac{|\phi_i(\mathbf{r})|^2 |\phi_k(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r} \tag{2.222}$$

where all the matrices are of dimension $N - 1$. If one uses this approximation for χ_s^{-1} one obtains the following approximation for the OPM exchange potential [35]

$$v_x^{KLI}(\mathbf{r}) = v_S(\mathbf{r}) + \sum_i^{N-1} w_i \frac{|\phi_i(\mathbf{r})|^2}{\rho(\mathbf{r})} \tag{2.223}$$

$$w_i = \langle \phi_i | v_x - v_i | \phi_i \rangle \tag{2.224}$$

where v_i is equal to the orbital dependent Hartree-Fock potential (only its self-consistent orbitals are different). The potential v_S is the Slater potential. These equations have been derived first by Krieger, Li and Iafrate in a different way. This approximation turns out to be very accurate [38, 39, 40, 41]

From these approximate equations for v_x we can easily see the structure of the exchange potential. It consists of a part v_S which has a Coulombic long range $-1/r$ behaviour and a part we will call the step potential. For atomic systems this step potential is constant within the atomic shells and changing rapidly at the atomic shell boundaries. It has therefore a step-like structure with the heigths of the steps approximately equal to the constants w_i. A more extensive discussion of all these features can be found in chapter 8 of this thesis.

2.12   Approximate density functionals: LDA and GGA

The simplest and most widely used density functional approximation for the exchange-correlation energy is the Local Density Approximation (LDA). In this approximation one uses the exchange-correlation energy density of the homogeneous electron gas $e_{xc}^{hom}(\rho_0)$ dependent on the homogeneous density ρ_0 and replaces this for the inhomogeneous system with density $\rho(\mathbf{r})$ by

$$e_{xc}^{LDA}(\rho(\mathbf{r})) = e_{xc}^{hom}(\rho_0)|_{\rho_0 = \rho(\mathbf{r})} \tag{2.225}$$

The exchange-correlation energy functional then becomes

$$E_{xc}^{LDA}[\rho] = \int e_{xc}^{LDA}(\rho(\mathbf{r})) d\mathbf{r} \tag{2.226}$$

At first sight this approximation seems to work only for systems with slowly varying densities, such as the weakly perturbed electron gas. However the approximation works surprisingly well even for systems which have very inhomogeneous electron densities such as atoms and molecules. This calls

for an explanation.

A more detailed look on LDA can be obtained from considerations of the xc-hole and the pair-correlation function. We have

$$\rho_{xc}^{LDA}(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_1) \left[\bar{g}^{hom}([\rho]; |\mathbf{r}_1 - \mathbf{r}_2|) - 1 \right] \tag{2.227}$$

where $\bar{g}^{hom}[\rho]$ is the coupling constant integrated pair-correlation function of the homogeneous electron gas. By now accurate expressions for $\bar{g}^{hom}[\rho]$ are known. The simplest part is the exchange part of the pair-correlation function which is unaffected by the coupling constant integration. It is given (in its spin unpolarized form) by

$$\bar{g}_x^{hom}([\rho]; \mathbf{r}_1, \mathbf{r}_2) = 1 - \frac{9}{2} \left[\frac{\sin(k_F(\mathbf{r}_1)|\mathbf{r}_1 - \mathbf{r}_2|) - k_F(\mathbf{r}_1)|\mathbf{r}_1 - \mathbf{r}_2| \cos(k_F(\mathbf{r}_1)|\mathbf{r}_1 - \mathbf{r}_2|)}{(k_F(\mathbf{r}_1)|\mathbf{r}_1 - \mathbf{r}_2|)^3} \right]^2 \tag{2.228}$$

where $k_F(\mathbf{r})$ is the local Fermi wave vector defined as

$$k_F(\mathbf{r}) = (\frac{3}{\pi})^{\frac{1}{3}} \rho(\mathbf{r})^{\frac{1}{3}} \tag{2.229}$$

It satisfies the scaling property

$$\bar{g}_x^{hom}([\rho]; \mathbf{r}_1, \mathbf{r}_2) = \bar{g}_x^{hom}([\rho_{1/\lambda}]; \lambda\mathbf{r}_1, \lambda\mathbf{r}_2) \tag{2.230}$$

where $\rho_{1/\lambda}(\mathbf{r}) = \lambda^{-3}\rho(\lambda^{-1}\mathbf{r})$. The corresponding exchange functional as can be calculated from the above pair-correlation function is then

$$E_x^{LDA}[\rho] = \frac{3}{4}(\frac{3}{\pi})^{\frac{1}{3}} \int \rho(\mathbf{r})^{\frac{4}{3}} d\mathbf{r} \tag{2.231}$$

which satisfies the correct exchange scaling. Corresponding expressions exist for the correlation part of $\bar{g}^{hom}[\rho]$ and for $E_c^{LDA}[\rho]$. The first important thing to note is that the LDA xc-hole is spherical around the reference electron

$$\rho_{xc}^{LDA}(\mathbf{r}_1, \mathbf{r}_2) = \rho_{xc}(\mathbf{r}_1, s) \tag{2.232}$$

where $s = |\mathbf{r}_1 - \mathbf{r}_2|$ and secondly it satisfies the sum rule

$$\int \rho_{xc}^{LDA}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = 4\pi \int_0^\infty \rho_{xc}^{LDA}(\mathbf{r}_1, s) s^2 ds = -1 \tag{2.233}$$

The above relations provide some insight in the unexpected success of LDA in strongly inhomogeneous systems. The coupling constant averaged xc-hole in real inhomogeneous systems is not spherically symmetric. Therefore the LDA xc-hole cannot adequately approximate this xc-hole. However the spherically averaged hole is reasonably well reproduced within the local density approximation. Consequently the exchange-correlation energy which is determined by the spherically averaged xc-hole is reasonably well produced as well. The satisfaction of the sum rule 2.233 further implies that if the LDA xc-hole $\rho_{xc}^{LDA}(\mathbf{r}, s)$ has positive errors for some values of s, it must have negative errors for other values of s leading to a systematic cancelation of errors.

An obvious way to go beyond the local density approximation is to extend the exchange-correlation functional with terms containing gradients of the electron density. The gradients then measure changes of the electron density and can therefore be expected to improve the local density approximation. Most of the gradient functionals developed are based on the response theory of the weakly

varying electron gas [42]. The approximations obtained in this way are called Gradient Expansion Approximations (GEA). This then leads to the expressions

$$E_x^{GEA}[\rho] = E_x^{LDA}[\rho] + \beta \int \frac{(\nabla \rho)^2}{\rho^{\frac{4}{3}}} d\mathbf{r} + \dots \tag{2.234}$$

$$E_c^{GEA}[\rho] = E_c^{LDA}[\rho] + \int C(\rho) \frac{(\nabla \rho)^2}{\rho^{\frac{4}{3}}} d\mathbf{r} + \dots \tag{2.235}$$

where β is a constant and $C(\rho)$ is a function determined by response theory. However in practice, the success of LDA nothwithstanding, the GEA provides no systematic improvement over LDA as realistic densities in atoms and molecules do not vary slowly over space. An analysis of the gradient expansion of the xc-hole shows that although the short range part (near the reference electron) is improved by the gradient expansion, the long range part is considerably worsened [43]. If one however by hand corrects the long range part of the gradient corrected xc-hole and enforces the sum rule property one can obtain xc-energy functionals which give a considerable improvement over LDA for energetics. The corresponding approximations are known as Generalized Gradient Approximations or GGA's. The most widely used GGA's are the Becke GGA [44] for the exchange energy and the exchange- and correlation GGA's by Perdew [45] and Perdew and Wang [46, 47, 43, 48]. We will give a short account of the ideas behind the derivation of these GGA's.
The Becke GGA correction to the LDA is of the form

$$E_x^{GGA}[\rho] = \int \rho(\mathbf{r})^{\frac{4}{3}} f(x(\mathbf{r})) d\mathbf{r} \tag{2.236}$$

where $x(\mathbf{r}) = |\nabla \rho(\mathbf{r})|/\rho(\mathbf{r})^{\frac{4}{3}}$ and f is a function to be determined. The variable x is a dimensionless quantity introduced to satisfy the scaling property of the exchange functional. In order to obtain the GEA for weakly varying densities we must have the following small x behaviour of f

$$f(x) \sim -\beta x^2 \qquad (x \downarrow 0) \tag{2.237}$$

We can write the exchange energy in terms of the potential of the exchange hole v_x^h as

$$E_x[\rho] = \frac{1}{2} \int \rho(\mathbf{r}) v_x^h(\mathbf{r}) d\mathbf{r} \tag{2.238}$$

In the above ansatz we have

$$v_x^h(\mathbf{r}) = v_x^{h,LDA}(\mathbf{r}) + 2\rho(\mathbf{r})^{\frac{1}{3}} f(x(\mathbf{r})) \tag{2.239}$$

To satisfy the correct asymptotics for v_x^h for finite systems ($v_x^h \sim -1/r$ $(r \to \infty)$) we must have [49, 32]

$$f(x) \sim -\frac{1}{6} \frac{x}{\ln x} \quad (x \to \infty) \sim -\frac{\rho(\mathbf{r})}{2} \frac{1}{r} \quad (r \to \infty) \tag{2.240}$$

which follows directly by inserting exponentially decaying densities. One of the simplest interpolations between the large and the small x behaviour of $f(x)$ is then found by taking

$$f(x) = -\frac{\beta x^2}{1 + 6\beta x \sinh^{-1} x} \tag{2.241}$$

The coefficient β was fitted by Becke to obtain the correct exchange energy of the noble gas atoms. A somewhat different approach to obtain GGA's has been used by Langreth and Mehl [50],

Perdew [45], and Perdew and Wang [43, 48]. One can write the spherical average of the coupling constant integrated xc-hole in momentum space as [50, 43]

$$\bar{\rho}_{xc}(\mathbf{r}, k) = \int \bar{\rho}_{xc}(\mathbf{r}, s)e^{i\mathbf{k}\cdot\mathbf{s}}d\mathbf{s} = \int_0^\infty 4\pi s^2 \frac{\sin(ks)}{ks}\bar{\rho}_{xc}(\mathbf{r}, s)ds \tag{2.242}$$

Its Fourier inverse is given by

$$\bar{\rho}_{xc}(\mathbf{r}, s) = \frac{1}{(2\pi)^3}\int_0^\infty 4\pi k^2 \frac{\sin(ks)}{ks}\bar{\rho}_{xc}(\mathbf{r}, k)dk \tag{2.243}$$

Langreth and Perdew [51] and Langreth and Mehl [50] have carried out an GEA wavevector analysis for the spherically averaged correlation hole within the random phase approximation (RPA) yielding

$$\bar{\rho}_c^{GEA}(\mathbf{r}, k) = \bar{\rho}_c^{LDA}(\mathbf{r}, k) + \frac{2\pi\sqrt{3}}{k_{FT}(\mathbf{r})}C(\infty)\frac{(\nabla\rho(\mathbf{r}))^2}{\rho(\mathbf{r})^{\frac{4}{3}}}\exp\left(-\frac{2\sqrt{3}}{k_{FT}(\mathbf{r})}\frac{C(\infty)}{C(\rho)}k\right) \tag{2.244}$$

where $k_{FT}(\mathbf{r}) = 2\sqrt{k_F(\mathbf{r})/\pi}$ is the Fermi-Thomas wavevector or inverse screening length and C is a local function of the electron density. From equation 2.242 we find

$$\lim_{k\to 0}\bar{\rho}_{xc}(\mathbf{r}, k) = \int_0^\infty 4\pi s^2 \bar{\rho}_{xc}(\mathbf{r}, s)ds = -1 \tag{2.245}$$

In particular we have for the correlation hole in momentum space

$$\bar{\rho}_c(\mathbf{r}, k = 0) = 0 \tag{2.246}$$

One can see from equation 2.244 that this property is violated by the GEA correlation hole (it is satisfied by the LDA correlation hole). To repair this feature Langreth and Mehl and Perdew propose a wavevector space cut-off. One replaces the gradient contribution to $\bar{\rho}_{xc}(\mathbf{r}, k)$ by zero for $k < k_c$ where $k_c = \tilde{f}|\nabla\rho(\mathbf{r})|/\rho(\mathbf{r})$ is a cut-off proportional to the inhomogeneity wavevector, i.e. one removes the long wave length part of the Fourier analysis of the xc-hole. This then leads to the following approximation for the correlation energy functional [45]

$$E_c^{GGA}[\rho] = E_c^{LDA}[\rho] - \int e^{-\Phi}C(\rho)\frac{(\nabla\rho(\mathbf{r}))^2}{\rho(\mathbf{r})^{\frac{4}{3}}}d\mathbf{r} \tag{2.247}$$

where

$$\Phi = 1.745\tilde{f}\frac{C(\infty)}{C(\rho)}\frac{|\nabla\rho|}{\rho^{\frac{7}{6}}} \tag{2.248}$$

Other GGA approximations for exchange and correlation have been obtained by Perdew and Wang [43, 48, 47] using real space cut-offs of the xc-hole. We will describe the procedure for the exchange energy functional. The correlation functional can be treated in a similar manner. The GEA exchange hole in the second order gradient expansion has the following structure

$$\bar{\rho}_x^{GEA}(\mathbf{r}, s) = k_F(\mathbf{r})^3\left[A_x(k_F(\mathbf{r})s) + x(\mathbf{r})^2 B_x(k_F(\mathbf{r})s)\right] \tag{2.249}$$

where $k_F(\mathbf{r}) = (3\pi^2\rho(\mathbf{r}))^{\frac{1}{3}}$ is the local Fermi wavevector and $x(\mathbf{r}) = |\nabla\rho|/\rho^{\frac{4}{3}}$ is a dimensionless inhomogeneity parameter. In the GGA one takes

$$\bar{\rho}_x^{GGA}(\mathbf{r}, s) = \bar{\rho}_x^{GEA}(\mathbf{r}, s)\theta(U(x) - k_F(\mathbf{r})s) \tag{2.250}$$

The Heaviside function θ provides a long range cutoff of the exchange hole. The cut-off radius is chosen such that the sum rule is satisfied

$$-1 = \int_0^\infty 4\pi s^2 \bar{\rho}_{xc}(\mathbf{r}, s) ds = \int_0^{U(x)} 4\pi u^2 \left[A_x(u) + x(\mathbf{r})^2 B_x(u) \right] du \tag{2.251}$$

One obtains

$$E_x^{GGA}[\rho] = \int \rho(\mathbf{r})^{\frac{4}{3}} F(x) d\mathbf{r} \tag{2.252}$$

with

$$F(x) = \int_0^{U(x)} 4\pi u \left[A_x(u) + x(\mathbf{r})^2 B_x(u) \right] du \tag{2.253}$$

The functions $U(x)$ and $F(x)$ can be evaluated numerically and fitted to an analytical expression for practical applications. A pleasing feature from a theoretical point of view is that the above approach does not contain any adjustable parameters which need to be fitted to known exchange energies.

Completely analogous to the exchange case one can also carry out a real-space cut-off procedure for the correlation hole leading to parameter free GGA's for the correlation energy. These GGA's turn out to yield accurate atomic correlation energies [52, 53, 54, 55]

2.13 An integral equation for the exchange-correlation potential

In this section we will establish some connections between density functional theory and Greens function theory. We will in particular derive an integral equation for the exchange-correlation potential in terms of the irreducible self-energy. This equation was first derived by Sham and Schlüter [56] and Sham [57].

We first discuss the key concepts of Greens function theory. The one-particle Greens function is defined as the following expectation value

$$
\begin{aligned}
iG(xt, x't') &= \theta(t - t')\langle \Psi(t)_S | \hat{\psi}(x) e^{-i\hat{H}(t - t')} \hat{\psi}^+(x') | \Psi(t')_S \rangle \\
&\quad - \theta(t' - t)\langle \Psi(t')_S | \hat{\psi}^+(x) e^{-i\hat{H}(t' - t)} \hat{\psi}(x') | \Psi(t)_S \rangle
\end{aligned}
\tag{2.254}
$$

where θ is the Heaviside function and

$$|\Psi(t)\rangle_S = e^{-iEt} |\Psi \rangle \tag{2.255}$$

is the ground state in the Schrödinger picture where E is the groundstate energy. The Greens function as defined above has a nice physical interpretation. For $t > t'$ we add by the action of $\hat{\psi}^+(x')$ a particle to the ground state $|\Psi(t')_S\rangle$ at spin-space point x'. The in this way created $N + 1$-particle state will then propagate under the influence of the Hamiltonoperator $\hat{H}$ from t' to t where we take overlap with the $N + 1$-particle state $\hat{\psi}^+(x)|\Psi(t)\rangle_S$. The Greens function for $t > t'$ therefore describes the transition amplitude that a particle which is added to the ground state of many-particle system will move in the time $t - t'$ from point x' to point x. In a similar manner we have for $t' > t$ a propagation of a $N - 1$-particle state. So for $t > t'$ we are dealing with particle propagation and for $t' > t$ with hole propagation.

The Greens function may be written in more compact notation by going to the Heisenberg picture

$$
\begin{aligned}
iG(xt, x't') &= \theta(t - t')\langle \Psi | \hat{\psi}(xt)_H \hat{\psi}^+(x't')_H | \Psi \rangle - \theta(t' - t)\langle \Psi | \hat{\psi}^+(x't')_H \hat{\psi}(xt)_H | \Psi \rangle \tag{2.256} \\
&= \langle \Psi | T[\hat{\psi}(xt)_H \hat{\psi}^+(x't')_H] | \Psi \rangle
\end{aligned}
$$

where we defined

$$
\begin{aligned}
\hat{\psi}(xt)_H &= e^{i\hat{H}t}\hat{\psi}(x)e^{-i\hat{H}t} \\
\hat{\psi}^+(xt)_H &= e^{i\hat{H}t}\hat{\psi}^+(x)e^{-i\hat{H}t}
\end{aligned}
\tag{2.257}
$$

and the time-ordered product of two fermion operators as

$$
T[\hat{A}(t)\hat{B}(t')] = \theta(t-t')\hat{A}(t)\hat{B}(t') - \theta(t'-t)\hat{B}(t')\hat{A}(t)
\tag{2.258}
$$

The time-ordered expression is especially suitable for a perturbation expansion where the time-ordering appears in a natural way.

As G describes the time propagation of a $N+1$-particle state and a $N-1$-particle state it contains information about the excitation spectrum of the system. This is made explicit by writing

$$
\begin{aligned}
iG(xt, x't') &= \theta(t-t')\sum_n e^{-i(E_n^{N+1}-E)(t-t')}\langle\Psi|\hat{\psi}(x)|\Psi_n^{N+1}\rangle\langle\Psi_n^{N+1}|\hat{\psi}^+(x')|\Psi\rangle \\
&\quad - \theta(t'-t)\sum_n e^{-i(E-E_n^{N-1})(t'-t)}\langle\Psi|\hat{\psi}^+(x')|\Psi_n^{N-1}\rangle\langle\Psi_n^{N-1}|\hat{\psi}(x)|\Psi\rangle
\end{aligned}
\tag{2.259}
$$

where E_n^{N+1} and E_n^{N-1} are the energies of the eigenstates $|\Psi_n^{N+1}\rangle$ and $|\Psi_n^{N-1}\rangle$ of the $N+1$- and $N-1$-particle systems. From this equation we can see that G only depends on x and x' and the time difference $t-t'$. We can therefore define the Fourier transform of G by

$$
G(x, x';\omega) = \int_{-\infty}^{+\infty} G(x, x';t-t')e^{-i\omega(t-t')}\frac{d\omega}{2\pi}
\tag{2.260}
$$

Using the following expression for the Heaviside step function

$$
\theta(\tau) = \lim_{\eta\downarrow 0}\frac{-1}{2\pi i}\int_{-\infty}^{+\infty}\frac{e^{-i\omega\tau}}{\omega+i\eta}d\omega
\tag{2.261}
$$

we find an expression for G which clearly reveals its analytical structure

$$
G(x, x';\omega) = \sum_n \frac{f_n(x_1)f_n^*(x_1')}{\omega-\alpha_n+i\eta} + \frac{g_n(x_1)g_n^*(x_1')}{\omega-\epsilon_n+i\eta}
\tag{2.262}
$$

where f_n and g_n are defined as

$$
\begin{aligned}
f_n(x) &= \langle\Psi|\hat{\psi}(x)|\Psi_n^{N+1}\rangle = \sqrt{N+1}\int\Psi^*(x_1\ldots x_N)\Psi_n^{N+1}(x_1\ldots x_N x)dx_1\ldots dx_N \quad (2.263)\\
g_n(x) &= \langle\Psi_n^{N-1}|\hat{\psi}(x)|\Psi\rangle = \sqrt{N}\int\Psi_n^{N-1*}(x_1\ldots x_{N-1})\Psi(x_1\ldots x_{N-1}x)dx_1\ldots dx_{N-1} \quad (2.264)
\end{aligned}
$$

and where $\alpha_n = E_n^{N+1}-E$ and $\epsilon_n = E-E_n^{N+1}$ correspond to the electron affinities $-\alpha_n$ and the ionisation energies $-\epsilon_n$. In the above derivation we assumed all the levels E_n^{N+1} and E_n^{N-1} of the $N-1$ and $N+1$ systems to be discrete. This is of course not true for a real electronic system such as an atom or a molecule where a continuous spectrum always exists. For an infinite system such a solid there is even no discrete spectrum at all. For these cases one has to replace the summation by an integral over the energies. This changes the analytical structure of G of equation 2.262 ,i.e. in the addition to the simple poles corresponding to the discrete levels we obtain branch cuts corresponding to the continuum states.

If we want to find approximations for G we need an equation of motion. Such an equation can be derived using the Heisenberg equation of motion

$$
i\partial_t\hat{\psi}(xt)_H = [\hat{\psi}(xt)_H, \hat{H}]
\tag{2.265}
$$

which using the commutation relations 2.13 and the form of the Hamiltonian 2.4 gives

$$i\partial_t \hat{\psi}(xt)_H = \left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int w(\mathbf{r}, \mathbf{r}')\hat{\psi}^+(x't)_H\hat{\psi}(x't)_H \, dx' \right] \hat{\psi}(xt)_H \tag{2.266}$$

With the above equation and the definition of G one finds

$$(i\partial_{t_1} + \frac{1}{2}\nabla_1^2 - v(\mathbf{r}_1))G(x_1t_1, x_1't_1') =$$

$$\delta(x_1 - x_1')\delta(t_1 - t_1') - i\int w(\mathbf{r}_1, \mathbf{r}_2)G_2(x_1t_1, x_2t_1, x_1't_1', x_2t_1^+)d\mathbf{r}_2 \tag{2.267}$$

where t_1^+ means $t_1 + \delta$ where $\delta \downarrow 0$. Here we defined the two-particle Green function as

$$G_2(x_1t_1, x_2t_2, x_1't_1', x_2't_2') = (-i)^2 \langle \Psi | T[\hat{\psi}(x_1t_1)_H\hat{\psi}(x_2t_2)_H\hat{\psi}^+(x_2't_2')_H\hat{\psi}^+(x_1't_1')_H] | \Psi \rangle \tag{2.268}$$

and where the time-ordered product of a arbitrary number of fermion operators has been defined as

$$T[\hat{A}_1(t_1)\hat{A}_2(t_2)\ldots\hat{A}_n(t_n)] = \text{sign}(P)\hat{A}_{P(1)}(t_{P(1)})\hat{A}_{P(2)}(t_{P(2)})\ldots\hat{A}_{P(n)}(t_{P(n)}) \tag{2.269}$$

where the permutaion P is chosen such that $t_{P(1)} > t_{P(2)} > \ldots > t_{P(n)}$. The two-partcle Greens function G_2 has a similar physical interpretation as G as a transition amplitude but now with two particles added or removed from the system. Another important quantity is the self-energy Σ defined by the equation

$$(i\partial_{t_1} + \frac{1}{2}\nabla_1^2 - v(\mathbf{r}_1))G(x_1t_1, x_1't_1') =$$

$$\delta(x_1 - x_1')\delta(t_1 - t_1') - i\int w(\mathbf{r}_1, \mathbf{r}_2)\Sigma(x_1t_1, x_2t_2)G(x_2t_2, x_1't_1')dx_2dt_2d\mathbf{r}_2 \tag{2.270}$$

The Fourier transform of the equation yields

$$[\omega - h(x_1)]G(x_1, x_1'; \omega) - \int \Sigma(x_1, x_2; \omega)G(x_2, x_1'; \omega)dx_2 = \delta(x_1 - x_1') \tag{2.271}$$

If we define the Greens function G_0 for a system of non-interacting particles by the equation

$$[\omega - h(x_1)]G_0(x_1, x_1'; \omega) = \delta(x_1 - x_1') \tag{2.272}$$

We obtain the equation

$$G(x_1, x_1'; \omega) = G_0(x_1, x_1'; \omega) + \int G_0(x_1, x_2; \omega)\Sigma(x_2, x_3; \omega)G(x_3, x_1'; \omega)dx_2dx_3 \tag{2.273}$$

This equation is known as Dysons equation. We will write this more symbolically as

$$G = G_0 + G_0\Sigma G \tag{2.274}$$

In the same operator notation we have

$$[\omega - h - \Sigma]G = 1 \tag{2.275}$$

We now split up the self-energy in Hartree part and a remainder Σ_{xc} as follows

$$\Sigma(x_1, x_2; \omega) = \delta(x_1 - x_2)v_H(x_1) + \Sigma_{xc}(x_1, x_2; \omega) \tag{2.276}$$

where v_H is the repulsive potential of the electronic cloud. Let us now rewrite the Dyson equation as

$$\left[\omega - h_s - \tilde{\Sigma}\right] G = 1 \tag{2.277}$$

where

$$h_s(\mathbf{r}) = -\frac{1}{2}\nabla^2 + v(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \tag{2.278}$$

is the Kohn-Sham Hamiltonian and we defined $\tilde{\Sigma}$ as

$$\tilde{\Sigma}(x_1, x_2; \omega) = \Sigma_{xc}(x_1, x_2; \omega) - \delta(x_1 - x_2)v_{xc}(x_1) \tag{2.279}$$

One obtains an alternative Dyson equation in terms of the Kohn-Sham Greens function G_s

$$G = G_s + G_s\tilde{\Sigma}G \tag{2.280}$$

where G_s satisfies

$$\left[\omega - h_s\right] G_s = 1 \tag{2.281}$$

By the definition of the Kohn-Sham system the diagonal of G_s yields the ground state density of the system

$$\rho(\mathbf{r}) = -iG_s(x_1 t, x_1 t+) = -i \int \frac{d\omega}{2\pi} G_s(x_1, x_1; \omega) \tag{2.282}$$

being equal to the diagonal Greens function of the fully interacting system, i.e.

$$\rho(\mathbf{r}) = -i \int \frac{d\omega}{2\pi} G(x_1, x_1; \omega) \tag{2.283}$$

From the Dyson equation it then follows that

$$\int dx_1 \int dx_2 \int \frac{d\omega}{2\pi} G(x, x_1; \omega)\tilde{\Sigma}(x_1, x_2; \omega)G(x_2, x; \omega) = 0 \tag{2.284}$$

which by definition of $\tilde{\Sigma}$ yields the following integral equation for v_{xc}.

$$\int K(\mathbf{r}_1, \mathbf{r}_2)v_{xc}(\mathbf{r}_2)d\mathbf{r}_2 = Q(\mathbf{r}_1) \tag{2.285}$$

where

$$K(\mathbf{r}_1, \mathbf{r}_2) = \int \frac{d\omega}{2\pi} G_s(\mathbf{r}_1, \mathbf{r}_2; \omega)G(\mathbf{r}_2, \mathbf{r}_1; \omega) \tag{2.286}$$

$$Q(\mathbf{r}) = \int \frac{d\omega}{2\pi} \int d\mathbf{r}_1 \int d\mathbf{r}_2 G_s(\mathbf{r}, \mathbf{r}_1; \omega)\Sigma_{xc}(\mathbf{r}_1, \mathbf{r}_2; \omega)G(\mathbf{r}_2, \mathbf{r}; \omega) \tag{2.287}$$

Which is an integral equation for the exchange-correlation potential. The equation can be solved self-consistently once an explicit expansion of G in terms of G_s is given. One of the simplest approximations is

$$G = G_s \tag{2.288}$$

$$\Sigma_{xc}(\mathbf{r}_1, \mathbf{r}_2) = -i \int \frac{d\omega}{2\pi} \frac{G_s(\mathbf{r}_1, \mathbf{r}_2; \omega)}{|\mathbf{r}_1 - \mathbf{r}_2|} = -\frac{\gamma_s(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{2.289}$$

yielding

$$K(\mathbf{r}_1, \mathbf{r}_2) = \int \frac{d\omega}{2\pi} G_s(\mathbf{r}_1, \mathbf{r}_2; \omega) G_s(\mathbf{r}_2, \mathbf{r}_1; \omega) = \chi_s(\mathbf{r}_1, \mathbf{r}_2) \tag{2.290}$$

$$Q(\mathbf{r}) = \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\gamma_s(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \int \frac{d\omega}{2\pi} G_s(\mathbf{r}, \mathbf{r}_1; \omega) G_s(\mathbf{r}_2, \mathbf{r}; \omega) \tag{2.291}$$

These equations together with

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r}) \right] \phi_1(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.292}$$

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{2.293}$$

yield the x-only OPM equations.

The integral equation for v_{xc} can also be derived in an alternative way. One can express the exchange-correlation energy functional as

$$E_{xc}[\rho] = i\text{tr}(\ln(1 - \tilde{\Sigma}G_s) + \tilde{\Sigma}G) - i\Phi_{xc}[\rho] \tag{2.294}$$

where in diagrammatic perturbation theory Φ_{xc} is expressed as a sum of so-called skeleton diagrams in terms of the Greens function G excluding two first order graphs. For a derivation we refer to Sham and Schlüter [56] and Sham [57]. Functional differentiation of the above expression then yields the integral equation for v_{xc} (see for example the references [58, 59]) The integral equation for v_{xc} has been used in the so-called GW-approximation [2, 60, 61, 62] to estimate band gap corrections to the LDA [63] for semiconductors and insulators.

Chapter 3

A review of functional calculus

3.1 Introduction

In this review we present some aspects of nonlinear functional analysis [64, 65] .We thereby concentrate on differential and integral calculus on function spaces. Of most practical importance is the finding of extrema of functionals. This is the central problem in the calculus of variations which is for instance widely used in the Lagrangian formalism in classical mechanics. In order to do calculus on more general spaces than the real numbers $\mathcal{R}$ and the complex numbers $\mathcal{C}$ those general spaces have to satisfy some requirements. Roughly said, they must look like the space of real numbers. The important features to be carried over to these spaces are its vectorspace structure (adding and multiplying), its topological structure (there is a distance between two points) and the completeness property (the real numbers form a continuum). General spaces with these properties are called Banach spaces. Special cases of Banach spaces are of course the real numbers $\mathcal{R}$ and the complex numbers $\mathcal{C}$ themselves.

3.2 Banach spaces

3.2.1 Definition

Many for theorems in ordinary calculus depend on three basic properties of the real numbers. Any extension of ordinary calculus to more general spaces should exhibit these basic properties. These are summed up in the following cryptic statements:

a] You can add and multiply points
b] There is a distance between points
c] There is a continuum of points

We will now discuss each statement separately. This then leads to the definition of Banach space.
First of all a Banach space B is a vectorspace which means that if $x, y \in B$ and $\lambda \in \mathcal{R}$ or $\mathcal{C}$ then:

$$x, y \in B \Rightarrow x + y \in B$$
$$x \in B \text{ and } \lambda \in \mathcal{R} \text{ or } \mathcal{C} \Rightarrow \lambda x \in B$$

We thus just add and multiply as for the real numbers.

Secondly a Banach space is a normed space which means that we have on B a function $\| \| : B \to \mathcal{R}^{\geq 0}$ from B to the positive real numbers including zero with the following properties:

$$\|\lambda x\| = |\lambda| \|x\|$$
$$\|x + y\| \leq \|x\| + \|y\|$$
$$\|x\| = 0 \Leftrightarrow x = 0$$

These properties are a copy of the properties of the length of an ordinary vector in three-dimensional space. The introduction of a norm now allows us to define the distance between two points by the following formula:

$$d(x, y) = \|x - y\|$$

We have for instance:

$$d(x, y) = 0 \Leftrightarrow x = y$$

So now we can talk about the nearness of points. This is a very important ingredient in calculus because now we can talk about the convergence of limits or about continuous functions which map nearby points to nearby points.

The third important property of the real numbers is its completeness. This is best illustrated with an example. Consider the space of rational numbers $\mathcal{Q}$ embedded in the space of real numbers $\mathcal{R}$. Both $\mathcal{Q}$ and $\mathcal{R}$ are normed vectorspaces but they differ in one property. Between two points in $\mathcal{Q}$ there are points which do not lie in $\mathcal{Q}$. We have for instance $\sqrt{2} \in [\frac{6}{5}, \frac{3}{2}]$ but $\sqrt{2}$ is not an element of $\mathcal{Q}$. This is not the case for the real numbers. Between two real numbers there is no no-real number. The real numbers thus form a continuum. How is this property characterised mathematically ?

We continue our example. We can approach $\sqrt{2} \in \mathcal{R}$ as close as we like with a sequence of rational numbers $x_n \in \mathcal{Q}$. The limit of this sequence is in $\mathcal{R}$ of course well-defined and given by $\sqrt{2}$. This limit is however not defined in $\mathcal{Q}$ as we walk out of the space of rational numbers into the space of real numbers. One can in $\mathcal{Q}$ however still notice 'convergence' as the difference $|x_n - x_m|$ goes to zero if n and m go to infinity. Sequences with this property are called Cauchy sequences. Cauchy sequences do not always converge in $\mathcal{Q}$ but they do converge in $\mathcal{R}$. This then finally leads to the definition of complete spaces. A complete space is a space in which every Cauchy sequence converges to an elemant of the space.

Now we can finally state the definition of a Banach space. A Banach space is a complete, normed vectorspace.

3.2.2   Example

To make things more concrete we will now give an example of a Banach space. Consider the collection of continuous functions defined on the interval $[a, b] \in \mathcal{R}$. This space we will denote by $C[a, b]$. This is a Banach space if the norm is appropriately chosen. One can check all requirements:

a] Adding of two continuous functions or multiplying a continuous function by a real number yields again a continuous function. So $C[a, b]$ is a vectorspace over the field of real or complex numbers.

b] We define the norm of a function by:

$$\|f\| = \max_{x \in [a,b]} |f(x)| \tag{3.1}$$

This makes $C[a, b]$ a normed space. One can easily check the requrements for the norm.

c] We still have to check that $C[a, b]$ with the given norm is complete space. This is always the most difficult part to check. Suppose we have a Cauchy sequence of continuous functions $\{f_n\}$. Then for any point $x \in [a, b]$:

$$|f_n(x) - f_m(x)| \leq \max_{x \in [a,b]} |f_n(x) - f_m(x)| = \|f_n - f_m\| \to 0 \qquad (n, m \to \infty)$$

This means that $f_n(x)$ for each $x \in [a, b]$ is Cauchy sequence in $\mathcal{R}$ and as $\mathcal{R}$ is complete this Cauchy sequence converges to some $f(x) \in \mathcal{R}$. This gives a limit function $f(x)$. We still have to prove that this limit function is continuous. For fixed n we have for all $x \in [a, b]$:

$$|f(x) - f_n(x)| = \lim_{m \to \infty} |f_m(x) - f_n(x)| \leq \max_{x \in [a,b]} \lim_{m \to \infty} |f_m(x) - f_n(x)| =$$

$$= \max_{x \in [a,b]} |f(x) - f_n(x)| = \|f - f_n\|$$

Thus the sequence of continuous functions converges uniformly to f and hence its limit function f is continuous and $f \in C[a, b]$.

3.2.3   The Banach spaces L^p and H^1

We now describe two types of Banach spaces which are relevant to quantum mechanics and density functional theory in particular. The first Banach space of importance is the space of functions for which the Lebesque integral

$$\int |f(x)|^p dx \tag{3.2}$$

exists where x is usually a n-dimensional real vector and p a positive real number. This space is usually denoted as $L^p(\mathcal{R}^n)$. Another important space is the space of real functions for which the Lebesque integral

$$\int |f(x)|^2 + |\nabla f(x)|^2 dx \tag{3.3}$$

exists which called the Sobolev space $H^1(\mathcal{R}^n)$ (the 1 in H^1 gives the order of the derivative). As both spaces are based on the concept of Lebesque integral we will give in the following a short account on the definition of the Lebesque integral based on measure theory. In measure theory one is interested in the question how to define a volume element for more general spaces than the real numbers. Such a definition is needed if one wants to perform integration on more general spaces Consider the space $\mathcal{B}$ and a family $\mathcal{A}$ of subsets of $\mathcal{B}$ containing the empty set $\emptyset$. A positive measure m (which is intuitively something like a volume) is then defined as a mapping from $\mathcal{A}$ into the extended real numbers $m : \mathcal{A} \to \mathcal{R}_+ \cup \{+\infty\}$ which is countably additive for every disjoint family of subsets $(a_1, \ldots, a_n, \ldots)$ in $\mathcal{A}$ with union in $\mathcal{A}$:

$$m(\cup_{i=1}^{\infty} a_i) = \sum_{i=1}^{\infty} m(a_i) \tag{3.4}$$

and

$$m(\emptyset) = 0 \tag{3.5}$$

The space $(\mathcal{B}, \mathcal{A}, m)$ endowed with measure m is called a measure space. The elements of $\mathcal{A}$ are called measurable subsets of $\mathcal{B}$ and $m(a)$ is the measure of $a \in \mathcal{A}$.

A property is said to hold almost everywhere if it holds for all points of $\mathcal{B}$ except possibly for points of a set a of measure $m(a) = 0$.

As an example we take the space of real numbers $\mathcal{R}$ and take for $\mathcal{A}$ the collection of subsets of $\mathcal{R}$ generated by the open intervals through union and intersection (which makes $\mathcal{R}$ a topological space). We define the measure m on an open interval $]a, b[, b > a$ by:

$$m(]a, b[) = b - a \tag{3.6}$$

We then have for instance for $a < b < c$:

$$m(]a, c[) = c - a = b - a + c - b = m(]a, b[) + m(]b, c[) \tag{3.7}$$

On the other hand we have:

$$m(]a, b[) + m(]b, c[) = m(]a, b[\cup]b, c[) = m(]a, c[\setminus\{b\}) \tag{3.8}$$

and

$$m(]a, c[) = m(]a, c[\setminus\{b\} \cup \{b\}) = m(]a, c[\setminus\{b\}) + m(\{b\}) \tag{3.9}$$

Hence it follows that $m(\{b\}) = 0$ and so this measure for a single point is zero. By countable addition it follows for example that the measure of the collection of rational numbers $\mathcal{Q}$ is zero:

$$m(\mathcal{Q}) = m(\cup_{i=1}^{\infty}\{q_i\}) = \sum_{i=1}^{\infty} m(\{q_i\}) = 0 \tag{3.10}$$

In this measure two functions which differ on the set $\mathcal{Q}$ are almost everywhere equal as they differ on a set of measure zero.

The measure we have defined here on the real numbers is known as the Lebesque measure and is the basis of the Lebesque integration theory. For a function $f : \mathcal{R} \to \mathcal{R}$ the Lebesque integral is defined as follows. Divide the range of f into a finite number of small intervals and find the set a_i of all $x \in \mathcal{R}$ for which $f(x)$ is in the i-th interval. Assign a measure $m(a_i)$ to the set a_i. Let k_i be some value of $f(x)$ in the i-th interval and let f_n be the step function equal to k_i when $x \in a_i$. The Lebesque integral is then the limit when it exists of:

$$\int f \, dm = \sum_i k_i m(a_i) \tag{3.11}$$

when the sequence of functions (f_n) tend to f in a sense we will make more precise in the following. In order to do this we need some further definitions.

A real function $f : \mathcal{B} \to \mathcal{R}$ on the measure space $(\mathcal{B}, \mathcal{A}, m)$ is said to be measurable when:

$$\{x | a < f(x) < b\} \in \mathcal{A}, \quad \forall a, b \in \mathcal{R} \tag{3.12}$$

A function on $(\mathcal{B}, \mathcal{A}, m)$ is called a step function or simple if it is zero except on a finite number n of disjoint sets $a_i \in \mathcal{A}$ of finite measure $m(a_i)$ where the function is equal to a finite constant k_i. The integral of a step function is by definition:

$$\int_{\mathcal{B}} f \, dm = \sum_{i=1}^{n} k_i m(a_i) \tag{3.13}$$

For example if we denote the rational numbers on the real interval $I = [0, 1]$ by $\mathcal{Q}$ and we define the step function to be equal to two on $\mathcal{Q}$ and one elsewhere then the integral in the Lebesque-measure is given by:

$$\int_I f\, dm = 2m(\mathcal{Q}) + m(I \backslash \mathcal{Q}) = 0 + 1 = 1 \tag{3.14}$$

We finally give the definition of the integral of a measurable function. We first consider the case of a positive function.Let f be a positive real valued function on the measure space $(\mathcal{B}, \mathcal{A}, m)$. Then we define the integral of f with respect to the measure m by:

$$\int_{\mathcal{B}} f\, dm \equiv \sup\left\{\int_{\mathcal{B}} \rho\, dm\right\} \tag{3.15}$$

as ρ ranges over all step functions with $0 \leq \rho \leq f$. If this integral is finite then f is said to be integrable. Now an arbitrary real valued function can always be written $f = f^+ - f^-$ with f^+ and f^- positive functions. Then f is integrable if both f^+ and f^- are integrable and its integral is given by:

$$\int_{\mathcal{B}} f\, dm \equiv \int_{\mathcal{B}} f^+\, dm - \int_{\mathcal{B}} f^-\, dm \tag{3.16}$$

The integral on the real numbers $\mathcal{R}$ or $\mathcal{R}^n$ with respect to the Lebesque measure is called Lebesque integral. The Lebesque integral of two functions is equal if the two functions are equal almost everywhere. In the remainder of this section we will regard these functions as equal. More precisely we look at the equivalence class of functions which are equal almost everywhere. In a somewhat loose notation we will denote the equivalence class to which function f belongs with the same f. We now define

$$L^p(\mathcal{R}) = \{f | \int |f(x)|^p dx < \infty\} \tag{3.17}$$

where the integral is a Lebesque integral. For $p \geq 1$ we can assign to this space a norm $\|\|_p$ which makes this space a Banach space

$$\|f\|_p = \left[\int |f(x)|^p dx\right]^{\frac{1}{p}} \tag{3.18}$$

This is a consequence of the Minkovski inequality

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \tag{3.19}$$

for $p \geq 1$. We further define the space $L^\infty(\mathcal{R})$ as the space of (classes of) measurable functions bounded almost everywhere with norm

$$\|f\|_\infty = \text{ess sup} |f(x)| \tag{3.20}$$

where the essential supremum 'ess sup' is defined as the smallest number M such that $|f(x)| \leq M$ almost everywhere. This a again a Banach space. An important inequality that can be proven for the L^p-spaces with $0 \leq p \leq \infty$ is the Hölder inequality

$$\|fg\|_1 \leq \|f\|_p \|g\|_q \tag{3.21}$$

with $1/p + 1/q = 1$ with $f \in L^p$ and $g \in L^q$. We further define the Sobolev space $H^1(\mathcal{R})$ as the Banach space

$$H^1(\mathcal{R}) = \{f \in L^2(\mathcal{R}) | \int |f(x)|^2 + |\nabla f(x)|^2 dx < \infty \tag{3.22}$$

with corresponding norm

$$\|f\| = \int \left[|f(x)|^2 + |\nabla f(x)|^2 dx \right]^{\frac{1}{2}} \tag{3.23}$$

(the derivative ∇f for $f \in L^2$ can be defined by means of a Fourier transform). We now have discussed the most important Banach spaces for applications in physics. We will in the next section discuss functions on Banach spaces which will be called operators or functionals.

3.3   Operators and functionals

Every map $F : B_1 \to B_2$ in which B_1 and B_2 are Banach spaces is called an operator. For the special case $B_2 = \mathcal{R}$ this operator is called a functional. The spaces B_1 and B_2 are in general different spaces with different norms. We can consider for example the operator $F : C[a,b] \to C[a,b]$ given by:

$$F[f] = f^2 \tag{3.24}$$

or the functional $F : C[a,b] \to \mathcal{R}$ given by:

$$F[f] = \int_a^b f^2(x)dx \tag{3.25}$$

Special operators are the linear operators from B_1 to B_2. Those operators have the property that if $f, g \in B_1$ and $\lambda \in \mathcal{R}$ then:

$$\begin{aligned} F[\lambda f] &= \lambda F[f] \\ F[f + g] &= F[f] + F[g] \end{aligned} \tag{3.26}$$

The space of linear operators from B_1 to B_2 is also a vectorspace. On this space we can also introduce a norm by:

$$\|F\| = \sup_{f \in B_1} \{ \frac{\|F[f]\|_2}{\|f\|_1} \} \tag{3.27}$$

in which $\|\|_1$ and $\|\|_2$ are the norms on the spaces B_1 and B_2. If the above operatornorm exists for an operator F then this operator is called bounded. The space of bounded linear operators from B_1 to B_2 is again a Banach space which is denoted by $[B_1, B_2]$. This space will reappear in the next section when we want to define higher order derivatives.

We now have obtained all the knowledge necessary to define integration and differentiation on Banach spaces. We can now go on and prove most of the theorems of ordinary calculus by imitating the proofs. This then yields ordinary calculus as a special case of calculus on Banach spaces.

3.4   Functional differentiation

3.4.1   The Fréchet derivative

In analogy with ordinary calculus on the vectorspace $\mathcal{R}^n$ we can define two types of derivatives. The first is the Fréchet-derivative, which is analogous to the total derivative in vector calculus, and the second is the Gateaux-derivative which is analogous to the directional derivative in vector calculus.We will start with the total derivative.

An operator $F : D_F \to B_2$ with $D_F \subseteq B_1$ in which B_1 and B_2 are Banach spaces with norms $\|\|_1$ and $\|\|_2$ is called differentiable in $f \in D_F$ if there exists a linear operator:

$$\frac{\delta F}{\delta f}(f) \in [B_1, B_2] \tag{3.28}$$

such that for $f + h \in D_F$:

$$\lim_{\|h\|_1 \to 0} \frac{\|F[f+h] - F[f] - \frac{\delta F}{\delta f}(f)[h]\|_2}{\|h\|_1} = 0 \tag{3.29}$$

If $\frac{\delta F}{\delta f}(f)$ exists for all $f \in D_F$ then this linear functional is called the Fréchet-derivative of F. This definition is analogous to the corresponding definition of the total derivative in vectorcalculus in $\mathcal{R}^n$ in which case $\frac{\delta F}{\delta f}(f)$ is a $n \times n$-matrix working on the n-component vector h.

We will now give some examples. Consider again the operator $F : C[a, b] \to C[a, b]$ defined by $F[f] = f^2$. Its Fréchet-derivative is a linear operator $\frac{\delta F}{\delta f}(f) : B_1 \to B_2$ defined by:

$$\frac{\delta F}{\delta f}(f)[h] = 2fh \in C[a, b] \tag{3.30}$$

This follows directly from the definition of $\frac{\delta F}{\delta f}$:

$$\lim_{\|h\| \to 0} \frac{\|(f+h)^2 - f^2 - \frac{\delta F}{\delta f}(f)[h]\|}{\|h\|} = \lim_{\|h\| \to 0} \frac{\|h^2\|}{\|h\|} =$$

$$= \lim_{\|h\| \to 0} \frac{1}{\|h\|} \max_{x \in [a,b]} |h^2(x)| \leq \lim_{\|h\| \to 0} \frac{1}{\|h\|} (\max_{x \in [a,b]} |h(x)|)^2 = \lim_{\|h\| \to 0} \|h\| = 0$$

Hence we have proven our statement. As a next example we will calculate the Fréchet-derivative of the functional $F : C[a, b] \to \mathcal{R}$ defined by:

$$F[f] = \int_a^b f^2(x)dx \tag{3.31}$$

Its Fréchet-derivative is given by the linear functional $\frac{\delta F}{\delta f}(f) : C[a, b] \to \mathcal{R}$ given by:

$$\frac{\delta F}{\delta f}(f)[h] = \int_a^b 2f(x)h(x)dx \tag{3.32}$$

We again check the definition:

$$\lim_{\|h\| \to 0} \frac{1}{\|h\|} |\int_a^b (f+h)^2 dx - \int_a^b f^2 dx - \int_a^b 2fh dx| = \lim_{\|h\| \to 0} \frac{1}{\|h\|} \int_a^b h^2 dx$$

$$\leq \lim_{\|h\| \to 0} \frac{b-a}{\|h\|} (\max_{x \in [a,b]} |h(x)|)^2 = \lim_{\|h\| \to 0} (b-a)\|h\| = 0$$

which proves our statement.

We now have defined the derivative of an operator as a linear operator. This is in accordance with the view of regarding the derivative of a function as a linear approximation to that function. In order to calculate higher order approximations we must carry out a Taylor-expansion and define higher order derivatives. Before addressing this question we will now give another definition of the derivative which is called the Gâteaux-derivative.

3.4.2 The Gâteaux-derivative

The Gâteaux-derivative of an operator $F : B_1 \to B_2$ in the direction $h \in B_1$ is a linear operator $\frac{\delta F}{\delta f} \in [B_1, B_2]$ defined by:

$$\frac{\delta F}{\delta f}(f)[h] = \lim_{t \to 0} \frac{F[f + th] - F[f]}{t} \tag{3.33}$$

The limit should be taken in the norm $\|\|_2$ of B_2, thus one has to check:

$$\lim_{t \to 0} \|\frac{F[f + th] - F[f]}{t} - \frac{\delta F}{\delta f}(f)[h]\|_2 = 0 \tag{3.34}$$

The existence of the Gâteaux-derivative is guaranteed if F is also Fréchet-differentiable and in that case they are equal. However existence of the Gateaux-derivative does not guarantee the existence of the Fréchet-derivative. This is easy to imagine as F might be differentiable in some directions but not in all directions, or those derivatives might not be equal.

We will now give some examples of the Gateaux-derivative. Consider $F : C[a, b] \to \mathcal{R}$ defined by:

$$F[f] = \int_a^b f^{\frac{4}{3}}(x)dx \tag{3.35}$$

Then the Gateaux-derivative $\frac{\delta F}{\delta f}(f) : C[a, b] \to \mathcal{R}$ is given by:

$$\frac{\delta F}{\delta f}(f)[h] = \lim_{t \to 0} \frac{1}{t}(\int_a^b (f + th)^{\frac{4}{3}}dx - \int_a^b f^{\frac{4}{3}}dx) =$$

$$= \lim_{t \to 0} \frac{1}{t} \int_a^b f^{\frac{4}{3}}(1 + \frac{4}{3}t\frac{h}{f} + O(t^2)) - f^{\frac{4}{3}}dx =$$

$$= \lim_{t \to 0} \int_a^b \frac{4}{3}f^{\frac{1}{3}}hdx + O(t) = \int_a^b \frac{4}{3}f^{\frac{1}{3}}hdx$$

Another important example is the derivation of the Euler-Lagrange equations. Consider the Banach space of continuously differentiable functions on some subspace V of three-dimensional space, denoted by $C^1(V)$. We can then define the functional $F : C^1(V) \to \mathcal{R}$ by:

$$F[f] = \int_V \mathcal{L}(f, \nabla f)d\mathbf{r} \tag{3.36}$$

in which $\mathcal{L}$ is a local function of f and ∇f. The functions f and ∇f are supposed to disappear on the edge of V. Then the Gateaux-derivative of F is given by:

$$\frac{\delta F}{\delta f}(f)[h] = \lim_{t \to 0} \frac{1}{t} \int_V \mathcal{L}(f + th, \nabla f + t\nabla h) - \mathcal{L}(f, \nabla f)d\mathbf{r} =$$

$$= \lim_{t \to 0} \frac{1}{t} \int_V t\frac{\partial \mathcal{L}}{\partial f}h + t\frac{\mathcal{L}}{\partial \nabla f} \cdot \nabla h + O(t^2)d\mathbf{r} =$$

$$= \int_V \frac{\partial \mathcal{L}}{\partial f}h + \frac{\partial \mathcal{L}}{\partial \nabla f} \cdot \nabla h d\mathbf{r} = \int_V (\frac{\partial \mathcal{L}}{\partial f} - \nabla \cdot \frac{\partial \mathcal{L}}{\partial \nabla f})h d\mathbf{r} \tag{3.37}$$

If we are looking for extreme values of F then $\frac{\delta F}{\delta f}(f)[h]$ must be zero for all variations h. In that case we must have:

$$\frac{\partial \mathcal{L}}{\partial f} - \nabla \cdot \frac{\partial \mathcal{L}}{\partial \nabla f} = 0 \tag{3.38}$$

which yields the famous Euler-Lagrange equations.

3.4.3 Higher order derivatives

We will now adress the question of higher order derivatives. We note that if $F : B_1 \to B_2$ then $\frac{\delta F}{\delta f}(f)$ is a linear operator. It is an element of the space $[B_1, B_2]$. So we can define an operator $\frac{\delta F}{\delta f} : B_1 \to [B_1, B_2]$ which assigns to each element $f \in B_1$ the functional derivative $\frac{\delta F}{\delta f}(f)$ in f. Thus:

$$\frac{\delta F}{\delta f} : f \mapsto \frac{\delta F}{\delta f}(f) \in [B_1, B_2] \tag{3.39}$$

As the space $[B_1, B_2]$ is also a Banach space we can differentiate $\frac{\delta F}{\delta f}$ again and obtain a linear operator:

$$\frac{\delta^2 F}{\delta f^2}(f) : B_1 \to [B_1, B_2] \tag{3.40}$$

thus $\frac{\delta^2 F}{\delta f^2}(f) \in [B_1, [B_1, B_2]]$. This is called the second order derivative of F. We can now start allover again and define an operator:

$$\frac{\delta^2 F}{\delta f^2} : f \mapsto \frac{\delta^2 F}{\delta f^2}(f) \in [B_1, [B_1, B_2]] \tag{3.41}$$

This operator can then be further differentiated and so on. If we go on like this , this procedure finally gives our n^{th}-order derivative which is a functional:

$$\frac{\delta^n F}{\delta f^n} : f \mapsto \frac{\delta^n F}{\delta f^n}(f) \in \underbrace{[B_1, [B_1, [B_1, \dots [B_1}_{n}, B_2] \dots]]] \tag{3.42}$$

Let us make this more concrete with an example. Take the functional $F : C[a, b] \to \mathcal{R}$ defined by:

$$F[f] = \int_a^b f^2(x) dx \tag{3.43}$$

The functional derivative is a linear functional $\frac{\delta F}{\delta f}(f) : C[a, b] \to \mathcal{R}$ given by:

$$\frac{\delta F}{\delta f}(f)[h] = \int_a^b 2f(x)h(x) dx \tag{3.44}$$

We will denote this functional by $'2f'$. We can now define the operator $\frac{\delta F}{\delta f} : C[a, b] \to [C[a, b], \mathcal{R}]$ by:

$$\frac{\delta F}{\delta f}[f] =' 2f' \tag{3.45}$$

This assigns to each function f a linear functional. We now can differentiate this operator again. This yields again a linear operator which is an element of $[C[a, b], [C[a, b], \mathcal{R}]]$ defined by:

$$\frac{\delta^2 F}{\delta f^2}(f)[h] = 2'h' \tag{3.46}$$

The proof of this statement is not difficult. If we denote the norm on the space $[C[a, b], \mathcal{R}]$ by $\| \|_3$ then we have:

$$\lim_{\|h\|_1 \to 0} \frac{1}{\|h\|_1} \| \frac{\delta F}{\delta f}(f + h) - \frac{\delta F}{\delta f}(f) - \frac{\delta^2 F}{\delta f^2}(f)[h] \|_3 =$$

$$= \lim_{\|h\|_1 \to 0} \frac{\|'2(f+h)' -' 2f' - 2'h'\|_3}{\|h\|_1} = \lim_{\|h\|_1 \to 0} \frac{\|'0'\|_3}{\|h\|_1} = 0$$

As $\frac{\delta^2 F}{\delta f^2}$ assigns to each point $f \in C[a,b]$ a linear operator working on a point $h \in C[a,b]$ which in its turn assigns a linear operator working on a point $k \in C[a,b]$ we can also regard $\frac{\delta^2 F}{\delta f^2}$ as an operator which assigns to each point $f \in C[a,b]$ a bilinear operator on $C[a,b]$:

$$\frac{\delta^2 F}{\delta f^2} : f \mapsto \frac{\delta^2 F}{\delta f^2}(f) \in [C[a,b] \times C[a,b], \mathcal{R}] \tag{3.47}$$

defined by:

$$\frac{\delta^2 F}{\delta f^2}(f)[h,k] = 2'h'[k] = \int_a^b 2h(x)k(x)dx \tag{3.48}$$

This is completely general. We can equivalently view $\frac{\delta^n F}{\delta f^n}$ as an operator:

$$\frac{\delta^n F}{\delta f^n} : B_1 \to \underbrace{[B_1 \times \ldots \times B_1}_{n}, B_2] \tag{3.49}$$

whichs assigns to each $f \in B_1$ a multilinear operator on the product space $B_1 \times \ldots \times B_1$ which is also a Banach space:

$$\frac{\delta^n F}{\delta f^n}(f)(h_1, \ldots, h_n) \mapsto \frac{\delta^n F}{\delta f^n}(f)[h_1, \ldots, h_n] \in B_2 \tag{3.50}$$

This is analogous to the case of vector calculus in which the nth-order derivatives are multilinear tensors. In our example we have as a first order derivative the linear functional:

$$\frac{\delta F}{\delta f}(f)[h] = \int_a^b 2f(x)h(x)dx \tag{3.51}$$

and as a second order derivative the bilinear functional:

$$\frac{\delta^2 F}{\delta f^2}(f)[h,k] = \int_a^b 2h(x)k(x)dx \tag{3.52}$$

Those operators are often given by their integral kernels when it is possible to regard the functional derivative as a linear integral operator. In the case of our example we have:

$$\frac{\delta F}{\delta f}(f)[h] = \int_a^b \frac{\delta F}{\delta f(x)}h(x)dx \tag{3.53}$$

$$\frac{\delta^2 F}{\delta f^2}(f)[h,k] = \int_a^b \int_a^b \frac{\delta^2 F}{\delta f(x)\delta f(y)}h(x)k(y)dxdy \tag{3.54}$$

where the integral kernels are given by:

$$\frac{\delta F}{\delta f(x)} = 2f(x) \tag{3.55}$$

$$\frac{\delta^2 F}{\delta f(x)\delta f(y)} = 2\delta(x-y) \tag{3.56}$$

In general we can write:

$$\frac{\delta^n F}{\delta f^n}(f)[h_1, \ldots, h_n] = \int \frac{\delta^n F}{\delta f(x_1) \ldots \delta f(x_n)}h_1(x_1) \ldots h_n(x_n)dx_1 \ldots dx_n \tag{3.57}$$

This completes our discussing of the higher order derivatives. We can now go on to define Taylor sequences, but before we do that we will first give extremal points of functionals and give an overview on functional integration.

3.5 Variational calculus

3.5.1 Extremal points of functionals

One of the most important applications of functional calculus concerns the determination of extremal points. One usually looks for local or global minima of functionals. Important applications can be found in almost all areas of physics.

In the following we will discuss local minima of a functional $F : B \mapsto \mathcal{R}$ from a Banach space B to the real numbers. All the results derived can easily be applied to the case of local maxima as they are local minima of the functional $-F$. We first state a definition. A functional F has a local minimum at $f_0 \in B$ when there exists a neigbourhood $V(f_0) = \{f \in B | \|f - f_0\| < \epsilon\}$ such that

$$F[f_0] \leq F[f] \qquad \forall f \in V(f_0) \tag{3.58}$$

Let F be Fréchet differentiable (Gâteaux differentiable in every direction $h \in B$). Then a necessary condition for F to have a local minimum at f_0 is

$$\frac{\delta F}{\delta f}(f_0) = 0 \tag{3.59}$$

For a proof we define $g : \mathcal{R} \mapsto \mathcal{R}$ by $g(t) = F[f_0 + th]$ for $h \in B$. As f_0 is a local minimum we have $g(0) = F[f_0]$ and $g(t) \geq F[f_0]$. Therefor $g : \mathcal{R} \mapsto \mathcal{R}$ has a local minimum at $t = 0$ and therefore $dg/dt(0) = 0$ and we find

$$0 = \frac{dg}{dt}(0) = \frac{\delta F}{\delta f}(f_0)[h] \qquad \forall h \in B \tag{3.60}$$

So $\delta F/\delta f(f_0) = 0$. The following theorem is useful for practical applications. Let $F : B \mapsto \mathcal{R}$ be twice differentiable. A sufficient condition for f_0 to have a minimum at $f_0 \in B$ is that

$$\frac{\delta F}{\delta f}(f_0) = 0 \tag{3.61}$$

and

$$\frac{\delta^2 F}{\delta f^2}(f)[h, h] \geq 0 \qquad \forall f, h \in B \tag{3.62}$$

To prove this we define $k : \mathcal{R} \mapsto \mathcal{R}$ by

$$k(t) = F[f_0 + th] + (1 - t)\frac{\delta F}{\delta f}(f_0 + th)[h] \tag{3.63}$$

Then $k(1) = F[f_0 + h]$ and $k(0) = F[f_0]$ because the derivative at f_0 vanishes

$$\frac{dk}{dt}(t) = (1 - t)\frac{\delta^2 F}{\delta f^2}(f_0 + th)[h, h] \tag{3.64}$$

Which yields

$$F[f_0 + h] - F[f_0] = k(1) - k(0) = \int_0^1 (1 - t)\frac{\delta^2 F}{\delta f^2}(f_0 + th)[h, h]dt \geq 0 \tag{3.65}$$

and we find that $F[f_0 + h] \geq F[f_0]$, and therefore f_0 must be a local minimum.

3.5.2   Convex functionals

An important class of functionals for variational calculations is the set of convex functionals. A functional $F : B \mapsto \mathcal{R}$ is convex when for $f_1, f_2 \in B$ and $0 \leq \lambda_1, \lambda_2 \leq 1$ with $\lambda_1 + \lambda_2 = 1$ we have

$$F[\lambda_1 f_1 + \lambda_2 f_2] \leq \lambda_1 F[f_1] + \lambda_2 F[f_2] \tag{3.66}$$

It is readily verified that an equivalent definition is

$$F[\sum_{i=1}^{N} \lambda_i f_i] \leq \sum_{i=1}^{N} \lambda_i F[f_i] \quad \sum_{i=1}^{N} \lambda_i = 1 \quad 0 \leq \lambda_i \leq 1 \tag{3.67}$$

The importance of convexity for variational problems can be inferred from the following statement. If $F : B \mapsto \mathcal{R}$ is convex and F is twice differentiable then

$$\frac{\delta^2 F}{\delta f^2}(f)[h, h] \geq 0 \qquad \forall f, h \in B \tag{3.68}$$

To prove this we first note that due to the convexity of F we have for $0 \leq t \leq 1$

$$F[f_0 + t(f - f_0)] = F[tf + (1 - t)f_0] \leq tF[f] + (1 - t)F[f_0] \tag{3.69}$$

and therefore

$$\frac{F[f_0 + t(f - f_0)] - F[f_0]}{t} \leq F[f] - F[f_0] \tag{3.70}$$

Taking the limit $t \downarrow 0$ and using the differentiability of F we find

$$\frac{\delta F}{\delta f}(f_0)[f - f_0] \leq F[f] - F[f_0] \tag{3.71}$$

Similarly by interchanging f and f_0 we find

$$\frac{\delta F}{\delta f}(f)[f_0 - f] \leq F[f_0] - F[f] \tag{3.72}$$

Adding both inequalities then yields

$$\left(\frac{\delta F}{\delta f}(f) - \frac{\delta F}{\delta f}(f_0) \right)[f - f_0] \geq 0 \tag{3.73}$$

This inequality means that $\delta F / \delta f$ is a monotonously increasing functional. If we take $f = f_0 + th$ and take the limit $t \downarrow 0$ and use the fact that F is twice differentiable we have

$$0 \leq \lim_{t \downarrow 0} \frac{1}{t} \left(\frac{\delta F}{\delta f}(f_0 + th) - \frac{\delta F}{\delta f}(f_0) \right)[h] = \frac{\delta^2 F}{\delta f^2}(f_0)[h, h] \tag{3.74}$$

which proves our statement. The converse, which we will not prove here, is also true. If F is twice differentiable and the second derivative is always larger than or equal to zero then F is convex. From the results of the previous section we can now conclude that if $\delta F / \delta f(f_0) = 0$ for a twice differentiable convex functional then f_0 is a local minimum of F.

We will now discuss one other quantity which is useful to prove differentiability of convex functionals, which is the tangent functional or subgradient. Suppose that $F : B \mapsto \mathcal{R}$ is a convex functional which is differentiable at f_0 then

$$F[f] - F[f_0] \geq \frac{\delta F}{\delta f}(f_0)[f - f_0] \qquad \forall f \in B \tag{3.75}$$

This equation which has been derived above is equivalent to equation 3.71. The above equation is a special case of the more general equation

$$F[f] - F[f_0] \geq L[f - f_0] \qquad \forall f \in B \tag{3.76}$$

where $L : B \mapsto \mathcal{R}$ is a bounded linear functional. If for a convex functional F such a linear functional exists at f_0 then F is called subdifferentiable at f_0. The functional L is called a subgradient or a tangent functional. One can prove the following statement. For a proof see [66]. If there is for a convex functional F a unique tangent functional at $f_0 \in B$ then F is Gâteaux differentiable at f_0 and

$$L = \frac{\delta F}{\delta f}(f_0) \tag{3.77}$$

Gâteaux differentiability of a convex functional is therefore equivalent to the uniqueness of a tangent functional. This can be illustrated with the following example. Let the Banach space B be the real numbers $B = \mathcal{R}$ and the function $F : \mathcal{R} \mapsto \mathcal{R}$ defined as $F(x) = |x|$. This is a convex function. The linear functions $L_1(x) = x$ and $L_2(x) = -x$ are both subgradients at the point $x = 0$. Therefore there is no unique subgradient at $x = 0$ and F is not differentiable at $x = 0$. For $x > 0$ there is the unique subgradient L_1 and for $x < 0$ there is a unique subgradient L_2. Therefore F is differentiable for $x \neq 0$.

3.6 Functional integration

3.6.1 The line integral

As a motivation for the definition of line integrals in Banach spaces we will first give a familiar example of line integration in ordinary vector calculus. Suppose we have a curve $\vec{\gamma} : \mathcal{R} \to \mathcal{R}^3$ in three-dimensional space starting at $\vec{a} = \vec{\gamma}(0)$ and ending in $\vec{b} = \vec{\gamma}(1)$. Suppose we also have a scalar function $F : \mathcal{R}^3 \to \mathcal{R}$ and we define $F_\gamma(t) = F(\vec{\gamma}(t))$ then it follows:

$$\frac{dF_\gamma}{dt}(t) = \nabla F(\vec{\gamma}(t)) \cdot \frac{d\vec{\gamma}}{dt}(t) \tag{3.78}$$

From this formula we can deduce:

$$F(\vec{b}) - F(\vec{a}) = \int_0^1 \frac{dF_\gamma}{dt}(t)dt = \int_0^1 \nabla F(\vec{\gamma}(t)) \cdot \frac{d\vec{\gamma}}{dt}(t)dt \tag{3.79}$$

In general we can for a given vectorfield $\vec{v}$ calculate the line integral along curve $\vec{\gamma}$ which is defined as:

$$\int_{\vec{\gamma}} \vec{v} \cdot d\vec{\gamma} \equiv \int_0^1 \vec{v}(\vec{\gamma}(t)) \cdot \frac{d\vec{\gamma}}{dt}(t)dt \tag{3.80}$$

In general the outcome of this integral is path dependent.Only if there exists some scalar function F with $\nabla F = \vec{v}$ then the outcome of the integral is path independent. In that case we have a conservative vectorfield.
Let us generalise this to Banach spaces. Consider a functional $F : B \to \mathcal{R}$ where B is Banach space with norm $\|\|$. We can then define for a given $f, h \in B$ the function $k : \mathcal{R} \to \mathcal{R}$ by:

$$k(t) = F[f + th] \tag{3.81}$$

Then:

$$\frac{dk}{dt}(t) = \lim_{s \to 0} \frac{k(t+s) - k(t)}{s} =$$

$$= \lim_{s \to 0} \frac{F[f + th + sh] - F[f + th]}{s} = \frac{\delta F}{\delta f}(f + th)[h]$$

This thus yields the Gateaux-derivative in point $f + th$. This then gives:

$$F[f + h] - F[f] = k(1) - k(0) = \int_0^1 \frac{dk}{dt}(t)dt = \int_0^1 \frac{\delta F}{\delta f}(f + th)[h]dt \qquad (3.82)$$

So we have derived the formula:

$$F[f_b] - F[f_a] = \int_0^1 \frac{\delta F}{\delta f}(f_a + t(f_b - f_a))[f_b - f_a]dt \qquad (3.83)$$

This defines the line integral of $\frac{\delta F}{\delta f} \in [B, [B, \mathcal{R}]]$ along the path $\gamma(t) = f_a + t(f_b - f_a) \in B$. This can be generalised to arbitrary elements $v \in [B, [B, \mathcal{R}]]$. Suppose we have an operator $v : B \to [B, \mathcal{R}]$ then the integral of v from f_a to f_b is defined by:

$$\int_{f_a}^{f_b} v[f]df \equiv \int_0^1 v(f_a + t(f_b - f_a))[f_b - f_a]dt \qquad (3.84)$$

(By $v(f)[h]$ we mean v in point $f \in B$ working on $h \in B$.) This integral can also be defined for more general paths by splitting up the general path in small straight pieces on which the integral is defined and then taking the limit of an infinite number of infinitesimal pieces. This then leads to the following definition:

$$\int_\gamma v[f]df \equiv \int_0^1 v(\gamma(t))[\frac{d\gamma}{dt}(t)]dt \qquad (3.85)$$

where $\gamma : \mathcal{R} \to B$ is a differentiable path in Banach space. Using this definition we can prove the following statements just as for ordinary line integrals:

$$\int_\gamma v[f]df = -\int_{-\gamma} v[f]df \qquad (3.86)$$

$$\int_{\gamma_1} v[f]df + \int_{\gamma_2} v[f]df = \int_{\gamma_1 + \gamma_2} v[f]df \qquad (3.87)$$

where $-\gamma$ is γ with reverse orientation thus with parameter t running from 1 to 0. Further is $\gamma_1 + \gamma_2$ the combined path obtained by first walking along path γ_1 and subsequently walking along path γ_2. If the integrals are path independent the outcome of the integrals can only depend on the endpoints of the path and we can denote the integral by:

$$\int_\gamma v[f]df = \int_{f_a}^{f_b} v[f]df \qquad (3.88)$$

where $f_a = \gamma(0)$ and $f_b = \gamma(1)$. In that case the above statements give:

$$\int_{f_a}^{f_b} v[f]df = -\int_{f_b}^{f_a} v[f]df \qquad (3.89)$$

$$\int_{f_a}^{f_c} v[f]df + \int_{f_c}^{f_b} v[f]df = \int_{f_a}^{f_b} v[f]df \tag{3.90}$$

Suppose we have an operator $v \in [B, [B, \mathcal{R}]]$ for which the line integral is path independent. We will prove for the case of a straight path (the particular path is not important as the integral is path independent) that if we define the functional $F : B \to \mathcal{R}$ by

$$F[f] = \int_{f_a}^{f} v[k]dk = \int_0^1 v(f_a + t(f - f_a))[f - f_a]dt \tag{3.91}$$

with $v \in [B, [B, \mathcal{R}]]$ then

$$\frac{\delta F}{\delta f}(f) = v(f) \in [B, \mathcal{R}] \tag{3.92}$$

The proof is analogous to the corresponding proof for the Riemann integral. We calculate:

$$\frac{F[f + th] - F[f]}{t} = \frac{1}{t}\int_f^{f+th} v[k]dk = \frac{1}{t}\int_0^t v(f + sh)[h]ds \tag{3.93}$$

Here we used the path independence property:

$$\int_{f_a}^{f+th} v[k]dk - \int_{f_a}^{f} v[k]dk = \int_f^{f+th} v[k]dk \tag{3.94}$$

Then it follows:

$$\left|\frac{F[f + th] - F[f]}{t} - v(f)[h]\right| = \left|\frac{1}{t}\int_0^t (v(f + sh)[h] - v(f)[h])ds\right|$$

$$\leq \frac{1}{t}\int_0^1 |v(f + sh)[h] - v(f)[h]|ds \leq \frac{1}{t}\max_{s \in [0,t]}\{|v(f + sh)[h] - v(f)[h]|t\} =$$

$$= \max_{s \in [0,t]}\{|v(f + sh)[h] - v(f)[h]|\} \to 0 \qquad (t \to 0)$$

Thus we have:

$$\frac{\delta F}{\delta f}(f)[h] = \lim_{t \to 0} \frac{F[f + th] - F[f]}{t} = v(f)[h] \tag{3.95}$$

and hence we have proven our statement. Thus path independency for line-integrals of v implies that $v(f)$ is the functional derivative of some functional $F : B \to \mathcal{R}$ We will now give some examples. Suppose we have an operator $v : C[a, b] \to [C[a, b], \mathcal{R}]$ defined by:

$$v[f] = \frac{4}{3}f^{\frac{1}{3}} \in [C[a, b], \mathcal{R}] \tag{3.96}$$

The action of this operator on a function $h \in C[a, b]$ is defined as:

$$v(f)[h] = \frac{4}{3}f^{\frac{1}{3}}[h] = \int_a^b \frac{4}{3}f^{\frac{1}{3}}(x)h(x)dx \tag{3.97}$$

Then the integral of v from f_a to f_b is given by:

$$\int_{f_a}^{f_b} v[k]dk = \int_0^1 \frac{4}{3}(f_a + t(f_b - f_a))^{\frac{1}{3}}[f_b - f_a]dt =$$

$$= \int_a^b dx \frac{4}{3}(f_b - f_a) \int_0^1 dt(f_a + t(f_b - f_a))^{\frac{1}{3}} =$$

$$= \int_a^b dx \frac{4}{3}(f_b - f_a)[\frac{3}{4}\frac{1}{f_b - f_a}(f_a + t(f_b - f_a))^{\frac{4}{3}}]_0^1 =$$

$$= \int_a^b dx(f_b^{\frac{4}{3}} - f_a^{\frac{4}{3}}) = F[f_b] - F[f_a]$$

where $F : C[a,b] \to \mathcal{R}$ is defined by:

$$F[f] = \int_a^b f^{\frac{4}{3}}(x)dx \tag{3.98}$$

One can check that the functional derivative of F yields $v(f)$ and hence the result is path independent. Our final example in this section also involves gradients. Define the functional $F : C^1(V) \to \mathcal{R}$ in which $C^1(V)$ is the space of continuous differentiable functions on some subset of three-dimensional space $\mathcal{R}^3$ by:

$$F[f] = \int_V \frac{(\nabla f)^2}{f^{\frac{4}{3}}}d\mathbf{r} \tag{3.99}$$

The functions f and ∇f are supposed to disappear on the edge of V. Using the Euler-Lagrange equations 3.37 we find for the functional derivative:

$$v(f) = \frac{\delta F}{\delta f}(f) = \frac{\partial}{\partial f}(\frac{(\nabla f)^2}{f^{\frac{4}{3}}}) - \nabla \cdot \frac{\partial}{\partial \nabla f}(\frac{(\nabla f)^2}{f^{\frac{4}{3}}}) =$$

$$= -\frac{4}{3}\frac{(\nabla f)^2}{f^{\frac{7}{3}}} - \nabla \cdot (2\frac{\nabla f}{f^{\frac{4}{3}}}) = -\frac{4}{3}\frac{(\nabla f)^2}{f^{\frac{7}{3}}} - 2\frac{\nabla^2 f}{f^{\frac{4}{3}}} + \frac{8}{3}\frac{(\nabla f)^2}{f^{\frac{7}{3}}} =$$

$$= \frac{4}{3}\frac{(\nabla f)^2}{f^{\frac{7}{3}}} - 2\frac{\nabla^2 f}{f^{\frac{4}{3}}} \in [C^1(V), \mathcal{R}] \tag{3.100}$$

We will now integrate this derivative from $0 \in C^1(V)$ to $f \in C^1(V)$. Then we must calculate:

$$\int_0^f v[k]dk = \int_0^1 v(tf)[f]dt = \int_V d\mathbf{r}f \int_0^1 dt \frac{4}{3}\frac{(t\nabla f)^2}{(tf)^{\frac{7}{3}}} - 2\frac{t\nabla^2 f}{(tf)^{\frac{4}{3}}} =$$

$$= \int_V d\mathbf{r}f(\frac{4}{3}\frac{(\nabla f)^2}{f^{\frac{7}{3}}}\int_0^1 dt t^{-\frac{1}{3}} - 2\frac{\nabla^2 f}{f^{\frac{4}{3}}}\int_0^1 dt t^{-\frac{1}{3}}) =$$

$$= \int_V d\mathbf{r}(2\frac{(\nabla f)^2}{f^{\frac{4}{3}}} - 3\frac{\nabla^2 f}{f^{\frac{1}{3}}}) = \int_V d\mathbf{r}\frac{(\nabla f)^2}{f^{\frac{4}{3}}} - 3\int_V d\mathbf{r}\nabla \cdot (\frac{\nabla f}{f^{\frac{1}{3}}}) =$$

$$= \int_V \frac{(\nabla f)^2}{f^{\frac{4}{3}}} = F[f] \tag{3.101}$$

which recovers our starting functional.

3.6.2 Integrability conditions

In the calculation of line integrals it is often useful to know whether a given line integral is path independent because in that case we can try to deform the path to a contour for which the integral is easy to calculate without changing the outcome of the integral. We thus like to have an easy to verify criterion which tells us whether a given line integral is path independent. For everyday three-dimensional vectorfields such a criterion is easy to give. The line integral along the field is path independent if this field is rotationless. Mathematically this means that its curl should vanish. This condition is readily derived. Suppose we have a vectorfield $v : \mathcal{R}^3 \to \mathcal{R}^3$. If the line integral along this vectorfield is path independent then this vectorfield should be the gradient of some scalar function $f : \mathcal{R}^3 \to \mathcal{R}$ thus:

$$v = \nabla f \tag{3.102}$$

Or in terms of vector components:

$$v_i = \partial_i f \tag{3.103}$$

If f is twice differentiable then it follows that:

$$\partial_k v_i - \partial_i v_k = \partial_k \partial_i f - \partial_i \partial_k f = 0 \tag{3.104}$$

and hence:

$$\nabla \times v = 0 \tag{3.105}$$

In the derivation of this condition we used the fact that we could interchange the differentiation with respect to different variables. In general it is true for a n-times differentiable function $f : \mathcal{R}^3 \to \mathcal{R}$ that:

$$\partial_{i_1} \partial_{i_2} \ldots \partial_{i_n} f = \partial_{i_{p(1)}} \partial_{i_{p(2)}} \ldots \partial_{i_{p(n)}} f \tag{3.106}$$

where p is an arbitrary permutation of the numbers $1, 2, \ldots, n$. A direct generalisation of this statement to n-times differentiable operators $F : B_1 \to B_2$ on Banach space would be:

$$\frac{\delta^n F}{\delta f^n}(f)[h_1, h_2, \ldots, h_n] = \frac{\delta^n F}{\delta f^n}(f)[h_{p(1)}, h_{p(2)}, \ldots, h_{p(n)}] \tag{3.107}$$

where $h_i \in B_1$. Using the definition of differentiation one can prove this statement. In particular we have for the bilinear operator $\frac{\delta^2 F}{\delta f^2}(f) \in [B_1 \times B_1, B_2]$ that:

$$\frac{\delta^2 F}{\delta f^2}(f)[h, k] = \frac{\delta^2 F}{\delta f^2}(f)[k, h] \tag{3.108}$$

for $h, k \in B_1$. If it is possible to view this operator as an integral operator then the integral kernel should be symmetric:

$$\frac{\delta^2 F}{\delta f(x) \delta f(y)} = \frac{\delta^2 F}{\delta f(y) \delta f(x)} \tag{3.109}$$

Suppose we have in point $f \in B_1$ a linear operator $v(f) : B_1 \to B_2$ which is the functional derivative of an operator $F : B_1 \to B_2$, thus:

$$v(f) = \frac{\delta F}{\delta f}(f) \in [B_1, B_2] \tag{3.110}$$

If we let this operator work on an element $h \in B_1$ and we take the derivative in direction $k \in B_1$ then:

$$\frac{\delta v}{\delta f}(f)[h][k] = \lim_{t \to 0} \frac{v(f + tk) - v(f)}{t}[h] =$$

$$= \lim_{t \to 0} \frac{1}{t}(\frac{\delta^2 F}{\delta f^2}(f + tk)[h] - \frac{\delta F^2}{\delta f^2}(f)[h]) = \frac{\delta^2 F}{\delta f^2}(f)[h, k] \tag{3.111}$$

Thus using the symmetry of the second order derivative (equation 3.108) we can conclude that if $v \in [B_1, [B_1, B_2]$ is the derivative of a twice differentiable operator $F : B_1 \to B_2$ then:

$$\frac{\delta v}{\delta f}(f)[h][k] = \frac{\delta v}{\delta f}(f)[k][h] \tag{3.112}$$

Using the integral kernel notation:

$$\frac{\delta v}{\delta f}(f)[h][k] = \int \frac{\delta v(x)}{\delta f(y)} h(x)k(y)dxdy \tag{3.113}$$

we must have for the integral kernels:

$$\frac{\delta v(x)}{\delta f(y)} = \frac{\delta v(y)}{\delta f(x)} \tag{3.114}$$

This is a necessary condition for the path independence for the line integral of v. We now give some examples. Consider a functional $F : C^2[a, b] \to \mathcal{R}$ on the space $C^2[a, b]$ of twice continuously differentiable functions on the interval $[a, b]$. The functions are also required to vanish at the endpoints of the interval so we are really working in a subspace of $C^2[a, b]$. Let the functional F be defined by:

$$F[f] = \int_a^b \frac{1}{2}(\frac{df}{dx}(x))^2 dx \tag{3.115}$$

Its functional derivative $v(f) = \frac{\delta F}{\delta f}(f) \in [C^2[a, b], \mathcal{R}]$ is given by:

$$v(f) = \frac{d^2 f}{dx^2} \in [C^2[a, b], \mathcal{R}] \tag{3.116}$$

This is easily derived using the Euler-Lagrange equations 3.37. Its action on a function $h \in C^2[a, b]$ is given by:

$$\frac{d^2 f}{dx^2}[h] = \int_a^b \frac{d^2 f}{dx^2}(x)h(x)dx \tag{3.117}$$

As $v(f) \in [C^2[a, b], \mathcal{R}]$ is the functional derivative of functional F the operator $\frac{\delta v}{\delta f} \in [C^2[a, b], [C^2[a, b], \mathcal{R}]]$ should be a symmetric linear operator. This is easily checked:

$$\frac{\delta v}{\delta f}(f)[h][k] = \lim_{t \to 0} \frac{1}{t} \int_a^b (\frac{d^2}{dx^2}(f + tk) - \frac{d^2 f}{dx^2})h(x)dx =$$

$$= \int_a^b \frac{d^2 k}{dx^2}(x)h(x)dx = -\int_a^b \frac{dk}{dx}(x)\frac{dh}{dx}(x)dx \tag{3.118}$$

In the last step we used partial integration and the fact that $h(a) = h(b) = 0$. In this derivation all the limits are taken in the norm sense. A more careful derivation using norms shows that this last

result is indeed correct. This last formula is obviously symmetric in h and k. If we use the integral kernel notation we have:

$$\frac{\delta v(x)}{\delta f(y)} = \frac{d^2}{dx^2}\delta(x - y) \tag{3.119}$$

which is also symmetric.

As a final example we consider on the same space an operator $v(f) \in [C^2[a,b], \mathcal{R}]$ with a nonsymmetric derivative. Let $v(f)$ be defined by:

$$v(f)[h] = \int_a^b (\frac{df}{dx}(x))^2 h(x)dx \tag{3.120}$$

Its derivative is given by:

$$\frac{\delta v}{\delta f}(f)[h][k] = \lim_{t \to 0} \frac{1}{t} \int_a^b ((\frac{d}{dx}(f + tk))^2 - (\frac{df}{dx})^2)h(x)dx =$$

$$= \int_a^b 2\frac{df}{dx}(x)\frac{dk}{dx}(x)h(x)dx \tag{3.121}$$

This formula is not symmetric in h and k. We can also write this as:

$$\frac{\delta v}{\delta f}(f)[h][k] = \int_a^b \int_a^b \frac{df}{dx}(x)\frac{d\delta(x - y)}{dy}h(x)k(y)dxdy \tag{3.122}$$

So the integral kernel is given by:

$$\frac{\delta v(x)}{\delta f(y)} = \frac{df}{dx}\frac{d}{dy}\delta(x - y) \tag{3.123}$$

This kernel is nonsymmetric so any line integral along v will be path dependent. This ends our discussion of functional integration. There is one important topic left which is very useful in the approximation of functionals and that is the definition of Taylor-sequences. We will deal with this question in the next section.

3.7   Taylor expansions

In this section we will derive the Taylor formula for operators $F : B_1 \to B_2$. For given $f, h \in B_1$ define the function $k : \mathcal{R} \to B_2$ by:

$$k(t) = \sum_{n=0}^N \frac{1}{n!}(1 - t)^n \frac{\delta^n F}{\delta f^n}(f + th)\underbrace{[h, \ldots, h]}_{n} \tag{3.124}$$

we now will use the following formula for the Riemann integral:

$$k(1) - k(0) = \int_0^1 \frac{dk}{dt}(t)dt \tag{3.125}$$

in order to calculate $\frac{dk}{dt}(t)$ we must calculate:

$$\frac{d}{dt}(\frac{\delta^n F}{\delta f^n}(f + th)) = \lim_{s \to 0} \frac{1}{s}(\frac{\delta^n F}{\delta f^n}(f + th + sh)[h, \ldots, h] - \frac{\delta^n F}{\delta f^n}(f + th)[h, \ldots, h]) =$$

$$= \frac{\delta^{n+1} F}{\delta f^{n+1}} (f + th) \underbrace{[h, \ldots, h]}_{n} [h] = \frac{\delta^{n+1} F}{\delta f^{n+1}} (f + th) \underbrace{[h, \ldots, h]}_{n+1} \tag{3.126}$$

It thus follows that:

$$\frac{dk}{dt}(t) = \sum_{n=1}^{N} \frac{-1}{(n-1)!} (1-t)^{n-1} \frac{\delta^n F}{\delta f^n} (f + th)[h, \ldots, h]$$

$$+ \sum_{n=0}^{N} \frac{1}{n!} (1-t)^n \frac{\delta^{n+1} F}{\delta f^{n+1}} (f + th)[h, \ldots, h] = \frac{1}{N!} (1-t)^N \frac{\delta^{N+1} F}{\delta f^{N+1}} (f + th)[h, \ldots, h] \tag{3.127}$$

We further have:

$$k(1) - k(0) = F[f + h] - \sum_{n=0}^{N} \frac{1}{n!} \frac{\delta^n F}{\delta f^n} (f)[h, \ldots, h] \tag{3.128}$$

Using equation 3.125 for the Riemann integral this yields:

$$F[f + h] = F[f] + \sum_{n=1}^{N} \frac{1}{n!} \frac{\delta^n F}{\delta f^n} (f)[h, \ldots, h] + \int_0^1 \frac{1}{N!} (1-t)^N \frac{\delta^{N+1} F}{\delta f^{N+1}} (f + th)[h, \ldots, h] dt \tag{3.129}$$

If $F : B_1 \to B_2$ is infinitely differentiable and if the restterm:

$$R_N(f, h) = \int_0^1 \frac{1}{N!} (1-t)^N \frac{\delta^{N+1} F}{\delta f^{N+1}} (f + th)[h, \ldots, h] dt \tag{3.130}$$

goes to zero for $N \to \infty$,thus if:

$$\lim_{N \to \infty} \|R_N(f, h)\|_2 = 0 \tag{3.131}$$

then we have within the convergence radius:

$$F[f + h] = F[f] + \sum_{n=1}^{\infty} \frac{1}{n!} \frac{\delta^n F}{\delta f^n} (f)[h, \ldots, h] \tag{3.132}$$

In the integral kernel notation we have:

$$F[f + h] = F[f] + \sum_{n=1}^{\infty} \frac{1}{n!} \int \frac{\delta^n F}{\delta f(x_1) \ldots \delta f(x_2)} h(x_1) \ldots h(x_n) dx_1 \ldots dx_n \tag{3.133}$$

We now have derived most formulas needed in practical application of functional calculus. This final chapter therefore ends our review of functional calculus. A more rigorous discussion on all the discussed topics can be found in [64, 65].

Chapter 4

Density functionals for Coulomb systems

4.1 Introduction

In this section we will discuss the several functional analytical properties of density functionals for electronic systems with Coulombic interparticle interactions. To derive continuity and differentiability of the density functionals we restrict ourselves to a certain set of external potentials, mathematically denoted as $L^\infty(\mathcal{R}^3) + L^{3/2}(\mathcal{R}^3)$ with contains the physically important Coulomb potential as well as any finite sum of Coulomb potentials relevant to molecular systems. The discussion in this section is largely based on the mathematical papers by Lieb [19] and Englisch and Englisch [67, 68]

4.2 Conditions on the electron density and external potentials

The electron density corresponding to a normalized N-electron wave function Ψ is defined as

$$\rho(\mathbf{r}) = N \sum_{\sigma_1...\sigma_N} \int |\Psi(\mathbf{r}\sigma_1, \mathbf{r}_2\sigma_2 \ldots \mathbf{r}_N\sigma_N)|^2 d\mathbf{r}_2 \ldots d\mathbf{r}_N \tag{4.1}$$

We first put some constraints on the wave function. First of all because of the probability interpretation one likes the wavefunction to be normalizable to one, so we require $\|\Psi\| < \infty$ where the norm is defined as

$$\|\Psi\| = \sum_{\sigma_1...\sigma_N} \int |\Psi(\mathbf{r}_1\sigma_1 \ldots \mathbf{r}_N\sigma_N)|^2 d\mathbf{r}_1 \ldots d\mathbf{r}_N \tag{4.2}$$

Secondly because of the superposition principle in quantum mechanics one requires that also some infinite linear combinations exist and are normalizable. More precisely we first define

$$\Psi_M = \sum_{i=1}^{M} a_m \Psi_m \tag{4.3}$$

with $\|\Psi_m\| = 1$ and

$$\sum_{i=1}^{M} |a_i|^2 = 1 \tag{4.4}$$

One can easily check this to be a Cauchy sequence

$$\lim_{M \to \infty} \|\Psi_{M+1} - \Psi_M\| \to 0 \quad (M \to \infty) \tag{4.5}$$

Then we would like the limit function to exist, i.e. we would like that our space of normalizable functions contains some limit functions Ψ such that

$$\lim_{M \to \infty} \|\Psi - \Psi_M\| = 0 \tag{4.6}$$

so we require the above Cauchy sequence to converge, more compactly we like our space of functions to be complete with respect to the square integral norm. It can be proven that the smallest space which is complete with respect to the above norm is the space L^2 of functions Ψ for which the Lebesque integral of $|\Psi|^2$ exists. For a definition of the Lebesque integral we refer to the previous chapter. A third requirement on wavefunctions is that their kinetic energy expectation value must be finite. The expectation values of the potential energy operators we will discuss later on. So we require that

$$T[\Psi] = \frac{1}{2} \sum_{i=1}^{N} \sum_{\sigma_1 \ldots \sigma_N} \int |\nabla_i \Psi(\mathbf{r}_1 \sigma_1 \ldots \mathbf{r}_N \sigma_N)|^2 d\mathbf{r}_1 \ldots d\mathbf{r}_N < \infty \tag{4.7}$$

Here $\nabla \Psi$ is defined almost everywhere (a precise definition of $\nabla \Psi$ for a L^2-function can be given using Fourier transforms). This implies that

$$\sum_{\sigma_1 \ldots \sigma_N} \int |\Psi(\mathbf{r}_1 \sigma_1 \ldots \mathbf{r}_N \sigma_N)|^2 + \sum_{i=1}^{N} |\nabla_i \Psi(\mathbf{r}_1 \sigma_1 \ldots \mathbf{r}_N \sigma_N)|^2 d\mathbf{r}_1 \ldots d\mathbf{r}_N < \infty \tag{4.8}$$

and so we now find that $\Psi \in H^1(\mathcal{R}^{3N})$. All these constraints on Ψ have consequences for the constraints on the electron density ρ. First of all because $\Psi \in L^2(\mathcal{R}^{3N})$ we find from equation 4.1 that $\rho \in L^1(\mathcal{R}^3)$ which means that the Lebesque integral of the electron density is finite

$$\int \rho(\mathbf{r}) d\mathbf{r} < \infty \tag{4.9}$$

We now will show that the finiteness of the kinetic energy implies that ρ must also be in $L^3(\mathcal{R}^3)$ which together with the previous results implies that $\rho \in L^1(\mathcal{R}^3) \cap L^3(\mathcal{R}^3)$.
If we consider a wave function $\Psi \in H^1(\mathcal{R}^{3N})$ which yields density ρ then

$$\nabla \rho(\mathbf{r}) = N \sum_{\sigma_1 \ldots \sigma_N} \int \Psi^*(\mathbf{r}\sigma_1 \ldots \mathbf{r}_N \sigma_N) \nabla \Psi(\mathbf{r}\sigma_1 \ldots \mathbf{r}_N \sigma_N) d\mathbf{r}_2 \ldots d\mathbf{r}_N + c.c. \tag{4.10}$$

Using the Schwarz inequality

$$|\int f(x_1 \ldots x_N) g^*(x_1 \ldots x_N) dx_1 \ldots dx_n|^2$$

$$\leq \int |f(x_1 \ldots x_N)|^2 dx_1 \ldots dx_N \int |g(x_1 \ldots x_N)|^2 dx_1 \ldots dx_N \tag{4.11}$$

we find

$$|\nabla \rho(\mathbf{r})|^2 \leq$$

$$4N^2 \sum_{\sigma_1 \ldots \sigma_N} \int |\Psi(\mathbf{r}\sigma_1 \ldots \mathbf{r}_N \sigma_N)|^2 d\mathbf{r}_2 \ldots d\mathbf{r}_N \sum_{\sigma_1 \ldots \sigma_N} \int |\nabla \Psi(\mathbf{r}\sigma_1 \ldots \mathbf{r}_N \sigma_N)|^2 d\mathbf{r}_2 \ldots d\mathbf{r}_N$$

$$= 8\rho(\mathbf{r})\frac{1}{2}N\sum_{\sigma_1\ldots\sigma_N}\int|\nabla\Psi(\mathbf{r}\sigma_1\ldots\mathbf{r}_N\sigma_N)|^2 d\mathbf{r}_2\ldots d\mathbf{r}_N \tag{4.12}$$

So we find that

$$T_W[\rho] = \frac{1}{2}\int|\nabla\sqrt{\rho(\mathbf{r})}|^2 d\mathbf{r} = \frac{1}{8}\int\frac{(\nabla\rho(\mathbf{r}))^2}{\rho(\mathbf{r})}d\mathbf{r} \leq T[\Psi] < \infty \tag{4.13}$$

Therefore $\sqrt{\rho(\mathbf{r})} \in H^1(\mathcal{R}^3)$. If we now use the Sobolev inequality for functions $f \in H^1(\mathcal{R}^3)$

$$\int|\nabla f(\mathbf{r})|^2 d\mathbf{r} \geq 3(\frac{\pi}{2})^{\frac{4}{3}}\left[\int|f(\mathbf{r})|^6 d\mathbf{r}\right]^{\frac{1}{3}} \tag{4.14}$$

we find that

$$\int\rho^3(\mathbf{r})d\mathbf{r} \leq \frac{1}{3}(\frac{2}{\pi})^{\frac{4}{3}}\int|\nabla\sqrt{\rho(\mathbf{r})}|^2 < \infty \tag{4.15}$$

so we conclude that $\rho \in L^3(\mathcal{R}^3)$ which together with $\rho \in L^1(\mathcal{R}^3)$ implies $\rho \in L^1(\mathcal{R}^3) \cap L^3(\mathcal{R}^3)$. Some definitions are useful for further discussions in this section. We define the sets

$$\mathcal{S}_N = \{\rho|\rho(\mathbf{r}) \geq 0, \sqrt{\rho} \in H^1(\mathcal{R}^3), \int\rho(\mathbf{r})d\mathbf{r} = N\} \tag{4.16}$$

and

$$\mathcal{R}_N = \{\rho|\rho(\mathbf{r}) \geq 0, \rho \in L^3(\mathcal{R}^3), \int\rho(\mathbf{r})d\mathbf{r} = N\} \tag{4.17}$$

The set $\mathcal{R}_N$ is convex, that is, if ρ_1 and $\rho_2 \in \mathcal{R}_N$ then $\rho = \lambda_1\rho_1 + \lambda_2\rho_2 \in \mathcal{R}_N$ with $0 \leq \lambda_1, \lambda_2 \leq 1$ and $\lambda_1 + \lambda_2 = 1$. This follows from

$$\int\rho(\mathbf{r})d\mathbf{r} = \lambda_1\int\rho_1(\mathbf{r})d\mathbf{r} + \lambda_2\int\rho_2(\mathbf{r})d\mathbf{r} = N \tag{4.18}$$

and

$$\|\lambda_1\rho_1 + \lambda_2\rho_2\|_3 \leq \lambda_1\|\rho_1\|_3 + \lambda_2\|\rho_2\|_3 < \infty \tag{4.19}$$

The set $\mathcal{S}_N$ is also convex. This follows from

$$\nabla\rho = \lambda_1\nabla\rho_1 + \lambda_2\nabla\rho_2 = 2\lambda_1\sqrt{\rho_1}\nabla\sqrt{\rho_1} + 2\lambda_2\sqrt{\rho_2}\nabla\sqrt{\rho_2} = \alpha_1\nabla\alpha_1 + \alpha_2\nabla\alpha_2 \tag{4.20}$$

with $\alpha_1 = \sqrt{2\lambda_1\rho_1}$ and $\alpha_2 = \sqrt{2\lambda_2\rho_2}$ and from the Schwarz inequality

$$(\alpha_1\nabla\alpha_1 + \alpha_2\nabla\alpha_2)^2 \leq (\alpha_1^2 + \alpha_2^2)((\nabla\alpha_1)^2 + (\nabla\alpha_2)^2) \tag{4.21}$$

This yields

$$(\nabla\rho)^2 \leq 4\rho(\lambda_1(\nabla\sqrt{\rho_1})^2 + \lambda_2(\nabla\sqrt{\rho_2})^2) \tag{4.22}$$

and we find

$$\int(\sqrt{\rho})^2 d\mathbf{r} \leq \lambda_1\int(\sqrt{\rho_1})^2 d\mathbf{r} + \lambda_2\int(\sqrt{\rho_2})^2 d\mathbf{r} < \infty \tag{4.23}$$

So $\rho \in \mathcal{S}_N$. In particular it follows that the von Weizsäcker kinetic energy functional 4.13 is a convex functional, i.e.

$$T_W[\lambda_1\rho_1 + \lambda_2\rho_2] \leq \lambda_1 T_W[\rho_1] + \lambda_2 T_W[\rho_2] \tag{4.24}$$

We will now prove the following statement. For any electron density $\rho \in \mathcal{S}_N$ there is a Slater determinant wavefunction $\Psi \in H^1(\mathcal{R}^{3N})$ which yields this density (i.e. with satisfies relation 4.1). The proof is by explicit construction. We write $\mathbf{r} = (r_1, r_2, r_3)$ and define

$$f(r_1) = \frac{2\pi}{N} \int_{-\infty}^{r_1} dv_1 \int_{-\infty}^{+\infty} dv_2 \int_{-\infty}^{+\infty} dv_3 \rho(v_1, v_2, v_3) \tag{4.25}$$

The function f is monotonously increasing with $f(-\infty) = 0$ and $f(+\infty) = 2\pi$. We define the orbitals ϕ_n by

$$\phi_n(\mathbf{r}) = \phi_n(r_1, r_2, r_3) = \sqrt{\frac{\rho(r_1, r_2, r_3)}{N}} \exp\left(i n f(r_1)\right) \tag{4.26}$$

for $n = 0, \ldots, N-1$. Then

$$|\phi_n(\mathbf{r})|^2 = \frac{\rho(\mathbf{r})}{N} \tag{4.27}$$

and for $n \neq m$ we have

$$\int \phi_m^*(\mathbf{r}) \phi_n(\mathbf{r}) d\mathbf{r} = \frac{1}{N} \int \rho(\mathbf{r}) e^{i(n-m)f(r_1)} d\mathbf{r} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{df}{dr_1} e^{i(n-m)f(r_1)} dr_1$$

$$= \frac{1}{2\pi i (n-m)} (e^{i(n-m)f(+\infty)} - e^{i(n-m)f(-\infty)}) = 0 \tag{4.28}$$

So the orbitals ϕ_n form an orthonormal set which sum to the prescribed density and therefore the corresponding Slater determinant wavefunction will yield the same density.

After having discussed the density and the expectation value of the kinetic energy operator we will now discuss the expectation value of the potential energy operators. Because $\Psi \in H^1(\mathcal{R}^{3N})$ it follows [19] that for

$$\hat{W} = \frac{1}{2} \sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{4.29}$$

we have

$$\langle \Psi | \hat{W} | \Psi \rangle < \infty \tag{4.30}$$

Also a lower bound for this quantity can be proven. If Ψ yields density ρ we have

$$\langle \Psi | \hat{W} | \Psi \rangle > \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - C \int \rho^{\frac{4}{3}}(\mathbf{r}) d\mathbf{r} \tag{4.31}$$

with $C = 1.68$ [69]. We now turn to the expectation values of the external potential $\hat{V}$. From the condition

$$\left| \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \right| < \infty \tag{4.32}$$

and $\rho \in L^1(\mathcal{R}^3) \cap L^3(\mathcal{R}^3)$ we can deduce some constraints on the external potential $v(\mathbf{r})$. If $\rho \in L^1(\mathcal{R}^3)$ then the above integral exists for bounded potentials,i.e. for potentials $v \in L^\infty(\mathcal{R}^3)$. This follows directly from

$$\left| \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \right| \leq \operatorname{ess\,sup} |v(\mathbf{r})| \int \rho(\mathbf{r}) d\mathbf{r} < \infty \tag{4.33}$$

if $v \in L^\infty(\mathcal{R}^3)$. If $\rho \in L^3(\mathcal{R}^3)$ then the above integral exists for potentials in the set $L^{\frac{3}{2}}(\mathcal{R}^3)$. This follows from the Hölder inequality

$$\|fg\|_1 \le \|f\|_p \|g\|_q \tag{4.34}$$

with $1/p + 1/q = 1$. If $v \in L^{\frac{3}{2}}(\mathcal{R}^3)$ then we find using this inequality

$$\left| \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \right| = \|\rho v\|_1 \le \|\rho\|_3 \|v\|_{\frac{3}{2}} < \infty \tag{4.35}$$

The most general set of potentials for which the expectation value $\langle \Psi | \hat{V} | \Psi \rangle$ exists is therefore the set

$$L^{\frac{3}{2}}(\mathcal{R}^3) + L^\infty(\mathcal{R}^3) = \{v | v = u + w, u \in L^{\frac{3}{2}}(\mathcal{R}^3), w \in L^\infty(\mathcal{R}^3)\} \tag{4.36}$$

This set is a Banach space with norm

$$\|v\| = \inf\{\|u\|_{\frac{3}{2}} + \|w\|_\infty | v = u + w\} \tag{4.37}$$

An important potential in this Banach space is Coulombic potential as it can be written as

$$\frac{1}{r} = u(r) + w(r) \tag{4.38}$$

where $r = |\mathbf{r}|$ and

$$u(r) = \frac{\theta(1-r)}{r} \qquad w(r) = \frac{\theta(r-1)}{r} \tag{4.39}$$

where θ is the Heaviside function, $\theta(x) = 0$ if $x \le 0$ and $\theta(x) = 1$ if $x > 0$. One can readily check that $u \in L^{\frac{3}{2}}(\mathcal{R}^3)$ and $w \in L^\infty(\mathcal{R}^3)$. The Banach space $L^{\frac{3}{2}}(\mathcal{R}^3) + L^\infty(\mathcal{R}^3)$ does not contain external potentials that go to infinity as $|\mathbf{r}| \to \infty$. This choice precludes some physically interesting potentials such as the harmonic oscillator potential. These potentials can be handled with the methods to be disussed but then one has to put additional restrictions on ρ such that the integral of ρv makes sense.

4.3 Properties of the energy functional $E[v]$

For the external potentials in the set $L^{\frac{3}{2}}(\mathcal{R}^3) + L^\infty(\mathcal{R}^3)$ we define the total energy functional $E[v]$ as

$$E[v] = \inf \langle \Psi | \hat{H} | \Psi \rangle \tag{4.40}$$

where $\Psi \in H^1(\mathcal{R}^{3N})$ and where the wave function Ψ is normalized to one $\|\Psi\| = 1$. We will prove some properties of this functional. First of all we have that $E[v]$ is concave, that is

$$E[\lambda_1 v_1 + \lambda_2 v_2] \ge \lambda_1 E[v_1] + \lambda_2 E[v_2] \tag{4.41}$$

for all $v \in L^{\frac{3}{2}}(\mathcal{R}^3) + L^\infty(\mathcal{R}^3)$ and $0 \le \lambda_1, \lambda_2 \le 1$ and $\lambda_1 + \lambda_2 = 1$. This follows directly from the variational principle. If Ψ is a wave function corresponding to the infimum in equation 4.40 then

$$\begin{aligned} E[v] &= \langle \Psi | \hat{T} + \hat{V} + \hat{W} | \Psi \rangle = \lambda_1 \langle \Psi | \hat{T} + \hat{V}_1 + \hat{W} | \Psi \rangle + \lambda_2 \langle \Psi | \hat{T} + \hat{V}_2 + \hat{W} | \Psi \rangle \\ &\ge \lambda_1 E[v_1] + \lambda_2 E[v_2] \end{aligned} \tag{4.42}$$

where we used $\lambda_1 + \lambda_2 = 1$. A simple application of the above relation is the following. Consider a molecular Hamiltonian with N atoms of positive nuclear charge Z_i at positions $\mathbf{R}_i$.

$$v(\mathbf{r}) = -\sum_{i=1}^{N} \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} \tag{4.43}$$

if we define the total nuclear charge

$$Z = \sum_{i=1}^{N} Z_i \tag{4.44}$$

and define $\lambda_i = Z_i/Z$ then we find using the concavity property of $E[v]$ that

$$
\begin{aligned}
E[v] &= E[\sum_{i=1}^{N} -\frac{\lambda_i Z}{|\mathbf{r} - \mathbf{R}_i|}] \geq \sum_{i=1}^{N} \lambda_i E[-\frac{Z}{|\mathbf{r} - \mathbf{R}_i|}] \\
&= \sum_{i=1}^{N} \lambda_i E[-\frac{Z}{|\mathbf{r}|}] = E[-\frac{Z}{|\mathbf{r}|}]
\end{aligned}
\tag{4.45}
$$

So we conclude that the total electronic energy of a molecule with total nuclear charge Z is always larger than or equal to the total energy of an atom with the same nuclear charge Z and the same number of electrons. For instance the electronic energy of the hydrogen molecule is larger than the electronic energy of the helium atom.

A second property of $E[v]$ is that it is monotonously decreasing, that is, if $v_1(\mathbf{r}) \leq v_2(\mathbf{r})$ for all $\mathbf{r}$ (almost everywhere) then $E[v_1] \leq E[v_2]$. This follows again from the variational property. If Ψ is a wave function corresponding to the infimum in equation 4.40 then

$$
\begin{aligned}
E[v_2] &= \langle \Psi | \hat{T} + \hat{V}_2 + \hat{W} | \Psi \rangle = \langle \Psi | \hat{T} + \hat{V}_1 + \hat{W} | \Psi \rangle + \langle \Psi | \hat{V}_2 - \hat{V}_1 | \Psi \rangle \\
&\geq E[v_1] + \int \rho(\mathbf{r})(v_2(\mathbf{r}) - v_1(\mathbf{r})) d\mathbf{r} \geq E[v_1]
\end{aligned}
\tag{4.46}
$$

We further state without proof a third property of $E[v]$, which is : $E[v]$ is continuous in the $L^{\frac{3}{2}} + L^{\infty}$-norm. So if we have a sequence of external potentials $\{v_i\}$ in the Banach space $L^{\frac{3}{2}} + L^{\infty}$ with norm $\|\|$ converging to $v \in L^{\frac{3}{2}} + L^{\infty}$ then

$$\lim_{i \to \infty} \|v_i - v\| = 0 \Rightarrow \lim_{i \to \infty} |E[v_i] - E[v]| = 0 \tag{4.47}$$

The proof of this can be found in the paper by Lieb [19].

4.4 The Hohenberg-Kohn functionals $F_{HK}[\rho]$ and $F_{EHK}[\rho]$

We will now discuss the Hohenberg-Kohn functional which has been introduced by Hohenberg and Kohn in their well-known paper. First we will prove the Hohenberg-Kohn theorem. Suppose that Ψ_1 and $\Psi_2 \in H^1(\mathcal{R}^{3N})$ are non-degenerate ground state wave functions corresponding to external potentials v_1 and $v_2 \in L^{\infty} + L^{\frac{3}{2}}$ with corresponding electron densities ρ_1 and ρ_2. Then if $v_1 \neq v_2 + C$ where C is constant then $\rho_1 \neq \rho_2$.

As a first step we have that if $v_1 \neq v_2 + C$ that $\Psi_1 \neq \Psi_2$. This follows by contradiction. Suppose $\Psi_1 = \Psi_2 = \Psi$ then by subtraction of the Hamiltonian for Ψ_1 and Ψ_2 we find that

$$(v_1 - v_2)\Psi = (E_1 - E_2)\Psi \tag{4.48}$$

If $v_1 - v_2$ is not constant in some region then Ψ must vanish in this region for the above equation to be true. However if $v_1, v_2 \in L^\infty + L^{\frac{3}{2}}$ then Ψ cannot vanish on an open set (a set with nonzero measure) by the unique continuation theorem. So we obtain a contradiction and we find $\Psi_1 \neq \Psi_2$. So different potentials (differing more than a constant) give different wavefunctions. In the following we will denote $\Psi_1 = \Psi[v_1]$ and $\Psi_2 = \Psi[v_2]$.

These different wave functions also yield different densities. This follows again by contradiction. If $\rho_1 = \rho_2 = \rho$ then

$$E[v_1] = \langle \Psi[v_1]|\hat{T} + \hat{V}_1 + \hat{W}|\Psi[v_1]\rangle < \langle \Psi[v_2]|\hat{T} + \hat{V}_1 + \hat{W}|\Psi[v_2]\rangle$$

$$= \langle \Psi[v_2]|\hat{T} + \hat{V}_2 + \hat{W}|\Psi[v_2]\rangle + \int \rho(\mathbf{r})(v_1(\mathbf{r}) - v_2(\mathbf{r}))d\mathbf{r}$$

$$= E[v_2] + \int \rho(\mathbf{r})(v_1(\mathbf{r}) - v_2(\mathbf{r}))d\mathbf{r} \tag{4.49}$$

Likewise

$$E[v_2] < E[v_1] + \int \rho(\mathbf{r})(v_2(\mathbf{r}) - v_1(\mathbf{r}))d\mathbf{r} \tag{4.50}$$

Adding the both inequalities yields the contradiction

$$E[v_1] + E[v_2] < E[v_1] + E[v_2] \tag{4.51}$$

So we conclude that $\rho_1 \neq \rho_2$, which proves our statement.

We now define the set $\mathcal{A}_N$ as

$$\mathcal{A}_N = \{\rho|\rho \text{ comes from a non-degenerate ground state}\} \tag{4.52}$$

where we only consider ground state densities from potentials in the set $L^\infty + L^{\frac{3}{2}}$. The set $\mathcal{A}_N$ is a subset of the previously defined set $\mathcal{S}_N$. The densities in the set $\mathcal{A}_N$ we will call pure state v-representable densities, shortly denoted as PS-V-densities. From the Hohenberg-Kohn theorem there is a unique external potential v (to within a constant) and a unique ground state wave function $\Psi[\rho]$ (to within a phase factor) which yields this density. On the set of PS-V densities we can therefore define the Hohenberg-Kohn functional F_{HK} as

$$F_{HK}[\rho] = \langle \Psi[\rho]|\hat{T} + \hat{W}|\Psi[\rho]\rangle \tag{4.53}$$

we can then define the energy functional E_v as

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \tag{4.54}$$

If ρ_0 is a ground state density corresponding to external potential v_0 and ρ an arbitrary other ground state density then

$$E_{v_0}[\rho] = \int \rho(\mathbf{r})v_0(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] = \langle \Psi[\rho]|\hat{T} + \hat{V}_0 + \hat{W}|\Psi[\rho]\rangle$$

$$\geq \langle \Psi[\rho_0]|\hat{T} + \hat{V}_0 + \hat{W}|\Psi[\rho_0]\rangle = E_{v_0}[\rho_0] \tag{4.55}$$

Therefore

$$E[v] = \inf_{\rho \in \mathcal{A}_N} \left\{ \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] \right\} \tag{4.56}$$

For application of the above formula we have to known F_{HK} on the set $\mathcal{A}_N$. As the set $\mathcal{A}_N$ is difficult to determine we need an extension of F_{HK} to a larger better known set. We will however first discuss the above functional. The functional F_{HK} has some mathematical inconveniences. First of all its domain $\mathcal{A}_N$ is not convex, that is, if $\rho_1, \rho_2 \in \mathcal{A}_N$ then not necessarily $\lambda_1\rho_1 + \lambda_2\rho_2 \in \mathcal{A}_N$ with $0 \le \lambda_1, \lambda_2 \le 1$ and $\lambda_1 + \lambda_2 = 1$. We will prove this in a later instant.
The functional F_{HK} is however convex, that is, if $\rho_1, \rho_2 \in \mathcal{A}_N$ and if $\lambda_1\rho_1 + \lambda_2\rho_2 \in \mathcal{A}_N$ with $0 \le \lambda_1, \lambda_2 \le 1$ then

$$F_{HK}[\lambda_1\rho_1 + \lambda_2\rho_2] \le \lambda_1 F_{HK}[\rho_1] + \lambda_2 F_{HK}[\rho_2] \tag{4.57}$$

This is readily proved. Suppose that the ground state densities $\rho_1, \rho_2, \lambda_1\rho_1 + \lambda_2\rho_2 \in \mathcal{A}_N$ correspond to the external potentials v_1, v_2 and v. Then

$$\int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \langle \Psi[\rho]|\hat{T} + \hat{W}|\Psi[\rho]\rangle =$$

$$= \langle \Psi[\rho]|\hat{H}|\Psi[\rho]\rangle = \lambda_1\langle \Psi[\rho]|\hat{H}|\Psi[\rho]\rangle + \lambda_2\langle \Psi[\rho]|\hat{H}|\Psi[\rho]\rangle$$

$$\le \lambda_1\langle \Psi[\rho_1]|\hat{T} + \hat{W}|\Psi[\rho_1]\rangle + \lambda_2\langle \Psi[\rho_2]|\hat{T} + \hat{W}|\Psi[\rho_2]\rangle + \int (\lambda_1\rho_1(\mathbf{r}) + \lambda_2\rho_2(\mathbf{r}))v(\mathbf{r})d\mathbf{r}$$

$$= \lambda_1 F_{HK}[\rho_1] + \lambda_2 F_{HK}[\rho_2] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \tag{4.58}$$

and we obtain the convexity of F_{HK}. The functional F_{HK} is defined on the set $\mathcal{A}_N$. This is not a convex set. There are convex combinations of PS-V ground state densities which are not in $\mathcal{A}_N$. An example of this is a convex combination of densities corresponding to a q-fold degenerate ground state multiplet

$$\rho = \sum_{i=1}^{q} \lambda_i\rho_i \qquad \sum_{i=1}^{q} \lambda_i = 1 \qquad 0 \le \lambda_i \le 1 \tag{4.59}$$

which in general is not in $\mathcal{A}_N$. We will demonstrate this at the end of this section. We can however readily extend our functional F_{HK} to this type of densities. We define

$$\mathcal{B}_N = \{\rho = \sum_i \lambda_i\rho_i | \rho_i \text{ comes from the same } v \text{ for all } i, \ \sum_i \lambda_i = 1, \ 0 \le \lambda_i \le 1\} \tag{4.60}$$

We call these densities ensemble v-representable densities or E-V-densities. We will now extend the functional F_{HK} to the set $\mathcal{B}_N$ of ensemble v-representable densities. The corresponding ensemble functional we will denote by F_{EHK}. For a degenerate ground state multiplet $\{\psi_i\}$ with q-fold degeneracy corresponding to some external potential v we define the ensemble density matrices

$$\hat{D} = \sum_{i=1}^{q} \lambda_i|\psi_i\rangle\langle\psi_i| \qquad \sum_{i=1}^{q} \lambda_i = 1, \qquad 0 \le \lambda_i \le 1 \tag{4.61}$$

the corresponding ensemble density is given by

$$\rho(\mathbf{r}) = \text{Tr}\hat{D}\hat{\rho}(\mathbf{r}) \tag{4.62}$$

where $\hat{\rho}(\mathbf{r})$ is the density operator. We now prove the following extension of the Hohenberg-Kohn theorem to arbitrary degenerate ground states. Suppose $\hat{D}_1$ and $\hat{D}_2$ are ground state ensemble density matrices belonging to v_1 and v_2 resp., with corresponding densities ρ_1 and ρ_2. If $v_1 \ne v_2 + C$

with C a constant then $\rho_1 \neq \rho_2$.

The proof is analogous to the proof of the non-degenerate case. First the sets of ground state ensemble density matrices corresponding to the two different potentials v_1 and v_2 are disjoint. Suppose v_1 generates the ground state multiplet $A_1 = \{|\phi_i\rangle, \ i = 1 \ldots q_1\}$ and v_2 generates the ground state multiplet $A_2 = \{|\psi_i\rangle, i = 1 \ldots q_2\}$. All the wave functions within these multiplets may without loss of generality be chosen orthonormal. Then none of the wave functions in the sets A_1 and A_2 are equal. This follows from the same argument as used in the proof of the Hohenberg-Kohn theorem for the non-degenerate case. In particular, as the sets A_1 and A_2 are only defined to within a unitary transformation no $|\psi_i\rangle$ in A_2 is a linear combination of the $|\phi_i\rangle$ in A_1. This then implies that two ground state ensemble density matrices constructed from the ground states in A_1 and A_2 are different

$$\hat{D}_1 = \sum_{i=1}^{q_1} \lambda_i |\phi_i\rangle\langle\phi_i| \neq \sum_{i=1}^{q_2} \mu_i |\psi_i\rangle\langle\psi_i| = \hat{D}_2 \tag{4.63}$$

where $\sum \lambda_i = \sum \mu_i = 1$. This follows for instance by taking the inproduct on both sides with $|\psi_m\rangle$ as the $|\psi_i\rangle$ are not linear combinations of the $|\phi_i\rangle$.

Secondly, if $\hat{H}_1 = \hat{T} + \hat{V}_1 + \hat{W}$ and $\hat{H}_2 = \hat{T} + \hat{V}_2 + \hat{W}$ then

$$\text{Tr}\,\hat{D}_1 \hat{H}_2 > \text{Tr}\,\hat{D}_2 \hat{H}_2 \tag{4.64}$$

This follows directly from

$$\text{Tr}\,\hat{D}_1 \hat{H}_2 = \sum_{i=1}^{q_1} \lambda_i \langle\phi_i|\hat{H}_2|\phi_i\rangle > \sum_{i=1}^{q_1} \lambda_i \langle\psi_i|\hat{H}_2|\psi_i\rangle =$$

$$= \sum_{i=1}^{q_1} \lambda_i E[v_2] = E[v_2] = \sum_{i=1}^{q_2} \mu_2 \langle\psi_i|\hat{H}_2|\psi_i\rangle = \text{Tr}\,\hat{D}_2 \hat{H}_2 \tag{4.65}$$

Now we can show that $\hat{D}_1$ and $\hat{D}_2$ yield different densities. We proceed again by reductio ad absurdum. Suppose $\rho_1 = \rho_2 = \rho$. Using the last result we find

$$E[v_1] = \text{Tr}\,\hat{D}_1 \hat{H}_1 = \text{Tr}\,\hat{D}_1(\hat{H}_2 + \hat{V}_1 - \hat{V}_2) =$$

$$= \text{Tr}\,\hat{D}_1 \hat{H}_2 + \int \rho(\mathbf{r})(v_1(\mathbf{r}) - v_2(\mathbf{r}))d\mathbf{r} > \text{Tr}\,\hat{D}_2 \hat{H}_2 + \int \rho(\mathbf{r})(v_1(\mathbf{r}) - v_2(\mathbf{r}))d\mathbf{r} =$$

$$= E[v_2] + \int \rho(\mathbf{r})(v_1(\mathbf{r}) - v_2(\mathbf{r}))d\mathbf{r} \tag{4.66}$$

Likewise we have

$$E[v_2] > E[v_1] + \int \rho(\mathbf{r})(v_2(\mathbf{r}) - v_1(\mathbf{r}))d\mathbf{r} \tag{4.67}$$

which added to the last inequality leads to the contradiction

$$E[v_1] + E[v_2] > E[v_1] + E[v_2] \tag{4.68}$$

Therefore $\hat{D}_1$ and $\hat{D}_2$ must give different densities, which proves the theorem.

Within the set of ensemble ground state density matrices corresponding to the *same* potential however, two different density matrices can yield the same density. The energy $\text{Tr}\,\hat{D}\hat{H}$ for those different

density matrices is however the same. For every E-V-density ρ we can therefore unambiguously define

$$F_{EHK}[\rho] = \text{Tr}\hat{D}[\rho](\hat{T} + \hat{W}) \tag{4.69}$$

where $\hat{D}[\rho]$ is any of the ground state ensemble density matrices corresponding to ρ. We can now define an extension of the energy functional E_v to the set of E-V-densities

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{EHK}[\rho] = \text{Tr}\hat{D}[\rho]\hat{H} \tag{4.70}$$

Similarly as for F_{HK} we easily can prove

$$E[v] = \inf_{\rho \in \mathcal{B}_N} \left\{ \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{EHK}[\rho] \right\} \tag{4.71}$$

The functional F_{EHK} is an extension of F_{HK} as we have

$$F_{EHK}[\rho] = F_{HK}[\rho] \ \text{ if } \ \rho \in \mathcal{A}_N \tag{4.72}$$

This follows directly from the fact that for a non-degenerate ground state $|\Psi[\rho]\rangle$ corresponding to ρ we have $\hat{D}[\rho] = |\Psi[\rho]\rangle\langle\Psi[\rho]|$,so

$$F_{EHK}[\rho] = \text{Tr}\hat{D}[\rho](\hat{T} + \hat{W}) = \langle\Psi[\rho]|\hat{T} + \hat{W}|\Psi[\rho]\rangle = F_{HK}[\rho] \tag{4.73}$$

We can furthermore prove that F_{EHK} is convex by the same proof as for F_{HK}. Nothing is however known on the the convexity of the set of E-V-densities $\mathcal{B}_N$ which constitute the domain of F_{EHK}. As we will now demonstrate the subset of PS-V-densities $\mathcal{A}_N$ of $\mathcal{B}_N$ is not convex. More precisely we will now show that there are E-V-densities which are not PS-V-densities. As any E-V-density is a convex combination of PS-V densities this then demonstrates the non-convexity of $\mathcal{A}_N$.

Consider an atom with total angular momentum quantum number $L > 0$ which has a $2L + 1$-degenerate ground state. The external potential v is the Coulomb potential. The degeneracy is due to the fact that the Hamiltonian of the system is invariant with respect to rotations. The ground state wave-functions then transform among one another according to a $2L + 1$-dimensional unitary representation of the rotation group. We assume that there is no accidental degeneracy. If we denote the ground state wave functions by $\{|\psi[\rho_i]\rangle = |\psi_i\rangle, i = 1 \ldots 2L + 1\}$ and the corresponding electron densities by ρ_i then the following convex combination

$$\bar{\rho} = \frac{1}{2L + 1} \sum_{i=1}^{2L+1} \rho_i \tag{4.74}$$

is invariant under all rotations and therefore spherically symmetric. However the ρ_i are not spherical. In fact not any of the densities corresponding to linear combinations of the ground states $|\psi_i\rangle$ is spherically symmetric. As the ρ_j is obtained from $|\psi_j\rangle$ which by a unitary transformation can be obtained from any other $|\psi_i\rangle$ and the external potential is invariant under rotations we find that

$$\int \rho_i(\mathbf{r})v(\mathbf{r})d\mathbf{r} = \int \rho_j(\mathbf{r})v(\mathbf{r})d\mathbf{r} = \int \bar{\rho}(\mathbf{r})v(\mathbf{r})d\mathbf{r} \tag{4.75}$$

for all $0 \leq i, j \leq 2L + 1$. Let us now suppose that $\bar{\rho}$ is generated from a ground state wave function $\psi[\bar{\rho}]$, then this wave function is not a linear combination of the $|\psi_i\rangle$ otherwise $\bar{\rho}$ would not be spherically symmetric. Then we find

$$\int \bar{\rho}(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{HK}[\bar{\rho}] = \int \bar{\rho}(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \langle\psi[\bar{\rho}]|\hat{T} + \hat{W}|\psi[\bar{\rho}]\rangle = \langle\psi[\bar{\rho}]|\hat{H}|\psi[\bar{\rho}]\rangle$$

$$> \sum_{i=1}^{2L+1} \frac{1}{2L+1} \langle \psi[\rho_i] | \hat{H} | \Psi[\rho_i] \rangle = \sum_{i=1}^{2L+1} \frac{1}{2L+1} \langle \psi[\rho_i] | \hat{T} + \hat{W} | \Psi[\rho_i] \rangle + \sum_{i=1}^{2L+1} \frac{1}{2L+1} \int \rho_i(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} =$$

$$= \sum_{i=1}^{2L+1} \frac{1}{2L+1} F_{HK}[\rho_i] + \int \bar{\rho}(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \tag{4.76}$$

This then gives

$$F_{HK}[\bar{\rho}] > \sum_{i=1}^{2L+1} \frac{1}{2L+1} F_{HK}[\rho_i] \tag{4.77}$$

But we already knew that F_{HK} was convex on the set of ground state densities which leads to a contradiction and therefore we must conclude that $\bar{\rho}$ is not a ground state density of any potential. The density $\bar{\rho}$ is however a convex combination of ground state densities corresponding to the same external potential and therefore by definition an ensemble v-representable density. We therefore have constructed a E-V-density which is not a PS-V-density. Therefore $\mathcal{A}_N$ is a real subset of $\mathcal{B}_N$ and moreover $\mathcal{A}_N$ is not convex.

4.5 The Levy and Lieb functionals $F_{LL}[\rho]$ and $F_L[\rho]$

The functionals F_{HK} and F_{EHK} have the unfortunate mathematical difficulty that their domains of definition $\mathcal{A}_N$ and $\mathcal{B}_N$,although they are well-defined, are difficult to characterize, i.e. it is difficult to know if a given density ρ belongs to $\mathcal{A}_N$ or $\mathcal{B}_N$. Although there are reasons to assume that all reasonably well-behaved (that is, twice differentiable, bounded and positive) densities belong to $\mathcal{B}_N$ this remains until now an unproven statement. It is therefore desirable to extend the domains of definition of F_{HK} and F_{EHK} to an easily characterizable (preferably convex) set of densities. This can be achieved using the constrained search procedure introduced by Levy. We define the Levy-Lieb functional F_{LL} as

$$F_{LL}[\rho] = \inf_{\psi \to \rho} \langle \psi | \hat{T} + \hat{W} | \psi \rangle \tag{4.78}$$

where the infimum is searched over all normalized anti-symmetric N-particle wave functions in $H^1(\mathcal{R}^{3N})$ yielding density ρ. As shown earlier such a density is always in the convex set $\mathcal{S}_N$ which is again a subspace of $L^1 \cap L^3$. One can furthermore show, as has been done by Lieb [19], that the infimum is always a minimum, i.e. there is always a minimizing wave function.

Let us discuss some properties of F_{LL}. The functional F_{LL} is an extension of the Hohenberg-Kohn functional F_{HK} which was defined on $\mathcal{A}_N$ to the larger set $\mathcal{S}_N$, i.e

$$F_{LL}[\rho] = F_{HK}[\rho] \quad \text{if} \quad \rho \in \mathcal{A}_N \tag{4.79}$$

This is readily derived. Suppose ρ is some ground state density corresponding to some external potential v and ground state wave function $\Psi[\rho]$ then

$$\int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + F_{HK}[\rho] = \langle \Psi[\rho] | \hat{H} | \Psi[\rho] \rangle = \inf_{\psi \to \rho} \langle \psi | \hat{H} | \psi \rangle =$$

$$= \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \inf_{\psi \to \rho} \langle \psi | \hat{T} + \hat{W} | \psi \rangle = \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + F_{LL}[\rho] \tag{4.80}$$

We define a corresponding energy functional

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{LL}[\rho] \tag{4.81}$$

If ρ_0 is the ground state density corresponding to v with corresponding ground state wave function $\psi[\rho_0]$ then

$$E_v[\rho] = \inf_{\psi \to \rho} \langle \psi|\hat{H}|\psi \rangle \geq \langle \psi[\rho_0]|\hat{H}|\psi[\rho_0]\rangle = E_v[\rho_0] \tag{4.82}$$

Minimizing E_v over the set $\mathcal{S}_N$ therefore yields the ground state density ρ_0 corresponding to external potential v. The functional F_{LL} has however one inconvenient property, it is not convex. We take the example of the previous section where we presented a density $\bar{\rho}$ which did not correspond to a ground state wavefunction. It was however a convex combination of degenerate ground state densities ρ_i with corresponding ground states $|\psi[\rho_i]\rangle$ corresponding to an external potential v. Then we find

$$\int \bar{\rho}(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{LL}[\bar{\rho}] = \inf_{\psi \to \bar{\rho}} \langle \psi|\hat{H}|\psi \rangle$$

$$> \frac{1}{2L+1} \sum_{i=1}^{2L+1} \langle \psi[\rho_i]|\hat{H}|\psi[\rho_i]\rangle = \frac{1}{2L+1} \sum_{i=1}^{2L+1} F_{LL}[\rho_i] + \int \bar{\rho}(\mathbf{r})v(\mathbf{r})d\mathbf{r} \tag{4.83}$$

and we find

$$F_{LL}[\bar{\rho}] > \frac{1}{2L+1} \sum_{i=1}^{N} F_{LL}[\rho_i] \tag{4.84}$$

which proves the non-convexity of F_{LL}. This is somewhat unfortunate as convexity is an important property which can be used to derive differentiability of functionals. We will therefore now define a different but related functional with the same domain $\mathcal{S}_N$ which is also convex. This is the Lieb functional F_L defined as

$$F_L[\rho] = \inf_{\hat{D} \to \rho} \text{Tr} \hat{D}(\hat{T} + \hat{W}) \tag{4.85}$$

where the infimum is searched over all N-particle density matrices

$$\hat{D} = \sum_{i=1} \lambda_i |\psi\rangle\langle\psi| \quad \sum_{i=1} \lambda_i = 1 \quad \psi \in H^1(\mathcal{R}^{3N}) \tag{4.86}$$

which yield the given density $\rho(\mathbf{r}) = \text{Tr} \hat{D}\hat{\rho}(\mathbf{r})$. One can also for this case prove the infimum to be a minimum, i.e. there is a minimizing density matrix. This functional is an extension of F_{EHK} to the larger set $\mathcal{S}_N$, that is

$$F_L[\rho] = F_{EHK}[\rho] \quad \text{if} \quad \rho \in \mathcal{B}_N \tag{4.87}$$

This follows directly from the fact that if $\rho \in \mathcal{B}_N$ then there is a potential v which generates a ground state ensemble density matrix $\hat{D}[\rho]$ which yields ρ. So

$$\int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{EHK}[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \text{Tr}\hat{D}[\rho](\hat{T} + \hat{W}) = \text{Tr}\hat{D}[\rho]\hat{H}$$

$$= \inf_{\hat{D} \to \rho} \text{Tr}\hat{D}\hat{H} = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \inf_{\hat{D} \to \rho} \text{Tr}\hat{D}(\hat{T} + \hat{W}) = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\rho] \tag{4.88}$$

We can again define an energy functional

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\rho] \tag{4.89}$$

which by a similar proof as for F_{LL} assumes its minimum at the ground state density corresponding to potential v. We further have the following relations

$$F_L[\rho] = F_{LL}[\rho] \quad \text{if} \quad \rho \in \mathcal{A}_N \tag{4.90}$$

and

$$F_L[\rho] < F_{LL}[\rho] \quad \text{if} \quad \rho \in \mathcal{B}_N \quad \text{and} \quad \rho \notin \mathcal{A}_N \tag{4.91}$$

The first relation follows from the fact that is the density ρ is a pure state v-representable density then the minimizing density matrix for F_L is a pure state density matrix. The second relation follows from the fact that if ρ is an ensemble v-representable density there is a ground state ensemble density matrix $\hat{D}[\rho]$ for which we have

$$\int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + \text{Tr}\hat{D}[\rho](\hat{T} + \hat{W}) = \langle \psi_i | \hat{H} | \psi_i \rangle \tag{4.92}$$

where $|\psi_i\langle$ is any of the ground states in the degenerate ground state multiplet. Any wave function yielding density ρ can not be a linear combination of these ground state wave-functions otherwise ρ would be pure state v-representable. Therefore its expectation value with the Hamiltonian must be larger, i.e

$$\langle \psi_i | \hat{H} | \psi_i \rangle < \inf_{\psi \to \rho} \langle \psi | \hat{H} | \psi \rangle = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_{LL}[\rho] \tag{4.93}$$

which proves our statement.
We will now demonstrate another important property of F_L , which is its convexity. If $\rho = \lambda_1\rho_1 + \lambda_2\rho_2$ then we have

$$\lambda_1 F_L[\rho_1] + \lambda_2 F_L[\rho_2] = \lambda_1 \inf_{\hat{D}_1 \to \rho_1} \text{Tr}\hat{D}_1(\hat{T} + \hat{W}) + \lambda_2 \inf_{\hat{D}_2 \to \rho_2} \text{Tr}\hat{D}_2(\hat{T} + \hat{W}) =$$

$$= \inf_{\hat{D}_1,\hat{D}_2 \to \rho_1,\rho_2} \text{Tr}(\lambda_1\hat{D}_1 + \lambda_2\hat{D}_2)(\hat{T} + \hat{W}) \geq \inf_{\hat{D} \to \rho} \text{Tr}\hat{D}(\hat{T} + \hat{W}) = F_L[\rho] \tag{4.94}$$

We therefore now have established that F_L is a convex functional on a convex space. This is important information which enables one to derive the Gâteaux differentiability of the functional F_L at the set $\mathcal{B}_N$ of ensemble v-representable densities. We will discuss this feature of F_L in the next section.

4.6   Gâteaux differentiability of F_L

We now turn to the question of the differentiability of the functional F_L. A convenient property of F_L which it shares with the functional F_{LL} is that it is defined on a convex set $\mathcal{S}_N$. This means that whenever $\rho, \rho_0 \in \mathcal{S}_N$ that also $\rho_0 + t(\rho - \rho_0) \in \mathcal{S}_N$ when $0 \leq t \leq 1$. In view of the definition of the Gâteaux derivative

$$\lim_{t \to 0} \frac{F_L[\rho_0 + t(\rho - \rho_0)] - F_L[\rho_0]}{t} = \frac{\delta F_L}{\delta \rho}(\rho_0)[\rho - \rho_0] \tag{4.95}$$

this is a convenient property. If the functional derivative $\delta F_L/\delta\rho(\rho_0)$ exists then it is a linear functional on the space $L^1 \cap L^3$. These linear functionals can be identified with the Banach space of potentials $L^{\frac{3}{2}} + L^\infty$. The potential is then regarded as the bounded linear functional with assigns to each density the value of the integral of ρv. More mathematically formulated, the dual space of $L^1 \cap L^3$ (which is the space of bounded linear functionals on $L^1 \cap L^3$) is the space $L^{\frac{3}{2}} + L^\infty$ [19]. If we use the convexity of F_L we find for $0 \leq t \leq 1$ that

$$F_L[\rho_0 + t(\rho - \rho_0)] = F_L[t\rho + (1-t)\rho_0] \leq tF_L[\rho] + (1-t)F_L[\rho_0] \tag{4.96}$$

and it follows that

$$\frac{F_L[\rho_0 + t(\rho - \rho_0)] - F_L[\rho_0]}{t} \leq F_L[\rho] - F_L[\rho_0] \tag{4.97}$$

If we assume Gâteaux differentiability at density ρ_0 then we find

$$F_L[\rho] - F_L[\rho_0] \geq \lim_{t \downarrow 0} \frac{F_L[\rho_0 + t(\rho - \rho_0)] - F_L[\rho_0]}{t} = \frac{\delta F_L}{\delta\rho}(\rho_0)[\rho - \rho_0] \tag{4.98}$$

Conversely, from the results of section 3.5.2 , we know because of the convexity of F_L that if there is a unique linear continuous functional (which is called a tangent functional or a subgradient) $L : L^1 \cap L^3 \mapsto \mathcal{R}$ such that

$$F_L[\rho] - F_L[\rho_0] \geq L[\rho - \rho_0] \tag{4.99}$$

is satisfied for all $\rho \in \mathcal{S}_N$ then F_L is Gâteaux differentiable at ρ_0 with

$$\frac{\delta F_L}{\delta\rho}(\rho_0) = L \tag{4.100}$$

To prove the differentiability of F_L at ρ_0 it is therefore sufficient to prove the existance of a unique continuous tangent functional at ρ_0. We will first prove the following statement. The functional F_L has a continuous tangent functional $-v \in L^{\frac{3}{2}} + L^\infty$ at $\rho_0 \in \mathcal{S}_N$ if there is a $v \in L^{\frac{3}{2}} + L^\infty$ such that

$$E[v] = \int \rho_0(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\rho_0] \tag{4.101}$$

is satisfied. This is readily derived. Suppose the above relation is satisfied for some v. We will show $-v$ to be tangent functional. We have

$$F_L[\rho_0] + \int \rho_0(\mathbf{r})v(\mathbf{r})d\mathbf{r} = E[v] \leq \inf_{\hat{D} \to \rho} \text{Tr}\hat{D}\hat{H} = F_L[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \tag{4.102}$$

Therefore

$$F_L[\rho] - F_L[\rho_0] \geq -\int v(\mathbf{r})(\rho(\mathbf{r}) - \rho_0(\mathbf{r}))d\mathbf{r} = L[\rho - \rho_0] \tag{4.103}$$

Where we define the linear functional $L : L^1 \cap \text{L}^3 \mapsto \mathcal{R}$ by

$$L : \rho \mapsto -\int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{4.104}$$

This functional is continuous. If we split up $v \in L^{\frac{3}{2}} + L^\infty$ as $v = u + w$ where $u \in L^{\frac{3}{2}}$ and $w \in L^\infty$ then we have

$$|L[\rho - \rho_0]| \leq \int |u(\mathbf{r})||\rho - \rho_0|d\mathbf{r} + \int |w(\mathbf{r})||\rho - \rho_0|d\mathbf{r} \leq \|u\|_{\frac{3}{2}}\|\rho - \rho_0\|_3 + \|w\|_\infty\|\rho - \rho_0\|_1 \tag{4.105}$$

So if $\rho \to \rho_0$ in the norm on $L^1 \cap L^3$ then $|L[\rho - \rho_0]|$ goes to zero, which proves the continuity of L. We will now show that if $-v$ is a tangent functional then equation 4.101 is satisfied. If equation 4.103 is satisfied then

$$E[v] \geq \inf_{\rho \in \mathcal{S}_N} \left[F_L[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \right] \geq \inf_{\rho \in \mathcal{S}_N} \left[F_L[\rho_0] + \int \rho_0(\mathbf{r})v(\mathbf{r})d\mathbf{r} \right] =$$

$$= F_L[\rho_0] + \int \rho_0(\mathbf{r})v(\mathbf{r})d\mathbf{r} \geq E[v] \tag{4.106}$$

which proves our statement.

We will now show that if $\rho_0 \in \mathcal{B}_N$ then F_L has a unique tangent functional $-v \in L^{\frac{3}{2}} + L^\infty$ at ρ_0. (Unique means here that v is determined to within an overall constant). Therefore F_L is Gâteaux differentiable on the set of ensemble v-representable densities. From the previous results it follows that to prove this we must show that for every $\rho_0 \in \mathcal{B}_N$ there is a unique potential v (to within a constant) which satisfies equation 4.101. If $\rho_0 \in \mathcal{B}_N$ then there is an external potential v and a ground state ensemble density matrix $\hat{D}[\rho_0]$ yielding density ρ_0. Then obviously

$$E[v] = \mathrm{Tr}\hat{D}[\rho_0]\hat{H} = \int \rho_0(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\rho_0] \tag{4.107}$$

We must now show its uniqueness. Suppose that we have a $\tilde{v} \neq v + C$ which satisfies

$$E[\tilde{v}] = \int \rho_0(\mathbf{r})\tilde{v}(\mathbf{r})d\mathbf{r} + F_L[\rho_0] \tag{4.108}$$

As we have $F_L[\rho_0] = \mathrm{Tr}\hat{D}[\rho_0](\hat{T} + \hat{W})$ we can also write

$$E[\tilde{v}] = \int \rho_0(\mathbf{r})\tilde{v}(\mathbf{r})d\mathbf{r} + \mathrm{Tr}\hat{D}[\rho_0](\hat{T} + \hat{W}) = \mathrm{Tr}\tilde{H}\hat{D}[\rho_0] = \sum_i \lambda_i \langle \psi_i | \tilde{H} | \psi_i \rangle \tag{4.109}$$

where $\tilde{H}$ is the Hamiltonian corresponding to external potential $\tilde{v}$ and $|\psi_i\rangle$ are ground states corresponding to potential v. However as the $|\psi_i\rangle$ are not ground states for $\tilde{H}$ we obtain

$$E[\tilde{v}] = \sum_i \lambda_i \langle \psi_i | \tilde{H} | \psi_i \rangle > E[\tilde{v}] \tag{4.110}$$

which is a contradiction and therefore v must be unique. We therefore have proven the differentiability of F_L on the set of ensemble-v-representable densities. We can further prove that F_L is not differentiable at the non-E-V-densities. As proven by Lieb, for any density $\rho \in \mathcal{S}_N$ there is a density matrix $\hat{D}[\rho]$ which minimizes the constrained search for F_L. Let us write

$$\hat{D}[\rho] = \sum_i \lambda_i |\psi_i\rangle\langle\psi| \tag{4.111}$$

Suppose that ρ is not E-V-representable, i.e. $\rho \notin \mathcal{B}_N$. Since ρ is not ensemble v-representable there is at least one $|\psi_i\rangle$ which cannot be a ground state for a potential v. Therefore

$$E[v] < \sum_i \lambda_i \langle \psi_i | \hat{H} | \psi_i \rangle = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\rho] \tag{4.112}$$

It follows that equation 4.101 can never be satisfied for non-E-V-densities, which yields that F_L is not differentiable at the non-E-V-densities.

Let us summarize the results of this section. The functional F_L is differentiable at the ensemble v-representable densities and nowhere else. This means that the Euler-Lagrange equation

$$\frac{\delta F_L}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) - \mu = 0 \tag{4.113}$$

is well-defined on the set $\mathcal{B}_N$. Here μ is a Lagrange multiplier to ensure the correct normalization of the density. The energy functional is minimized for the ground state ensemble densities corresponding to external potential v. As the Euler-Lagrange equation is well-defined for these densities, they are found as the solution of the above Euler-Lagrange equation. This then puts the variational equations within density functional theory on rigorous grounds.

4.7 The Kohn-Sham equations

Very important in any practical application of density functional theory are the Kohn-Sham equations. The Kohn-Sham system is introduced as the non-interacting system of particles which yield the same density as the fully interacting system of particles. If we consider a system non-interacting particles then $\hat{W} = 0$. For this system our functionals F_{HK}, F_{EHK}, F_{LL} and F_L then only contain the kinetic energy operator. We therefore denote these functionals as T_{HK}, T_{EHK}, T_{LL} and T_L and define

$$T_{HK}[\rho] = \langle \psi[\rho] | \hat{T} | \psi[\rho] \rangle \tag{4.114}$$

$$T_{EHK}[\rho] = \text{Tr}\hat{D}[\rho]\hat{T} \tag{4.115}$$

$$T_{LL}[\rho] = \inf_{\psi \to \rho} \langle \psi | \hat{T} | \psi \rangle \tag{4.116}$$

$$T_L[\rho] = \inf_{\hat{D} \to \rho} \text{Tr}\hat{D}\hat{T} \tag{4.117}$$

The functionals T_{LL} and T_L are defined on the same set $\mathcal{S}_N$ as F_{LL} and F_L for the interacting system. The functionals T_{HK} and T_{EHK} are now defined on the sets $\mathcal{A}'_N$ and $\mathcal{B}'_N$ of PS-V and E-V densities for the non-interacting system. The set $\mathcal{A}'_N$ is not convex, and convexity is unknown for the set $\mathcal{B}'_N$. The various convexity or non-convexity properties of the functionals F_{HK}, F_{EHK}, F_{LL} and F_L carry directly over to the functionals T_{HK}, T_{EHK}, T_{LL} and T_L. For T_L we can similarly to F_L prove that it is Gateaux differentiable on the set of non-interacting ensemble v-representable densities $\mathcal{B}'_N$.

Let us now define the exchange-correlation functional $E_{xc,L}$

$$E_{xc,L}[\rho] = F_L[\rho] - \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - T_L[\rho] \tag{4.118}$$

The functional $E_{xc,L}$ is defined on the set $\mathcal{S}_N$. As F_L is differentiable on $\mathcal{B}_N$ and T_L is differentiable on $\mathcal{B}'_N$ the exchange-correlation functional $E_{xc,L}$ is differentiable on the intersection of both sets $\mathcal{B}_N \cap \mathcal{B}'_N$.

Suppose now that an external potential in a non-interacting system v_s generates density $\rho \in \mathcal{B}'_N$ then the variational equation corresponding to the energy functional of the non-interacting system

$$E_{v_s}[\rho] = T_L[\rho] + \int \rho(\mathbf{r})v_s(\mathbf{r})d\mathbf{r} \tag{4.119}$$

is given by

$$\frac{\delta T_L}{\delta \rho(\mathbf{r})} + v_s(\mathbf{r}) - \mu_s = 0 \tag{4.120}$$

where μ_s is a Lagrange-multiplier to ensure the correct normalization of the density. As ρ is an enesemble v-representable ground state we also have

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{4.121}$$

$$\rho(\mathbf{r}) = \text{Tr}\hat{D}[\rho]\hat{\rho}(\mathbf{r}) = \sum_i \lambda_i|\phi_i(\mathbf{r})|^2 \tag{4.122}$$

On the other hand if ρ is also in $\mathcal{B}_N$ and is generated by an external potential v then ρ is obtained from the solution of the Euler-Lagrange equation

$$\frac{\delta F_L}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) - \mu = 0 \tag{4.123}$$

with μ again a Lagrange multiplier for the density normalization. So if $\rho \in \mathcal{B}_N \cap \mathcal{B}'_N$ we find

$$\frac{\delta T_L}{\rho(\mathbf{r})} + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{xc,L}}{\delta\rho(\mathbf{r})} - \mu = 0 \tag{4.124}$$

If we denote $v_{xc,L}(\mathbf{r}) = \delta E_{xc,L}/\delta\rho(\mathbf{r})$ we obtain (to within a constant)

$$v_s(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + v_{xc,L}(\mathbf{r}) \tag{4.125}$$

We therefore can find the ground state density for the interacting system by the solution of the equations

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' + v_{xc,L}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{4.126}$$

$$\rho(\mathbf{r}) = \text{Tr}\hat{D}[\rho]\hat{\rho}(\mathbf{r}) = \sum_i \lambda_i|\phi_i(\mathbf{r})|^2 \tag{4.127}$$

which constitute the Kohn-Sham equations. Note however that the above equations are only valid for densities in $\mathcal{B}_N \cap \mathcal{B}'_N$. Englisch and Englisch have however proved that for interacting ensemble v-representable densities the above Kohn-Sham equations always have a solution, which means $\mathcal{B}_N = \mathcal{B}'_N$. It further means that T_L is differentiable for all densities which are ensemble v-representable with respect to the interacting system. This then puts the Kohn-Sham equations on rigorous grounds.

Chapter 5

An exchange-correlation potential with correct asymptotic behaviour

In this work we analyse the properties of the exchange-correlation potential in the Kohn-Sham form of density functional theory, which leads to requirements for approximate potentials. Fulfilment of these requirements is checked for existing gradient-corrected potentials. In order to examine the behaviour of approximate potentials over all space we compare these potentials with exact Kohn-Sham potentials calculated from correlated densities using a newly developed iterative procedure. The main failures in the existing gradient corrected potentials arise in the asymptotic region of the atom where these potentials decay too fast and at the atomic nucleus where the potentials exhibit a Coulomb-like singular behaviour. We show that the main errors can be corrected by a simple potential in terms of the density and its gradients leading to considerably improved one-electron energies compared to the local density approximation. For Be and Ne it is shown that the electron density is improved in the outer region.

5.1 Introduction

In the past few years there has been considerable progress within density functional theory [5] in the calculation of properties of electronic systems ranging from solids to atoms and molecules. This progress is due to the introduction of gradient-corrected density functionals [44, 45, 43, 50] that give an overall improvement to the exchange-correlation energies of the local density approximation (LDA). For instance the atomisation energies of a standard set of molecules are improved by an order of magnitude compared to LDA[70], thereby correcting the overbinding of LDA. There are also succesful applications of nonlocal corrections in transition metal chemistry [71] and solid state and surface physics [47]. (Although the gradient corrected potentials are still local, we follow the conventional nomenclature of "nonlocal corrections" to distinguish from the LDA.) However several other features are not improved by the present day nonlocal corrections. This is especially the case for properties that are sensitively dependent on the behaviour of the exchange-correlation *potential*. One can for instance prove rigorously [31, 72] that the eigenvalue of the highest occupied Kohn-Sham orbital represents the ionisation energy of the system. However, typical errors in LDA for this quantity are 5 eV. This same error prevents the calculations of bound state solutions for negative ions as LDA gives an unbound outer electron with positive eigenvalue. The origin of this error can be traced to the incorrect asymptotic decay of the LDA exchange-correlation potential. This potential has an exponential decay into the vacuum as can directly be seen from the expontial decay

of the density itself. On physical grounds however (see for a proof [31]) the outer electron should experience the mean field of the ion it leaves behind, i.e. a potential that decays Coulombically like $-1/r$. The LDA electron is therefore too weakly bound and for negative ions even unbound. This breakdown of LDA in the outer region of the atom or molecule is also reflected in the exchange-correlation energy per particle $\epsilon_{xc}(\mathbf{r})$ which can equivalently be seen als the potential due to the exchange-correlation hole and which has an asymptotic decay like $-1/2r$. The LDA in this case gives again an exponential decay. In the electron gas this quantity ϵ_{xc} is usually expressed in terms of the Wigner-Seitz radius r_s representing the mean electronic distance which is proportional to $\rho^{-\frac{1}{3}}$. If the local density approximation is applied to the outer regions of atoms and molecules r_s grows exponentially and loses its meaning as a mean interelectronic distance which should grow linearly. If one believes (for intuitive physical reasons) that the mean electronic distance determines ϵ_{xc} then the bad representation of this quantity by LDA explains the failure of LDA in this region. One might now wonder if the present day nonlocal gradient corrections give any improvement for this asymptotic failure of LDA. Somewhat surprisingly, this is not the case. Although they give large improvements in energies they give little improvement in the asymptotic behaviour of the exchange-correlation potential. This is immediately apparent from the fact, undoubtedly noted by many DFT practitioners, that the gradient corrected potentials yield almost no improvement in the LDA eigenvalues, which are generally in error by 5-6 eV. In this paper we discuss, apart from the fairly well known asymptotic behaviour, other requirements which are to be fulfilled by the exact exchange or exchange-correlation potential. Those requirements apply to limiting or special situations ($r \to \infty$, $r \downarrow 0$, transition regions from one atomic shell to the next, limit of homogeneous electron gas) and also comprise invariance conditions (translational, rotational). They are useful in constructing approximate Kohn-Sham potentials. In section 5.3 we use the formulated requirements to examine some of the presently used nonlocal functionals, in particular the Becke correction for exchange and the Perdew correction for correlation. In order to investigate the potentials at arbitrary $\vec{r}$, we need the exact Kohn-Sham potential over all space. In section 5.4 we discuss a general procedure to construct the Kohn-Sham exchange-correlation potential from a given electron density ρ. The procedure is simple and is applicable to both atoms and molecules and to systems with an arbitrary number of electrons. This procedure is used to generate exact Kohn-Sham potentials from very accurate (highly correlated) densities of Be and Ne. This affords detailed insight in the strengths and weaknesses of existing approximations. As a first step towards improved potentials, we propose in section 5.5 a model Kohn-Sham potential which exhibits the correct asymptotic behaviour and also displays atomic shell structure. This potential gives a large improvement over the LDA eigenvalues and is in fact capable of yielding good ionisation energies as determined from the highest occupied Kohn-Sham orbital energy. It also improves the asymptotic decay of the electron density. In section 5.6 we present a summary and conclusions.

5.2 Nonlocal exchange-correlation potentials : requirements

In this section we will discuss some of the requirements that approximate exchange-correlation potentials have to satisfy. In connection with this we shall demonstrate in the next section which of these requirements are lacking with some of the presently used nonlocal potentials.

Scaling. The exchange energy within density functional theory may be defined as [23]:

$$E_x[\rho] = \int \rho(\mathbf{r})\epsilon_x(\mathbf{r})d\mathbf{r} \tag{5.1}$$

in which the the potential of the exchange-hole $\epsilon_x(\mathbf{r})$ or equivalently the exchange energy density per electron is defined as:

$$\epsilon_x(\mathbf{r}) = -\frac{1}{2\rho(\mathbf{r})} \int \frac{|\gamma_s(\mathbf{r}, \mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \tag{5.2}$$

Here γ_s is the Kohn-Sham one-particle density matrix, constructed from a determinant of Kohn-Sham orbitals. The exchange functional of eq. 5.1 is a functional of the density as the Kohn-Sham orbitals of which the one-particle density matrix γ_s is composed are uniquely determined by the density (an explicit scheme for doing this is presented in section 5.4). The exchange functional satisfies the following *scaling* relation [22]:

$$E_x[\rho_\lambda] = \lambda E_x[\rho] \tag{5.3}$$

in which ρ_λ is the following scaled density:

$$\rho_\lambda(\mathbf{r}) = \lambda^3 \rho(\lambda \mathbf{r}) \tag{5.4}$$

The exchange potential which is defined as the functional derivative of the exchange functional satisfies the following scaling relation [23, 22]:

$$v_x([\rho_\lambda]; \mathbf{r}) = \lambda v_x([\rho]; \lambda \mathbf{r}) \tag{5.5}$$

Using the scaling relation 5.3 one can prove the Levy-Perdew relation [22]:

$$E_x[\rho] = -\int \rho(\mathbf{r})\mathbf{r} \cdot \nabla v_x(\mathbf{r}) d\mathbf{r} \tag{5.6}$$

Any approximate exchange potential should satisfy equation 5.5. It is then possible to define an approximate exchange energy using the Levy-Perdew relation 5.6. We will return to this relation between potential and energy below.

There are no known scaling relations for the correlation energy functional defined as:

$$E_c[\rho] = E_{xc}[\rho] - E_x[\rho] \tag{5.7}$$

in which E_{xc} is the Kohn-Sham exchange-correlation energy which can be defined for instance using the coupling strength integration method [73, 24, 25, 26]

Asymptotic behaviour. An approximate functional or potential also has to satisfy some requirements with respect to asymptotic behaviour. First of all it follows from equation 5.2 using the sum rule property of the one-particle density matrix γ_s that:

$$\epsilon_x(\mathbf{r}) \sim -\frac{1}{2r} \qquad (r \to \infty) \tag{5.8}$$

Thus the potential of the exchange hole has a Coulombic asymptotic behaviour. The exchange potential has to fulfil a similar type of relation [35, 49]:

$$v_x(\mathbf{r}) \sim -\frac{1}{r} \qquad (r \to \infty) \tag{5.9}$$

The asymptotic behaviour of the potential due to the (coupling strength integrated) Coulomb hole is not known. We can only say that the correlation potential must decay faster than a Coulombic

potential at infinity. This follows from the asymptotic $-1/r$ decay of the total Kohn-Sham potential [74, 72]:

$$v_{xc}(\mathbf{r}) \sim -\frac{1}{r} \qquad (r \rightarrow \infty) \tag{5.10}$$

and the similar behaviour of the exchange potential of eq. 5.9. This result is consistent with the fact that the Coulomb hole of the coupling strength integrated two-particle density matrix integrates to zero electrons.

Weak inhomogeneity. Further known properties of the exchange functional follow from the gradient expansion of the weakly varying electron gas [10] or from the semi-classical expansion of the Kohn-Sham one-particle density matrix [75]. This gives the following approximate nonlocal correction to the exchange energy (for the spin unpolarised case):

$$E_x[\rho] = -\beta \int \frac{(\nabla\rho)^2}{\rho^{\frac{4}{3}}} d\mathbf{r} \tag{5.11}$$

This is the nonlocal correction used in the $X\alpha\beta$-approximation [76]. For weakly varying densities there is a known gradient expansion for the correlation energy. The nonlocal correction is (up to second order) given by:

$$E_c[\rho] = \int C(\rho) \frac{(\nabla\rho)^2}{\rho^{\frac{4}{3}}} d\mathbf{r} \tag{5.12}$$

in which $C(\rho)$ is a local function of the electron density [42].

Translational and rotational invariance. Two requirements which must be fulfilled by any density functional representing a physical quantity are translational and rotational invariance. The question of translational invariance for instance arises naturally in the discussion of the asymptotic $-1/r$ behaviour of the energy densities and the exchange or exchange-correlation potentials, where r should not refer to the distance from the arbitrary origin of the coordinate system. Denoting a physical quantity by A, translational invariance means:

$$A[\rho'] = A[\rho] \tag{5.13}$$

with

$$\rho'(\mathbf{r}) = \rho(\mathbf{r} + \vec{R}) \tag{5.14}$$

in which $\vec{R}$ is an arbitrary translation vector. This means physically that A should not change when we translate our coordinate system. If we define:

$$a([\rho]; \mathbf{r}) = \frac{\delta A}{\delta \rho(\mathbf{r})}[\rho] \tag{5.15}$$

then it follows that:

$$a([\rho']; \mathbf{r}) = a([\rho]; \mathbf{r} + \vec{R}) \tag{5.16}$$

This does not hold for arbitrary $a([\rho]; \mathbf{r})$ but it does hold if a is the functional derivative of a translational invariant functional A. This equation is easily proved if one uses:

$$\delta A[\rho] = \int \delta\rho(\mathbf{r}) a([\rho]; \mathbf{r}) d\mathbf{r} = \int \delta\rho(\mathbf{r} + \vec{R}) a([\rho]; \mathbf{r} + \vec{R}) d\mathbf{r} = \int \delta\rho'(\mathbf{r}) a([\rho]; \mathbf{r} + \vec{R}) d\mathbf{r} \tag{5.17}$$

Comparing this equation with:

$$\delta A[\rho'] = \int \delta \rho'(\mathbf{r}) a([\rho']; \mathbf{r}) d\mathbf{r} \qquad (5.18)$$

yields equation 5.16 as both variations should be equal for any variation $\delta\rho'(\mathbf{r}) = \delta\rho(\mathbf{r} + \vec{R})$. We will now take $A = E_x$ and $a = v_x$. Noting the explicit $\mathbf{r}$ dependence in the Levy-Perdew relation between E_x and v_x one might wonder whether this relation is translationally invariant. The exact exchange potential of course satisfies the translational invariance equation 5.16. Then it follows from the Levy-Perdew relation 5.6, if we insert ρ', that:

$$E_x[\rho'] = E_x[\rho] + \vec{R} \cdot \int \rho(\mathbf{r}) \nabla v_x([\rho]; \mathbf{r}) d\mathbf{r} \qquad (5.19)$$

Translational invariance requires the last term in this equation to be zero. As this should be true for any vector $\vec{R}$ we obtain (after carrying out a partial integration):

$$\int v_x(\mathbf{r}) \nabla \rho(\mathbf{r}) d\mathbf{r} = 0 \qquad (5.20)$$

Translational invariance thus gives an additional condition on the potential. The above formula is a special case of a more general result which follows directly from equation 5.13:

$$A[\rho] = A[\rho'] = A[\rho + \vec{R} \cdot \nabla \rho + O(R^2)] = A[\rho] + \int \frac{\delta A}{\delta \rho(\mathbf{r})} \vec{R} \cdot \nabla \rho(\mathbf{r}) d\mathbf{r} + O(R^2) \qquad (5.21)$$

As this equation should be valid for any translation vector $\vec{R}$ it follows that:

$$\int \frac{\delta A}{\delta \rho(\mathbf{r})} \nabla \rho(\mathbf{r}) d\mathbf{r} = 0 \qquad (5.22)$$

This equation has also been noted without proof in reference [23]. Equation 5.20 is then obtained by taking $A = E_x$ in the last equation. In particular for $A = E_{xc}$ we obtain:

$$\int v_{xc}(\mathbf{r}) \nabla \rho(\mathbf{r}) d\mathbf{r} = 0 \qquad (5.23)$$

For the case of rotational invariance we require equation 5.13 to be valid for:

$$\rho'(\mathbf{r}) = \rho(R\mathbf{r}) \qquad (5.24)$$

in which R is a rotation operation within three-dimensional coordinate space. For the functional derivative of A we then have:

$$a([\rho']; \mathbf{r}) = a([\rho]; R\mathbf{r}) \qquad (5.25)$$

This equation is proven in the same way as equation 5.16. If we take $a = v_x$ and use equation 5.25 we see that the Levy-Perdew equation 5.6 is already rotationally invariant. In general rotational invariance gives the following constraint on the functional derivative of A:

$$\int \frac{\delta A}{\delta \rho(\mathbf{r})} \mathbf{r} \times \nabla \rho(\mathbf{r}) d\mathbf{r} = 0 \qquad (5.26)$$

which can be proved by performing an infinitisimal rotation instead of a translation in equation 5.21. This gives :

$$\int v_{xc}(\mathbf{r}) \mathbf{r} \times \nabla \rho(\mathbf{r}) d\mathbf{r} = 0 \qquad (5.27)$$

We have now summarised some properties of the exact exchange-correlation potential, which are at the same time requirements to be fulfilled by approximate potentials. We will concentrate in the remainder of this paper on the potential rather than on the energy. One of the reasons is that the potential, being a unique, local function of the position $\vec{r}$, is more easily obtained, analysed and modelled (see below). Of course, even if one has obtained a good model potential that very closely approximates the exact Kohn-Sham potential, the exchange-correlation energy still has to be determined. This problem will be adressed in a subsequent paper.

5.3 The nonlocal exchange and correlation potentials of Becke and Perdew

Using the conditions formulated in the previous section we shall now discuss some of the currently used exchange-correlation functionals and potentials that include nonlocal corrections. These are the potentials derived from Becke's nonlocal exchange functional [44] and Perdews nonlocal correlation functional [45]. First we discuss Becke's nonlocal exchange functional [44] which in spin polarised form is given by:

$$E_x[\rho_\sigma] = \sum_\sigma \int \rho_\sigma^{\frac{4}{3}}(\mathbf{r}) f(x_\sigma(\mathbf{r})) d\mathbf{r} \tag{5.28}$$

in which $x_\sigma = |\nabla \rho_\sigma|/\rho_\sigma^{\frac{4}{3}}$ is a dimensionless quantity and σ is a spin index. The function f is given by:

$$f(x) = -\beta \frac{x^2}{1 + 6\beta x \sinh^{-1}(x)} \tag{5.29}$$

The form of equation 5.28 is chosen such that the exchange functional satisfies the scaling relation 5.3. The function f is chosen such that the potential of the exchange hole $\epsilon_x(\mathbf{r})$ or equivalently the exchange energy density per electron behaves asymptotically as $-1/2r$ (eq. 5.8). To enforce relation 5.8 the function f in equation 5.28 must satisfy the following asymptotic relation:

$$f(x) \sim -\frac{1}{6} \frac{x}{\log(x)} \qquad (r \to \infty \Rightarrow x \to \infty) \tag{5.30}$$

which is easily verified if exponentially decaying densities are inserted. If one also wants to obtain the gradient expansion result of equation 5.11 for slowly varying densities f must satisfy:

$$f(x) \sim -\beta x^2 \qquad (x \downarrow 0) \tag{5.31}$$

The form chosen by Becke is one of the simplest interpolations of f between these two limits and therefore his energy functional satisfies the important requirements of correct $r \to \infty$ and $x \to 0$ behaviour as well as translational and rotational invariance. However, the potential of the Becke energy expression decays asymptotically like [77, 49]:

$$v_x^{Becke}(\mathbf{r}) \sim \frac{k}{r^2} \qquad (r \to \infty) \tag{5.32}$$

in which k is some constant instead of the exact [35, 49]:

$$v_x(\mathbf{r}) \sim -\frac{1}{r} \tag{5.33}$$

In the following we will prove that any functional of the form of equation 5.28 satisfying the asymptotic relation 5.30 will not satisfy asymptotic relation 5.33. Such a proof has also been given by Engel et al. [49] and is only presented here for reasons of clarity and completeness. We will however also use the following short proof to make some statements about the behaviour of the exchange potential near the atomic nucleus for exchange functionals of the form of equation 5.28

For simplicity we use the spin unpolarised form but everything goes exactly through for the spin polarised case by just adding a spin index everywhere in the derivation. This is due to the fact that the exchange energy in the spin polarised case is just a simple sum of contributions of α and β spins in which both contributions have the same structure.
We will take the functional of the form:

$$E_x[\rho] = \int \rho^{\frac{4}{3}}(\mathbf{r}) f(x(\mathbf{r})) d\mathbf{r} \tag{5.34}$$

with

$$x = \frac{|\nabla \rho|}{\rho^{\frac{4}{3}}} \tag{5.35}$$

The second derivative of the function f is assumed to exist. The functional derivative of E_x of 5.34 is then given by:

$$\frac{\delta E_x}{\delta \rho(\mathbf{r})} = v_x(\mathbf{r}) = \frac{4}{3}\rho^{\frac{1}{3}}(f - x\frac{df}{dx} + x^2\frac{d^2 f}{dx^2})$$

$$+ \sum_{i,j} \frac{\partial_i \rho \partial_i \partial_j \rho \partial_j \rho}{|\nabla \rho|^3}(\frac{df}{dx} - x\frac{d^2 f}{dx^2}) - \frac{\nabla^2 \rho}{|\nabla \rho|}\frac{df}{dx} \tag{5.36}$$

We insert for the density the exponential $\rho(\mathbf{r}) = Ne^{-\alpha r}$. This exponential is exact in two regions of the atom: near the atomic nucleus where we have $\alpha = 2Z$ with Z the nuclear charge and for the outer asymptotic region where $\alpha = 2\sqrt{-2\mu}$ with μ the chemical potential (negative of the ionisation energy). This gives:

$$v_x(\mathbf{r}) = \frac{4}{3}\rho^{\frac{1}{3}}(f - x\frac{df}{dx} + x^2\frac{d^2 f}{dx^2}) - \alpha x\frac{d^2 f}{dx^2} + \frac{2}{r}\frac{df}{dx} \tag{5.37}$$

In order to satisfy the asymptotic behaviour of ϵ_x (equation 5.8) f must satisfy the asymptotic relation 5.30. If we now insert the large x behaviour of f (equation 5.30) into equation 5.37 we get retaining only the terms of lowest order in $1/\log x$:

$$v_x(\mathbf{r}) \sim -\frac{1}{3r}\frac{1}{\log x} \sim -\frac{1}{3r}\frac{1}{\log \alpha \rho^{-\frac{1}{3}}} \sim -\frac{1}{\alpha}\frac{1}{r^2} \qquad (r \to \infty) \tag{5.38}$$

We thus see that the exchange potential has a $-\alpha^{-1}/r^2$ behaviour instead of a Coulombic behaviour. Hence it follows that the requirements of correct asymptotic behaviour of ϵ_x (relation 5.8) and of v_x (relation 5.33) are incompatible for exchange functionals of the form 5.34.
From equation 5.37 we can also draw the following conclusion concerning the behaviour for $r \to 0$:
If $\frac{df}{dx}|_{x=x(\vec{0})} \neq 0$ then:

$$v_x(\mathbf{r}) \sim \frac{k}{r} \qquad (r \downarrow 0) \tag{5.39}$$

in which the constant k is given by:

$$k = 2\frac{df}{dx}\big|_{x=x(\vec{0})} \tag{5.40}$$

This is an unphysical behaviour of the exchange potential near the nucleus. In practice however (that is for the Becke functional) the constant k is very small compared to the nuclear charge (a typical value for k is 0.02 for the noble gases).

We will now discuss some properties of the Perdew nonlocal correlation functional which is of the form [45]:

$$E_c[\rho] = \int f(\rho_\alpha, \rho_\beta) e^{-g(\rho)|\nabla\rho|} \frac{(\nabla\rho)^2}{\rho^{\frac{4}{3}}} d\mathbf{r} \tag{5.41}$$

In this formula f is a local function of ρ_α and ρ_β which are spin densities and g is a local function of the total density. The form of this functional is adapted from the correlation part of the Langreth-Mehl functional in such a way that the functional reduces to the gradient expansion expression of equation 5.12 for slowly varying densities. The corresponding potential decays exponentially to infinity. This is not a bad feature as the correlation potential should decay faster than Coulombic although maybe one should expect a correlation potential decaying like $-\alpha/r^4$ in which α is the polarisability of the system. This term occurs in the exact exchange-correlation potential [74] and describes the polarisation of the system by an asymptotic electron, which is clearly a correlation effect.

Near the nucleus the Perdew potential also has a Coulombic singular behaviour. The origin of this unphysical singularity can be traced, like in the Becke potential, to terms in the potential that contain the Laplacian of the density, which due to the Slater-like behaviour of the atomic density near the nucleus leads to a Coulombic potential. These terms also occur in potentials of other nonlocal density functionals such as the one from Langreth and Mehl [50] and in the Generalized Gradient Approximation (GGA) of Perdew and Wang [46]. These functionals consequently also suffer from this same deficiency.

5.4 Constructing the potential from the density

We have demonstrated that the most widely used gradient corrected exchange and correlation potentials suffer from incorrect asymptotic behaviour at $r \to 0$ and $r \to \infty$. In order to investigate these potentials over the whole r range one needs the exact Kohn-Sham potential at arbitrary position. In this section we present an iterative scheme of obtaining the Kohn-Sham potential from a given electron density ρ. This scheme is different from most other schemes used thus far [78, 74, 79] and in particular is not limited to two-electron systems but is applicable to systems with any number of electrons. Recently during the writing of this paper we have noticed that work along similar lines has been carried out [80]. However our procedure is somewhat simpler and as we tested has the same convergence rate. In contrast to reference [80] in which calculations within a basis set were performed we use a completely numerical approach obtaining a basis set free representation of the potential.

We start from the Kohn-Sham-equations:

$$(-\frac{1}{2}\nabla^2 + v_s(\mathbf{r}))\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{5.42}$$

in which v_s is the Kohn-Sham-potential. The Kohn-Sham-orbitals are required to satisfy:

$$\sum_i^N |\phi_i(\mathbf{r})|^2 = \rho(\mathbf{r}) \tag{5.43}$$

where N is the number of electrons in the system. Multiplying equation 5.42 by ϕ_i^* and summing over i gives after dividing by ρ:

$$v_s(\mathbf{r}) = \frac{1}{\rho(\mathbf{r})} \sum_i^N \frac{1}{2} \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) + \epsilon_i |\phi_i(\mathbf{r})|^2 \tag{5.44}$$

We now define an iterative scheme using this equation. We want to calculate the potential corresponding to the density ρ. Suppose that at some stage in the iteration we have calculated orbitals ϕ_i^o with eigenvalues ϵ_i^o and density ρ^o and potential v^o. In the next step we define the new potential:

$$v^n(\mathbf{r}) = \frac{1}{\rho(\mathbf{r})} \sum_i^N \frac{1}{2} \phi_i^{o*}(\mathbf{r}) \nabla^2 \phi_i^o(\mathbf{r}) + \epsilon_i^o |\phi_i^o(\mathbf{r})|^2$$

$$= \frac{\rho^o(\mathbf{r})}{\rho(\mathbf{r})} v^o(\mathbf{r}) \tag{5.45}$$

Using this potential we calculate new orbitals and a new density and define in the same way a new potential. This procedure is continued until the density calculated from the orbitals is the same as the given density. In practice until:

$$\max_{\mathbf{r}} |1 - \frac{\rho^o(\mathbf{r})}{\rho(\mathbf{r})}| < \epsilon \tag{5.46}$$

with ϵ a given threshold. To achieve convergence one should take care to keep the prefactor in the last term of equation 5.45 in each iteration within an acceptable range:

$$1 - \delta < \frac{\rho^o(\mathbf{r})}{\rho(\mathbf{r})} < 1 + \delta \tag{5.47}$$

for example with $\delta = 0.05$. We noticed in the application of this procedure that one has to make sure that the potential is set to zero in infinity otherwise one might fail to converge. This is due to the fact that one can add an arbitrary constant to both the potential and the one electron eigenvalues without changing the density. This may also play a role in the practical application of the procedure of reference [80] where explicit use is made of the one electron eigenvalues in the potential construction procedure. The fixing of the potential in infinity is however easily carried out in our program in which differential equations are solved numerically. In that case boundary conditions at infinity and at the nucleus immediately fix the solution.

The scheme is not guaranteed to converge as there are densities which are not non-interacting v-representable (however many densities which are not v-representable by one determinant are v-representable by a linear combination of determinants [81] which can be accomplished in our scheme by using fractional occupation numbers). However if the procedure converges then its limit is unique as guaranteed by the Hohenberg-Kohn theorem applied to a non-interacting electron system [10].

We have calculated exact potentials for the beryllium and the neon atoms as for these atoms accurate densities from configuration interaction calculations are available. Both densities have been published by Bunge and Esquivel [82, 83] in a large basis of Slater functions. The use of Slater functions is advantageous because they give a much better representation of the density tail in the outer regions of the atom than for instance Gaussian functions. For the calculation of the potentials from these densities we used a modification of the Herman-Skilman atomic program [84] based on numerical integration. For the results of table 1 we have used our atomic and molecular density functional package based on Slater Type Orbitals (STO's) with which it is possible to carry out self-consistent calculations using the Becke-Perdew potential. In the local density approximation we use for the correlation potential the VWN-parametrisation of the electron gas data [85]. For the open shell atoms discussed in this article we performed spherically averaged spinpolarised calculations. The exact exchange-correlation potentials are displayed in figure 5.1a and figure 5.2a for Be and Ne respectively and r times these potentials in figs.5.1c and 5.2c. Both potentials have the same structure, a characteristic peak between the atomic shells (in our case between the K and the L-shell) and a Coulombic asymptotic behaviour. These features are most clearly displayed in the plots of rV. The appearance of the intershell peak has been observed before [74, 79] and can be understood from the work of Buijse et al. [86]. In ref. [86] it has been observed that an important contribution to the Kohn-Sham potential is the socalled kinetic potential V_{kin} defined in terms of the conditional amplitude Φ:

$$\Phi(x_2, \ldots, x_N | x_1) = \frac{\Psi(x_1, \ldots, x_N)}{\sqrt{\rho(x_1)/N}} \tag{5.48}$$

$$V_{kin}(x_1) = \int \Phi^*(-\frac{1}{2}\nabla_1^2)\Phi dx_2 \ldots dx_N = +\frac{1}{2}\int |\nabla_1 \Phi|^2 dx_2 \ldots dx_N \tag{5.49}$$

The conditional amplitude $\Phi(x_2, \ldots, x_N | x_1)$ describes the system of $N-1$ electrons with positions $x_2, \ldots, x_N$ when one electron is known to be at position x_1 and is the amplitude connected with the conditional probability of finding the other electrons when one electron is known to be at x_1. V_{kin} makes a significant (positive) contribution to the Kohn-Sham potential at those positions x_1 of the reference electron where the conditional amplitude changes rapidly, so that $\nabla_1 \Phi$ is large. As discussed in ref. [86], this is the case when x_1 crosses the border region between two atomic shells, since the exchange hole is localized within one atomic shell if the reference position is in that shell but "jumps" to the next shell when the reference position crosses the border (see refs. [87, 88, 89]). The intershell peak of V_{xc} is therefore an exchange effect and reflects the fact that the atomic shell structure has its origin in the antisymmetry principle. The peak in the potential helps to build the shell structure in the total density by reducing the density in the intershell region. It is obvious from the figures that the LDA potentials almost completely lack this important feature of the exact potential. This is particularly clear from the pronounced appearence of peaks in the difference plots of $(V_{XC} - V_{LDA})$ and $r(V_{XC} - V_{LDA})$ in figs. 5.1b,d and 5.2b,d. The LDA potential also has a wrong asymptotic behaviour for $r \to \infty$, as is evident from the fact that rV_{LDA} does not go to -1 but to 0. The exact potential has a much improved asymptotic behaviour. However, the quality of the exact potential we generate depends on the quality of the correlated wavefunction and density on which it is based. The asymptotic region is notoriously difficult to describe accurately, since wavefunctions are almost invariably obtained from energy optimization algorithms which have a strong bias towards improving the energetically important inner region. In the case of Neon we observe that for large r ($r \geq 3$ bohr) there is a spurious minimum in the curve for the exact V_{XC} (fig.5.2c) which we ascribe to inaccuracy of the correlated density at such large r. It is nevertheless

Figure 5.1: Exchange-correlation potentials of the beryllium atom.

Figure 5.2: Exchange-correlation potentials of the neon atom.

clear that the "exact" rV_{XC} approaches -1 much better than the LDA potential does. For Be it is not evident that the calculated "exact" V_{XC} suffers from poor accuracy of the asymptotic behaviour of the CI density, but we do feel that some suspicion is warranted concerning the (too?) slow approach by rV_{XC} to the limiting value of -1 (fig.5.1c). This suspicion is aggravated by the strange minimum between 6 and 8 bohr in the curve of $r(V_{XC} - V_{LDA})$ in fig.5.1d.

In panels a and c of figures 5.1 and 5.2 we have also added Beckes gradient-corrected exchange potential [44] and Perdews correlation potential[45] to the LDA potential. The plots demonstrate that these potentials make rather small corrections to the LDA potential. This makes it understandable that the one-electron energies shown in table 1, which are rather poor for LDA when compared to the experimental ionisation energies, do not improve when the Becke-Perdew non-local corrections are included in the SCF calculation. Panels b and d of the figures display directly the comparison between the non-local corrections to the potential and the difference $(V_{XC} - V_{LDA})$ to which they should be equal. It can be seen that the nonlocal Becke-Perdew potential behaves singularly near the atomic nucleus. This potential has in this region the Coulombic singular behaviour discussed earlier. At somewhat larger distances it crosses the horizontal axis and gives a small positive peak which is located at the right spot in the intershell region. For Be this does not lead to good agreement with $(V_{XC} - V_{LDA})$ in the intershell region, but for Ne $V_{BECKE-PERDEW}$ quite nicely approximates $(V_{XC} - V_{LDA})$ just at the position of the intershell peak. At larger distances the Becke-Perdew potential is almost zero and there is no correction to the LDA potential, in agreement with the failure to give improvement of the LDA eigenvalues. The large deviation in the asymptotic region between the Becke-Perdew potential and the exact nonlocal potential is most clearly exhibited in panels d: the exact nonlocal corrections correctly tend to a constant which approximates -1, whereas the Becke-Perdew potential tends to zero. The panels d also demonstrate that the Becke-Perdew potential multiplied by the radial distance does not tend to zero at the atomic nucleus but to a finite value, indicating the Coulombic behaviour of this potential in this region.

5.5 A model potential

In this section we will make a first step towards the construction of model potentials that more closely resemble the exact Kohn-Sham potential. The advantage in modeling potentials instead of energy expressions is the fact that potentials are uniquely defined by the exact density. The quality of model potentials can therefore be judged by comparing to exact potentials calculated from accurate densities.

In modeling the potentials one should incorporate the general features of atomic shell structure and asymptotic Coulombic behaviour. One might also wish the potential to satisfy some scaling properties. To incorporate the shell structure we use for our model potential the dimensionless parameter $x = |\nabla\rho|/\rho^{\frac{4}{3}}$. This parameter is proportional to the length of the gradient of the local Wigner-Seitz radius and can be interpreted as the change in mean electronic distance (at least in regions with slowly varying density where $r_s \sim \rho^{-1/3}$ is meaningful, i.e. in the regions where LDA is applicable). As an illustration of the behaviour of this parameter we display both this parameter and $r^2\rho(\mathbf{r})$ in figure 5.3 for the krypton atom. Whereas the electron density $\rho(\mathbf{r})$ is monotically decreasing in an atom, $r^2\rho(\mathbf{r})$ exhibits the shell structure. We observe that x also oscillates with the atomic shells, with maxima approximately at the inflection points of $r^2\rho(\mathbf{r})$, and thus seems a suitable parameter to model the shell structure of the exact potential. For the asymptotic behaviour of the potential we then have to take into account that x behaves for an exponentially decaying

Figure 5.3: The electron density times the radial distance squared and the parameter $x = |\nabla \rho| / \rho^{\frac{4}{3}}$ for the krypton atom.

density $\rho \sim e^{-\alpha r}$ asymptotically as $\alpha \rho^{-1/3}$, i.e. increases exponentially (see fig. 5.3). We thus choose our nonlocal correction to the LDA potential of the form

$$v_{xc}(\mathbf{r}) = \rho^{\frac{1}{3}}(\mathbf{r}) f(x(\mathbf{r})) \tag{5.50}$$

This form of the potential scales like an exchange potential [22, 23]:

$$v_{xc}([\rho_\lambda]; \mathbf{r}) = \lambda v_{xc}([\rho]; \lambda \mathbf{r}) \tag{5.51}$$

We do not know the scaling behaviour of the correlation part of the potential but as the major part of the potential comes from the exchange we take the above form as an approximation. For systems with small density variations (small values of x) we want the nonlocal correction potential also to be small. We therefore require $f(0) = 0$. We further know that [31, 72]:

$$v_{xc}(\mathbf{r}) \sim -\frac{1}{r} \qquad (r \to \infty) \tag{5.52}$$

This means that f must asymptotically satisfy:

$$f(x) \sim -\frac{1}{3} \frac{x}{\log(x)} \qquad (x \to \infty) \tag{5.53}$$

Our problem of finding a smooth interpolation between these limiting situations is similar to the one that Becke faced for the exchange energy density. Inspired by his solution we take:

$$f(x) = -\beta \frac{x^2}{1 + 3\beta x \sinh^{-1}(x)} \tag{5.54}$$

ATOM	HF	LDA	NL(BP)	MODEL	EXPT
H	0.500	0.234	0.280	0.440	0.500
He	0.918	0.571	0.585	0.851	0.903
Be	0.309	0.206	0.209	0.321	0.343
Ne	0.850	0.490	0.496	0.788	0.792
Ar	0.591	0.381	0.380	0.577	0.579
Kr	0.524	0.346	0.344	0.529	0.517
Xe	0.457	0.310	0.308	0.474	0.446
ION	HF	LDA	NL(BP)	MODEL	EXPT
F^-		-0.097	-0.099	0.128	0.125
Cl^-		-0.022	-0.023	0.140	0.133
Br^-		-0.008	-0.009	0.140	0.124
I^-		+0.005	+0.004	0.139	0.112
MOLEC	HF	LDA	NL(BP)	MODEL	EXPT
N_2	0.622	0.328	0.322	0.557	0.573
F_2	-	0.339	0.334	0.607	0.582
CO	0.551	0.334	0.336	0.529	0.515

Table 5.1: ionisation energies and electron affinities from the highest occupied Kohn-Sham orbital

This gives the final form of our model potential for the nonlocal corrections. In spin polarised form:

$$v_{xc}^{\sigma}(\mathbf{r}) = -\beta \rho_{\sigma}^{\frac{1}{3}}(\mathbf{r}) \frac{x_{\sigma}^2}{1 + 3\beta x_{\sigma} \sinh^{-1}(x_{\sigma})} \tag{5.55}$$

In order to check if with this potential we have captured the main features of the Kohn-Sham potential which the LDA potential is lacking, i.e. shell-structure and a Coulombic asymptotic behaviour, we compare to exact potentials in figs. 5.1 and 5.2. The parameter β in our model potential was fitted in such a way that our model potential resembled as closely as possible the difference between the exact and the LDA potential for the beryllium atom. For this procedure we choose the beryllium atom instead of the neon atom because its density appears to be the most accurate of the two. This leads to a value of β of 0.05. The model potential for neon in figure 5.2 uses the same parameter β. The model potentials in figures 5.1 and 5.2 have been calculated selfconsistently using our density functional program package.

As can be seen from figure 5.1 for the beryllium atom our potential is in reasonable agreement with the exact potential. In particular figs. 5.1c,d show that the intershell peak is fairly well represented and the asymptotic behaviour is essentially correct in that $r(V_{LDA} + V_{MODEL})$ approaches -1. Concerning the remaining difference, there is obviously room for improvement, although we suspect that part of the difference between our model potential and the exact potential in the asymptotic region is due to the possible inaccuracy in the latter mentioned before. The model potential does clearly improve upon the Becke-Perdew potential. For the neon atom (figure 5.2) the asymptotic behaviour of the model potential appears to be quite good (assuming that the minimum of the "exact" potential between 5 and 6 bohr is incorrect). However, in the inner region of the neon

atom the model potential is not so well behaved. Although it does exhibit an intershell peak, the approximation of this feature to the true intershell peak is rather poor. As a matter of fact the Becke-Perdew potential gives a better approximation to the exact nonlocal potential precisely at the intershell peak. Nevertheless, considering the whole r range we can say that for both the beryllium and the neon atom our simple model potential gives a considerable improvement of the LDA potential, especially in the asymptotic region were the Becke-Perdew potential gives almost no correction.

We might now ask whether these improvements are reflected in the quality of the eigenvalues. In order to investigate this question we calculated for several atoms and molecules the eigenvalue of the highest occupied Kohn-Sham orbital. This eigenvalue should be equal to the ionisation energy of the system (or electron affinity for negative ions). We have done this for both LDA, LDA with the Becke-Perdew potential added and for LDA with our model potential ($\beta = 0.05$) added. The results for H and a number of atoms with noble gas configurations are presented in table 5.1. In this table we also compare with the self-interaction free eigenvalues of the Hartree-Fock approximation. From the table we can see that the LDA eigenvalues have a large discrepancy with experiment, with a mean absolute error of 5.40 eV. We can also see that these values are not improved by inclusion of the Becke-Perdew potential. The model potential on the other hand gives a considerable improvement compared to the LDA eigenvalues, with a mean absolute error of 0.56 eV. A nice feature of the model potential is also that it yields bound state solutions for the negative ions. We see that the improvements are not restricted to atoms but also occur for molecules. Table 5.2 shows results for the alkali and alkaline earth atoms and ions. The same quality is obtained as for the noble gas atoms and ions.

Does the improvement of the asymptotic behaviour achieved by the model potential, apart from showing up clearly in the one-electron energies, also have observable consequences for the density? In tables 5.2,5.3 and 5.4 we investigate a number of moments of orbital and total densities. Table 5.2 demonstrates that the radial extent of the highest occupied orbital in the alkali and alkaline earth atoms is significantly affected by the asymptotic correction introduced by the model potential. Maximum errors of the LDA potential amount to 9% and 18% for $\langle r \rangle$ and $\langle r^2 \rangle$ respectively of the alkalis and 6% and 13% for the alkaline earths. The first members of the two series, Li and Be, show comparitively small errors of 0.2% and 0.3% for Li and 1% and 2% for Be. The errors are much smaller for the ions, probably since they possess very tight closed shells. Although the effects are rather small for Be, making this atom perhaps not a good test case, we happen to have a very accurate CI density available for this atom and therefore we compare in table 5.3 for Be the moments of the total density as obtained from various calculations. Also the Hartree-Fock data are shown in this table, in order to see how much of the effect of correlation is taken into account by the LDA or model potential. Judging by this example, for $\langle r^p \rangle$ with p positive the LDA potential corrects the Hartree-Fock result in spite of the wrong asymptotic behaviour of this potential, the model potential gives improvement over LDA but in general only something like 50% of the difference between Hartree-Fock and exact is covered by the model potential. For $\langle r^{-1} \rangle$ and $\langle r^{-2} \rangle$ there is no improvement over the Hartree-Fock results.

For Neon (table 5.4) it is seen that for $\langle r^p \rangle$ with p positive the model potential considerably improves the LDA result. With respect to the difference with Hartree-Fock the picture is different from that for Be: the LDA potential does not correct the Hartree-Fock results but gives as expected a too diffuse density. For $\langle r^{-1} \rangle$ and $\langle r^{-2} \rangle$ there is again no improvement over the Hartree-Fock results. The LDA potential appears to give a too diffuse, the model potential a too contracted density in the inner region.

ATOM	LDA			MODEL			EXPT
	$< r >$	$< r^2 >$	ϵ	$< r >$	$< r^2 >$	ϵ	ϵ
Li $< 2s >$	3.822	17.345	0.12	3.815	17.410	0.19	0.20
Na $< 3s >$	3.995	18.753	0.11	3.703	16.256	0.21	0.19
K $< 4s >$	4.839	26.962	0.10	4.428	22.739	0.18	0.16
Rb $< 5s >$	5.125	30.041	0.09	4.664	25.036	0.18	0.15
Cs $< 6s >$	5.674	36.481	0.08	5.170	30.475	0.16	0.14
Be $< 2s >$	2.621	8.263	0.20	2.595	8.118	0.32	0.34
Mg $< 3s >$	3.137	11.572	0.17	2.985	10.528	0.29	0.28
Ca $< 4s >$	3.991	18.304	0.14	3.771	16.437	0.24	0.22
Sr $< 5s >$	4.339	21.468	0.13	4.079	19.066	0.23	0.21
Ba $< 6s >$	4.880	26.909	0.12	4.588	23.897	0.21	0.19
ION	LDA			MODEL			EXPT
	$< r >$	$< r^2 >$	ϵ	$< r >$	$< r^2 >$	ϵ	ϵ
$Li^+ < 1s >$	0.585	0.468	2.19	0.576	0.453	2.65	2.78
$Na^+ < 2p >$	0.803	0.839	1.34	0.795	0.821	1.70	1.74
$K^+ < 3p >$	1.428	2.421	0.92	1.430	2.428	1.15	1.16
$Rb^+ < 4p >$	1.720	3.428	0.80	1.713	3.404	1.01	1.00
$Cs^+ < 5p >$	2.087	4.959	0.69	2.078	4.920	0.88	0.85
$Be^{2+} < 1s >$	0.421	0.240	4.81	0.417	0.235	5.45	5.66
$Mg^{2+} < 2p >$	0.684	0.599	2.46	0.681	0.592	2.88	2.95
$Ca^{2+} < 3p >$	1.260	1.866	1.58	1.263	1.879	1.85	1.87
$Sr^{2+} < 4p >$	1.555	2.775	1.34	1.553	2.772	1.58	1.60
$Ba^{2+} < 5p >$	1.912	4.129	1.14	1.909	4.123	1.35	1.35

Table 5.2: expectation values of $< r >$ and $< r^2 >$ for the highest occupied orbitals with corresponding eigenvalue

	$< r^{-2} >$	$< r^{-1} >$	$< r >$	$< r^2 >$	$< r^3 >$	$< r^4 >$	$< r^5 >$
HF	57.618	8.409	6.129	17.319	63.151	270.656	1325.49
LDA	56.766	8.339	6.091	17.019	61.475	261.641	1276.385
MODEL	57.837	8.446	6.022	16.704	59.796	250.937	1198.763
EXACT	57.597	8.427	5.975	16.284	56.946	233.167	1085.87

Table 5.3: density moments for the beryllium atom

	$< r^{-2} >$	$< r^{-1} >$	$< r >$	$< r^2 >$	$< r^3 >$	$< r^4 >$	$< r^5 >$
HF	414.890	31.113	7.891	9.372			
LDA	411.916	30.998	8.016	9.853	15.967	32.500	80.692
MODEL	416.406	31.275	7.895	9.477	14.799	28.576	66.239
EXACT	414.967	31.110	7.935	9.545	14.941	29.006	67.904

Table 5.4: density moments for the neon atom

5.6 Summary and conclusions

In this work we have formulated a number of conditions for approximate exchange-correlation potentials. Potentials of gradient-corrected density functionals that are currently used with much success to calculate atomic exchange and correlation energies as well as bond energies of molecules do not obey some of these conditions. In order to make comparisons over all space to (almost) exact Kohn-Sham potentials we devised an iterative scheme to obtain the corresponding Kohn-Sham potential from a given density. The scheme is applicable to both atoms and molecules and is not limited to systems with few electrons. In this paper highly accurate CI densities for Be and Ne have been used to generate accurate Kohn-Sham potentials. Comparison to potentials derived from existing gradient-corrected functionals demonstrated that these potentials cover only a small part of the difference between the LDA potential and the exact one. This is at first sight a little surprising: how can we have large improvements in energies and almost no improvement in the potential? Several explanations may be advanced. First of all the approximate nonlocal functional might "oscillate" around the exact functional giving a good approximation of the energies but a bad approximation of its functional derivative. Other deficiencies are inherent to the derivation of the functionals. For instance both the nonlocal exchange-correlation functional of Langreth and Mehl [50] and the Perdew-Wang generalised gradient expansion [46] use a long-range cut-off of the exchange-correlation energy density. Langreth and Mehl use a low-k (large distance) cut-off in the momentum distribution of the exchange-correlation energy and Perdew and Wang perform a real-space cut-off in the exchange-correlation hole. This neglect of this asymptotic region reflects itself in the potential which thereby loses its asymptotic Coulombic behaviour. However the Becke functional shows that even a correct behaviour of the exchange hole potential does not garantee a good behaviour of the exchange potential.

A model potential has been presented which corrects some of the deficiencies of the current potentials, notably the asymptotic behaviour. The improved asymptotic behaviour shows up very clearly in the one-electron energies. The error in the highest occupied LDA eigenvalue (which should represent in the exact case the ionisation energy) is reduced by almost an order of magnitude by using the model potential. The model potential also corrects significantly the higher moments of the density, notably for diffuse outer orbitals. Because of this correction of the highest occupied orbitals, the model potential may be useful in the calculation of highly excited diffuse states and in general for density dependent properties such as dipole and quadrupole moments and derivatives thereof (infrared intensities).

In spite of some success, the present model potential is clearly deficient in some respects, notably the behaviour in the atomic intershell region. It will therefore require further improvement. This problem will be addressed in a subsequent paper of this series. We feel that the most significant success of the nonlocal corrections to LDA apply to bond energies of molecules. In particular the potential derived from the Becke gradient correction to the exchange energy density displays very interesting structure in molecules [90] that may explain the success of the Becke energy expression for bond energies. We feel that further modeling of exchange-correlation potentials should not only try to optimize the potential for atoms but should take into account the special effects of chemical bonding.

Chapter 6

Structure of the optimized effective Kohn-Sham exchange potential and its gradient approximations

An analysis of the structure of the optimized effective Kohn-Sham exchange potential v_x and its gradient approximations is presented. The potential is decomposed into the Slater potential v_S and the response of v_S to the density variations v_{resp}, the latter exhibits the distinct atomic shell structure. The peaks of v_{resp} are interpreted as a consequence of the Fermi hole localization. It is the approximation of v_{resp} that raises the main problem for the gradient approaches. A direct fit of v_x is made with the gradient-dependent Padé approximant form that possesses proper asymptotic and scaling properties and reproduces its shell structure.

6.1 Introduction

The optimized potential model (OPM) [34, 33, 91, 35] provides a direct way to calculate local potential $v(\mathbf{r})$ of the Kohn-Sham equations (Hartree atomic units will be used throughout the paper)

$$\left[-\frac{1}{2}\nabla^2 + v(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{6.1}$$

Within the OPM the optimized effective potential (OEP) $v(\mathbf{r})$ is defined as a potential of (6.1) with the eigenfunctions $\phi_i(\mathbf{r})$, whose Slater determinant ψ minimizes the expectation value of the total Hamiltonian H of a many-electron system

$$\langle\psi|\hat{H}|\psi\rangle \tag{6.2}$$

$v(\mathbf{r})$ can be subdivided in the usual way

$$v(\mathbf{r}) = v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_x(\mathbf{r}) \tag{6.3}$$

where $v_{ext}(\mathbf{r})$ is the external potential, $v_H(\mathbf{r})$ is the Hartree potential of electrons

$$v_H(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}' \tag{6.4}$$

with the electron density $\rho(\mathbf{r})$ built from the occupied orbitals $\phi_i(\mathbf{r})$

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 \tag{6.5}$$

and $v_x(\mathbf{r})$ is called the exchange optimized effective potential. The important features of atomic v_x such as its asymptotic behaviour at short and long distances from a nucleus [33, 57, 31, 35, 92, 49] and the shell structure [33, 50, 93, 35, 94] were established in the OPM.

The correct long-distance asymptotics was proved for the v_x of the OPM in the pioneer work of Talman and Shadwick [33]. On the other hand, an interpretation of the shell structure, the "peaks" of v_x in the atomic intershell regions remained an open problem. Another problem is finding approximations to v_x. Even for atomic systems the rigorous OPM calculations are rather involved [93, 35, 36], because they require the simultaneous self-consistent solution of the one-electron Kohn-Sham equations and an integral equation for v_x. Owing to this, it is desirable to develop an effective approximation to v_x within the density functional theory (DFT). Recent progress of the DFT is due to the generalized gradient approximation (GGA) [50, 51, 45, 95, 44, 43, 47, 96] that gives a remarkable improvement to the exchange-correlation energies of the local density approximation (LDA) [10, 11, 97, 98, 99]. It was shown [92, 49, 36, 94, 32], however, that the standard GGA exchange potentials do not possess the proper asymptotic behaviour and shell structure. The corresponding gradient corrections add little to the LDA exchange potential.

In this paper we analyze a structure of the exchange OEP v_x and its gradient approximations. In Section 6.2 properties of v_x are discussed. v_x is decomposed into the Fermi hole or Slater potential v_S and the response of v_S to the density variation v_{resp} and the peaks of atomic v_x are interpreted as a consequence of the Fermi hole localization. In Section 6.3 the atomic OEP and its components v_S and v_{resp} are compared with those of the LDA, GGA and the pioneer gradient expansion approximation (GEA) of Herman et al. [76, 100]. In Section 6.4 a gradient fit to v_x is presented, which possesses the proper asymptotic and scaling properties and reproduces the shell structure of the OEP. The results of atomic calculations with the fitted potentials are compared with those of the OPM, LDA, GGA and GEA. The comparison is made for the form of potentials, orbital eigenvalues and also for the exchange energies, the latter have been calculated for the fitted potential via the Levy-Perdew relation [22].

6.2 Properties of the exchange optimized effective potential

The exchange OEP $v_x(\mathbf{r})$ possesses the following scaling property [22, 23]

$$v_x([\rho_\lambda]; \mathbf{r}) = \lambda v_x([\rho]; \lambda\mathbf{r}) \tag{6.6}$$

where $\rho_\lambda(\mathbf{r}) = \lambda^3 \rho(\lambda\mathbf{r})$. For finite systems it has the asymptotic form [33, 57, 31, 35]

$$v_x(\mathbf{r}) \rightarrow -\frac{1}{|\mathbf{r}|} \qquad (|\mathbf{r}| \rightarrow \infty) \tag{6.7}$$

Using 6.1,6.2 the OPM provides a radial integral equation [33, 91] for the atomic OEP $v(r)$ (r is the radial coordinate), which can be solved numerically. Figure 1 illustrates the typical features of the OEP exchange part v_x for Kr extracted from [91]. One can see from this Figure the finiteness of v_x at the nucleus

$$v_x(r) = \text{const} \qquad (r = 0) \tag{6.8}$$

Figure 6.1: Exchange optimized effective potential for Kr

and its shell structure, i.e. the non-monotonous dependence of v_x on r with the local maxima ("peaks") in the regions between the atomic shells.

In order to trace the origin of the intershell peaks of v_x, we use its definition as the functional derivative of the exchange energy $E_x[\rho]$

$$v_x(\mathbf{r}) = \frac{\delta E_x[\rho]}{\delta \rho(\mathbf{r})} \tag{6.9}$$

that holds true in the OPM and express the exchange part of 6.2 as an integral of the Fermi hole or Slater potential v_S [101]

$$E_x = \frac{1}{2} \int \rho(\mathbf{r}) v_S(\mathbf{r}) d\mathbf{r} \tag{6.10}$$

$$v_S(\mathbf{r}) = - \int \frac{f(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \tag{6.11}$$

In equation 6.11 $f(\mathbf{r}, \mathbf{r}')$ is the Fermi hole density [101, 102] built from the occupied Kohn-Sham orbitals $\phi_i(\mathbf{r})$

$$f(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r})^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \phi_i(\mathbf{r}) \phi_i^*(\mathbf{r}') \phi_j^*(\mathbf{r}) \phi_j(\mathbf{r}') \tag{6.12}$$

The expression 6.10 formally looks like an energy of the electronic interaction with an external potential. In reality, being an internal potential of the electronic system, v_S has the non-zero

response v_{resp} to density variations and, according to equations 6.9-6.11, v_x can be expressed as the arithmetical mean of v_S and its response v_{resp}

$$v_x(\mathbf{r}) = \frac{1}{2}[v_S(\mathbf{r}) + v_{resp}(\mathbf{r})] \tag{6.13}$$

$$v_{resp}(\mathbf{r}) = \int \rho(\mathbf{r}')\frac{\delta v_s(\mathbf{r}')}{\delta \rho(\mathbf{r})}d\mathbf{r}' = -\int \rho(\mathbf{r}')d\mathbf{r}' \int \frac{\delta f(\mathbf{r}',\mathbf{r}'')}{\delta \rho(\mathbf{r})}\frac{d\mathbf{r}''}{|\mathbf{r}'-\mathbf{r}''|} \tag{6.14}$$

with $\frac{\delta v_s(\mathbf{r}')}{\delta \rho(\mathbf{r})}$ and $\frac{\delta f(\mathbf{r}',\mathbf{r}'')}{\delta \rho(\mathbf{r})}$ being the linear response functions [6].

In Figure 2 OEP v_x [91] and its components $\frac{1}{2}v_S$ and $\frac{1}{2}v_{resp}$ are plotted for *Ca*, *Kr* and *Cd*. The relative contribution of v_{resp} to v_x ranges from 0.5, its asymptotic value at long distances r, and from ~ 0.35 within the shell regions to the near zero values in the intershell regions. In all cases v_S is a monotonous function of r, while v_{resp} possesses the distinct peaks in the intershell regions and, hence, the shell structure of v_x originates mainly from v_{resp}.

From Figures 2a-c and eq.6.14 the shell structure can be interpreted in terms of the different response of the Fermi hole density $f(\mathbf{r}_2,\mathbf{r}'')$ to the perturbations $\delta\rho(\mathbf{r})$ of the electron density in the shell and intershell regions. At the intershell points v_x is closer to $\frac{1}{2}v_S$ and v_S looks in fact like an external potential which has, by definition, zero response to the density perturbations. In the sense of integral contribution to 6.14, intershell points appear as the points outside the area of the Fermi hole 6.12, which determines v_{resp}. But this is true, because, owing to an exchange effect, the Fermi repulsion of electrons with like spins, for various pairs $\mathbf{r}'$ and $\mathbf{r}''$ the Fermi hole $f(\mathbf{r}',\mathbf{r}'')$ is mainly localized within the shell regions [102, 103]. One can expect (and we can see it from Figure 2) the smaller response of $f(\mathbf{r}',\mathbf{r}'')$ when $\rho(\mathbf{r})$ is perturbed somewhere outside its localization region, in particular, at the intershell points. To sum up, the intershell peaks of v_x can be interpreted as a consequence of localization of the Fermi hole and successful approximations to v_x should reflect this feature.

It is interesting to note, that the shell structure of the exchange potential is only the secondary effect of the Fermi repulsion. The main effect is contained in the Pauli potential [104], which appears as a kinetic contribution to the total potential of the Euler equation for $\rho^{\frac{1}{2}}(\mathbf{r})$ and has much more pronounced intershell peaks [105, 106, 107]. The analysis of the shell structures in the Euler equation will be given elsewhere [108].

6.3 Comparison of the OPM, LDA, GGA and GEA exchange potentials

In the recent papers [92, 49, 35, 94] OEP v_x has already been compared with various DFT approximations. In this paper the comparison is made not only for v_x, but also for its components v_S and v_{resp}. All the standard DFT schemes approximate the exchange energy density $\epsilon_x(\rho;\mathbf{r})$

$$E_x^{approx}[\rho] = \int \epsilon_x^{approx}([\rho];\mathbf{r})d\mathbf{r} \tag{6.15}$$

and, according to equations 6.10 and 6.15, the corresponding approximation to v_S can be simply related to ϵ_x^{approx}

$$v_S^{approx}(\mathbf{r}) = \frac{2\epsilon_x^{approx}([\rho];\mathbf{r})}{\rho(\mathbf{r})} \tag{6.16}$$

Figure 6.2: Exchange optimized effective potential and its components for a) Ca, b) Kr and c) Cd

Figure 6.3: Exchange potentials of OPM, LDA, GEA and GGA for a) Ca, b) Kr and c) Cd

while v_{resp}^{approx} can be obtained by the functional differentiation of equation 6.16.
In the LDA all the potentials v_x^{LDA}, v_S^{LDA} and v_{resp}^{LDA} have the same form

$$v_x^{LDA}(\mathbf{r}) = -2(\frac{3}{8\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r}); \ v_S^{LDA}(\mathbf{r}) = -3(\frac{3}{8\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r});$$

$$v_{resp}^{LDA}(\mathbf{r}) = -(\frac{3}{8\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r}) \tag{6.17}$$

In the GEA [76, 100] v_S^{LDA} is modified with the density gradient correction

$$v_S^{GEA}(\mathbf{r}) = v_S^{LDA}(\mathbf{r}) - 0.015(\frac{3}{4\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r})\xi = v_S^{LDA}(\mathbf{r}) + \tilde{v}_S^{GEA}(\mathbf{r}) \tag{6.18}$$

$$\xi = 2^{\frac{2}{3}}(\frac{\nabla\rho(\mathbf{r})}{\rho^{\frac{4}{3}}(\mathbf{r})})^2 \tag{6.19}$$

and in v_{resp}^{GEA} the additional Laplacian term appears

$$v_{resp}^{GEA}(\mathbf{r}) = v_{resp}^{LDA}(\mathbf{r}) - 0.005(\frac{3}{4\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r})(\xi - 6\eta) = v_{resp}^{LDA}(\mathbf{r}) + \tilde{v}_{resp}^{GEA}(\mathbf{r}) \tag{6.20}$$

$$\eta = 2^{\frac{2}{3}}\frac{\nabla^2\rho(\mathbf{r})}{\rho^{\frac{5}{3}}(\mathbf{r})} \tag{6.21}$$

$$v_x^{GEA}(\mathbf{r}) = v_x^{LDA}(\mathbf{r}) - 0.01(\frac{3}{4\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r})(\xi - \frac{3}{2}\eta) = v_x^{LDA}(\mathbf{r}) + \tilde{v}_x^{GEA}(\mathbf{r}) \tag{6.22}$$

All the gradient corrections in eqs.6.18,6.20,6.22 contain the empirical parameter determined in [76, 100].
In the GGA [44] v_S^{GEA} is further modified with the correction function $f(\xi)$ in order to provide the correct long-distance asymptotics of v_S^{GGA}

$$v_S^{GGA}(\mathbf{r}) = v_S^{LDA}(\mathbf{r}) + f(\xi)\tilde{v}_S^{GEA}(\mathbf{r}) \tag{6.23}$$

$$f(\xi) = \frac{0.56(\frac{4\pi}{3})^{\frac{1}{3}}}{1 + 0.0252\xi^{\frac{1}{2}}\sinh^{-1}\xi^{\frac{1}{2}}} \tag{6.24}$$

and in v_{resp}^{GGA} the function $f(\xi)$ appears together with its first and second derivatives

$$v_{resp}^{GGA}(\mathbf{r}) = v_{resp}^{LDA}(\mathbf{r}) + f(\xi)\tilde{v}_{resp}^{GEA}(\mathbf{r}) - 0.03(\frac{3}{4\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r})[\frac{df}{d\xi}(\xi\eta + 2\tau) - \frac{d^2f}{d\xi^2}\xi\tau] \tag{6.25}$$

$$v_x^{GGA}(\mathbf{r}) = v_x^{LDA}(\mathbf{r}) + f(\xi)\tilde{v}_x^{GEA}(\mathbf{r}) - 0.015(\frac{3}{4\pi})^{\frac{1}{3}}\rho^{\frac{1}{3}}(\mathbf{r})[\frac{3}{2}\frac{df}{d\xi}(\xi\eta + 2\tau) - \frac{3}{2}\frac{d^2f}{d\xi^2}\xi\tau] \tag{6.26}$$

$$\tau = 2^{\frac{4}{3}}\frac{\nabla\rho(\mathbf{r})\nabla\xi(\mathbf{r})}{\rho^{\frac{5}{3}}(\mathbf{r})} \tag{6.27}$$

Eq. 6.24 contains an empirical parameter determined in [44]. Being the functional derivatives of the corresponding exchange energy functionals, v_x^{LDA}, v_x^{GEA}, v_x^{GGA} and also their components have the proper scaling property 6.6.

Figure 6.4: Slater potentials of OPM, LDA, GEA and GGA for a) Ca, b) Kr and c) Cd

Figure 6.5: Response potentials of OPM, LDA, GEA and GGA for a) Ca, b) Kr and c) Cd

In Figures 3-5 OEP v_x [91] and its components v_S and v_{resp} are compared for Ca, Kr and Cd with the corresponding potentials of LDA, GEA and GGA. The potentials were calculated with the OPM densities ρ. All the approximate total exchange potentials fail to reproduce the correct asymptotics 6.7 of the OEP. At long distances v_x^{LDA} falls off exponentially, v_x^{GEA} diverges (See Figure 3b where the potentials for Kr are represented in logarithmic scale for larger distances), v_x^{GGA} decreases similarly to v_x^{LDA} and only at very long distances it is proportional to $-\frac{1}{r^2}$ [49]. At the nucleus v_x^{LDA} is finite, while the gradient potentials v_x^{GEA} and v_x^{GGA} diverge, being proportional to $-\frac{1}{r}$ [92, 32].

Both LDA and GGA potentials do not possess the distinct shell structure. v_x^{LDA} is an everywhere monotonous function of r and v_x^{GGA} exhibits a peak only for the outermost intershell regions of Ca and Cd, being a monotonous function of r for the rest of the space. Contrary to this, v_x^{GEA} possesses clear shell structure. Amplitudes of the peaks obtained within the GEA differ from those of the OPM, but the locations of the peaks are nearly the same as in the OPM (See Figure 3a-c). The abovementioned discrepancies between v_x and v_x^{approx} originate mainly from those between v_{resp} and v_{resp}^{approx}, while the correspondence between v_S and v_S^{approx} is definitely better. All considered approximations to v_S are finite at the nucleus and approximate rather closely v_S at short and intermediate distances r (See Figure 4). In particular, the GGA provides especially good approximation for v_S. By the construction [44], v_S^{GGA} has the correct asymptotics 6.7 for $r \to \infty$ and it is very close to v_S for all distances. Hence, the well-known success of the GGA in calculation of the integral characteristics, such as E_x, can be attributed to its locally very good approximation to v_S and, which is the same, to the exchange energy density ϵ_x. v_S^{LDA} and v_S^{GEA} are not so good: v_x^{LDA} falls of exponentially and v_S^{GEA} diverges at long distances and both potentials exhibit appreciable discrepancies with v_S at $r \sim 2 a.u.$

However, all the approaches show more pronounced deficiencies in approximating v_{resp} (See Figure 5). Due to the presence of the unbalanced Laplacian contributions to v_{resp}^{approx}, both GEA and GGA diverge at the nucleus, which causes the abovementioned divergence of the corresponding total exchange potentials. All the approaches fail to reproduce the correct asymptotics $v_{resp}(r) \to -\frac{1}{r}$ for $r \to \infty$: v_{resp}^{LDA} decreases exponentially, v_{resp}^{GEA} diverges and v_{resp}^{GGA} yields the positive asymptotics, $v_{resp}^{GGA}(r) \to \frac{1}{r}$, that cancels the correct asymptotics of v_S^{GGA} and produces the incorrect asymptotics of the total exchange GGA potential v_x^{GGA}. In the intershell regions v_{resp}^{LDA} is a monotonous function of r, v_{resp}^{GGA} has the peaks, but for all intershell regions, with the exception of the outermost one, they are much smaller than those of the OPM. Contrary to the LDA and GGA, GEA reproduces the peaks of v_{resp}, but with substantial deviations of amplitudes for v_{resp}^{GEA} and v_{resp} peaks.

One can see from the above analysis and discussion, that the methods to obtain v_x as functional derivative of the current approximate expressions for E_x do not meet the high demand to provide satisfactory approximation not only for v_S, but also for its functional derivative. The alternative way of a direct fitting of v_x to a gradient function with a due account of the accurate asymptotic and scaling properties is presented in the next section.

6.4   A gradient fit to v_x

To fit v_x, we use the Padé approximant approach, which has been used previously for the fitting of ϵ_x [49, 109, 85, 110, 94]. The following Padé approximant form is considered as a fitting potential v_x^f

$$v_x^f(\mathbf{r}) = v_x^{LDA}(\mathbf{r}) - 2^{\frac{1}{3}} \rho^{\frac{1}{3}}(\mathbf{r})[P_{n+1}^{n(\xi)}(\xi^{\frac{1}{2}} \sinh^{-1} \xi^{\frac{1}{2}})\xi - P_{m+1}^{m(\eta)}(|\eta|)\eta] \tag{6.28}$$

where $P_l^{k(y)}(x)$ is a $[k/l]$-Padé function of degrees k and l [111]

$$P_l^{k(y)}(x) = \frac{a_0^y + a_1^y x + ... + a_k^y x^k}{1 + b_1^y x + ... + b_l^y x^l} \tag{6.29}$$

The simplest Padé function $P_{n+1}^{n(\xi)}(\xi^{\frac{1}{2}} \sinh^{-1} \xi^{\frac{1}{2}})$ with $n = 0$ is the Becke's $f(\xi)$ (eq.6.24), which has also been used in [32] as a model of the exchange-correlation potential. The whole expression 6.28 can be considered as the Padé approximant modification of the GEA exchange potential 6.22. It appears that with the following restriction for the coefficients a_n^ξ and b_{n+1}^ξ

$$b_{n+1}^\xi = 3a_n^\xi \tag{6.30}$$

the Padé approximant form eqn.6.28 of arbitrary orders n and m possesses the scaling property 6.6, accurate long-distance asymptotics 6.7 and finiteness 6.8 at the nucleus.

The forms 6.28 with $n, m = 0 - 2$ have been used to fit v_x with the coefficients determined from the least squares minimization of the weighted difference $[v_x(r) - v_x^f(r)]$ for the closed-shell atoms $Be, Ne, Mg, Ar, Ca, Zn, Kr, Cd$

$$\sum_A \sum_i \{w^A(r_i)[v_x(r_i) - v_x^f(r_i)]\}^2 = \min \tag{6.31}$$

In equation 6.31 the inner summation is over the radial mesh points r_i, the outer one is over atoms and weighting functions $w^A(r_i)$ of the following form have been chosen

$$w_1^A(r_i) = \frac{p_i r_i^2 \rho^A(r_i)}{E_x^A} \tag{6.32}$$

$$w_2^A(r_i) = \frac{p_i r_i^2 [3\rho^A(r_i) + r \frac{d\rho^A}{dr}|_{r=r_i}]}{E_x^A} \tag{6.33}$$

where p_i are the weighting parameters. The radial densities are inserted in equation 6.32 to provide weights to the points r_i according to their contributions to the sum of the orbital energies of the Kohn-Sham equations. The weights in eqn.6.33 represent the corresponding contributions to the atomic exchange energy E_x^A, which can be calculated for a given v_x via the Levy-Perdew relation [22]

$$E_x = \int [3\rho(\mathbf{r}) + \mathbf{r}\nabla\rho(\mathbf{r})]v_x(\mathbf{r})d\mathbf{r} \tag{6.34}$$

Energies E_x^A are inserted in equations 6.32, 6.33 in order to make an uniformly representative functional for both light and heavy atoms.

The minimized value of the functional (eqn. 6.31) decreases considerably with increasing of degrees n and m of the fit of eqn. 6.28 from the least $n = 0$, $m = 0$ to $n = 0$, $m = 1$, but higher degrees appear to produce little effect. As a result, the following simple particular case of the form of eqn. 6.28 has been chosen as a fitting potential

$$v_x^f(\mathbf{r}) = v_x^{LDA}(\mathbf{r}) - 2^{\frac{1}{3}} \rho^{\frac{1}{3}}(\mathbf{r})[\frac{a_0^\xi \xi}{1 + 3a_0^\xi \xi^{\frac{1}{2}} \sinh^{-1} \xi^{\frac{1}{2}}} - \frac{(a_0^\eta + a_1^\eta \eta)\eta}{1 + b_2^\eta \eta^2}] \tag{6.35}$$

The optimized form of v_x^f, naturally, depends essentially on the actual weighting function $w^A(r_i)$ chosen. To illustrate this, in Figure 6 the non-local potential correction v_{nl} of the OPM (a difference between v_x and v_x^{LDA}) is compared for Ca, Kr and Cd with the corresponding corrections for

ORBITAL	OPM	v_x^{f1}	v_x^{f2}	v_{xc}^{f}	GGA	LDA
1s	144.50	144.60	145.11	146.32	144.38	143.85
2s	15.26	15.20	15.36	15.68	15.05	14.98
2p	12.43	12.37	12.54	12.89	12.26	12.22
3s	1.77	1.73	1.77	1.85	1.67	1.66
3p	1.08	1.05	1.09	1.18	0.99	0.98
4s	0.195	0.144	0.165	0.206	0.116	0.111

Table 6.1: Orbital energies for Ca (with the opposite sign and in a.u.) calculated within OPM, LDA, GGA and with the fitting potentials v_x^{f1}, v_x^{f2} and v_{xc}^{f}

the GEA, GGA and for two fits v_x^{f1} and v_x^{f2} of the form of eqn. 6.35 with the different sets of the optimized coefficients: $a_0^{\xi 1} = 0.0123$, $a_0^{\eta 1} = 0.0087$, $a_1^{\eta 1} = 0.0004$, $b_1^{\eta 1} = 0.0011$ for v_x^{f1} and $a_0^{\xi 2} = 0.0204$, $a_0^{\eta 2} = 0.0086$, $a_1^{\eta 2} = 0.0005$, $b_1^{\eta 2} = 0.0022$ for v_x^{f2}. For both fits the radial mesh $\{r_i\}$ of Numerov method [112] with the adaptive step size has been used. v_x^{f1} has been optimized for the weighting function $w_1^A(r_i)$ (eqn. 6.32) with all the weighting parameters p_i set equal, i.e. $p_i = 1$, while v_x^{f2} has been optimized for the weighting function $w_2^A(r_i)$ (eqn. 6.33) with the mesh weights of the Numerov quadrature being used as weighting parameters p_i. Due to a large concentration of the mesh points near the nucleus, the function $w_1^A(r_i)$ provides the greater contribution of the inner radial points in equation 6.31 as compared to $w_2^A(r_i)$. As a result, v_{nl}^{f1} reproduces the shell structure of the OPM potential better than any other potential presented (See Figure 4) and is close to v_{nl} in the core region, while v_{nl}^{f2} goes closer to v_{nl} in the valence region.

It is interesting to note that, as regards to a_0^{ξ} value, v_x^{f2} is placed in between v_x^{f1} and the model exchange-correlation potential v_{xc}^{f} of [32]. The latter is a particular case of eqn.6.35 with $a_0^{\eta} = a_1^{\eta} = b_2^{\eta} = 0$ and with $a_0^{\xi} = 0.05$ chosen to reproduce the first ionisation potential of Be atom. As it was shown in [32], v_{xc}^{f} yields a good estimate of ionisation energies for various atomic and molecular systems and it goes close to the accurate Kohn-Sham potential in the outer valence regions.

The abovementioned features of the potentials are reflected in Tables 1-3 where the orbital energies ϵ_i of eq.6.1 calculated with various exchange potentials are presented for Ca, Kr and Cd. LDA underestimates electron exchange and its orbital energies are smaller (in absolute magnitude) than those of the OPM. GGA yields energies, which are close to the LDA ones for valence and outer core orbitals, and only for the deep core orbitals the GGA energies grow definitely greater. In all cases (with the only exceptions of 2s- and 2p- orbitals of Cd) v_x^{f1} provides an improvement to the GGA orbital energies. Typical ϵ_i errors vary within $0.05 - 0.15$ a.u. for v_x^{f1}, which means small relative errors for the core orbital energies and quite appreciable errors for the valence ones. v_x^{f2} provides further improvement for the valence and outer core energies, though its deep core orbital energies are worse than those of v_x^{f1}. v_{xc}^{f} gives the best energies of the highest occupied orbitals for Ca and Kr, while the deep core energies are definitely overestimated. For Cd the highest orbital energy for v_{xc}^{f} deviates more from OPM than v_x^{f2} does, but it is quite close to its accurate ionisation energy 0.330 a.u. v_{xc}^{f}, being an approximation to the exchange-correlation rather than exchange-only potential, should indeed yield the ionisation energy. Note that, by the construction (eqn.6.35), all fitted potentials have the accurate long-distance asymptotics (eqn.6.7). One can see, that this property by itself cannot guarantee high quality of the calculated energies of the highest occupied orbitals that depends critically on the v_x^{f} behaviour at shorter distances in the outer valence region.

In Table 6.4 the non-local corrections E_x^{nl} of the OPM to the LDA exchange energies are compared for the closed-shell atoms Be, Ne, Mg, Ar, Ca, Zn, Kr, Cd with those of the approximate schemes.

Figure 6.6: Non-local corrections of OPM, GEA, GGA and of the fitting potentials to the LDA exchange potential for a) Ca, b) Kr and c) Cd

Figure 6.7: Dimensionless arguments ξ and η for Kr

ORBITAL	OPM	v_x^{f1}	v_x^{f2}	v_{xc}^f	GGA	LDA
1s	511.09	511.21	512.08	514.22	510.87	509.89
2s	66.64	66.52	66.76	67.32	66.35	66.21
2p	60.27	60.17	60.42	61.05	60.02	59.94
3s	9.65	9.41	9.50	9.68	9.30	9.25
3p	7.37	7.15	7.24	7.44	7.05	7.02
3d	3.30	3.10	3.20	3.41	3.02	3.01
4s	0.97	0.83	0.87	0.96	0.78	0.77
4p	0.501	0.358	0.399	0.480	0.305	0.300

Table 6.2: Orbital energies for Kr (with the opposite sign and in a.u.) calculated within OPM, LDA, GGA and with the fitting potentials v_x^{f1}, v_x^{f2} and v_{xc}^f

ORBITAL	OPM	v_x^{f1}	v_x^{f2}	v_{xc}^f	GGA	LDA
1s	943.01	943.10	944.24	947.13	942.70	941.38
2s	137.00	137.12	137.41	138.18	136.94	136.76
2p	127.68	127.82	128.14	129.00	127.67	127.56
3s	25.56	25.49	25.65	25.98	25.35	25.31
3p	21.76	21.71	21.86	22.21	21.61	21.57
3d	14.75	14.71	14.86	15.21	14.64	14.62
4s	3.70	3.61	3.65	3.74	3.56	3.54
4p	2.50	2.40	2.44	2.54	2.35	2.34
4d	0.56	0.47	0.52	0.63	0.43	0.42
5s	0.267	0.208	0.250	0.321	0.170	0.168

Table 6.3: Orbital energies for Cd (with the opposite sign and in a.u.) calculated within OPM, LDA, GGA and with the fitting potentials v_x^{f1}, v_x^{f2} and v_{xc}^f

ATOM	OPM	GGA	v_x^{f1}	v_x^{f2}	v_{xc}^f	GEA
Be	0.360	0.345	0.175	0.193	0.367	0.826
Ne	1.121	1.102	0.908	1.183	2.147	2.275
Mg	1.402	1.387	1.233	1.611	2.921	2.766
Ar	2.339	2.287	2.023	2.663	5.034	4.409
Ca	2.614	2.597	2.457	3.238	6.047	4.915
Zn	4.112	4.215	4.651	6.464	12.193	7.611
Kr	5.274	5.241	5.649	8.024	15.760	9.351
Cd	7.419	7.376	8.400	11.776	22.758	12.859

Table 6.4: Non-local corrections to the LDA exchange energies (with the opposite sign and in a.u.) calculated within OPM, GGA, GEA and via the Levy-Perdew relation with the fitting potentials v_x^{f1}, v_x^{f2} and v_{xc}^f

For the fitted potentials E_x^{nl} are calculated from the Levy-Perdew relation (eqn.6.34) and for GGA and GEA they are calculated directly from the corresponding energy expressions (eqn.6.15).

Though non-variational, v_x^{f1} yields a reasonable estimation of E_x^{nl} for the atoms considered. Its quality is worse, than the highly superior quality of the GGA [44] energies, but better than that of GEA [76, 100]. v_x^{f2} also yields somewhat better E_x^{nl} values than the GEA, though in this case for all atoms heavier than Ne the calculated E_x^{nl} values are substantially overestimated. v_{xc}^f provides the largest overestimation of E_x^{nl} for Mg and heavier atoms. A comparison of the results for fitted potentials shows, not unexpectedly, that, in order to provide a reasonable estimate for the total exchange energy, it is more important to have a better fit for the inner region than for the outer one.

A natural question arises whether it possible to construct a potential that would combine the advantages of v_x^{f1}, v_x^{f2} and v_{xc}^f and would provide a close fit to v_x at all distances. Our experience gives a negative answer to this question for the case of the Padé approximant form (eqn. 6.28). The reason for this is the concerted oscillatory behaviour of the dimensionless arguments ξ (eq. 6.19) and η (eq. 6.21) as functions of r (See Figure 7 for Kr as a typical example). ξ and η are the natural parameters of the gradient expansion for ϵ_x and v_x and their utilization ensures the accurate scaling property (eqn. 6.6) for various gradient approximations. However, because of their oscillations, there always exist core and valence regions where both ξ and η have a similar behaviour, so that

the gradient approximations always produce concerted, related to each other non-local corrections for these regions. Due to this, the local flexibility of the fit (eqn. 6.28) is substantially restricted and better fitting for the inner part of v_x produces at the same time worse fitting for the outer part and vice versa. As a result, the increase of a_0^ξ and removal of the η-dependent part of (eqn.6.35), when passing on from v_x^{f1} through v_x^{f2} to v_{xc}^f, brings better energies of the highest occupied orbitals but worse E_x^{nl} values.

From the above discussion v_x^{f1} seems to be the more optimal fit to v_x as compared with v_x^{f2}. The former provides systematical improvement for the orbital energies of the GGA, while keeping reasonable estimation of the exchange energies and it can be used in the Kohn-Sham equations as a simple approximation to v_x and also as a starting point for more refined approximations. For such a refinement one can consider (in addition to the standard ξ and η) other gradient parameters that should have distinctly different behaviour in the core and valence regions. Using these parameters for the interpolation between v_x^{f1} in inner regions and v_{xc}^f in the outer ones, one can hope to construct uniformly close fit to v_x, or indeed to v_{xc} if required. The corresponding work is in progress.

6.5   Conclusions

In this paper an analysis of the structure of the optimized effective Kohn-Sham exchange potential and its gradient approximations has been presented. The potential has been decomposed into the Slater potential and its response to the density variations, the latter exhibits the distinct atomic shell structure. The intershell peaks of the potential originate from the smaller response of the Fermi hole density, which is localized in the shell regions, to the density perturbations at the intershell points. It has been shown, that the approximation of the response potential is a major problem for the gradient approach. A direct fit of the potential has been made with the the Padé approximant form that possesses proper asymptotic and scaling properties and reproduces shell structure of the OEP.

Chapter 7

Analysis of electron interaction and atomic shell structure in terms of local potentials

The Kohn-Sham potential v_s of an N-electron system and the potential v_{eff} of the Euler-Lagrange equation for the square root of the electron density are expressed as the sum of the external potential plus potentials related to the electronic structure, such as the potential of the electron Coulomb repulsion, including the Hartree potential and the screening due to exchange and correlation, a potential representing the effect of Fermi-Dirac statistics and Coulomb correlation on the kinetic functional, and additional potentials representing "response" effects on these potentials. For atoms several of these potentials have distinct atomic shell structure: one of them has peaks between the shells, while two others are step functions. In one of those step functions the steps represent characteristic shell energies. Examples of the potentials extracted from the optimized potential model (OPM) are presented for Kr and Cd. Correlation potentials, obtained by subtracting the exchange potential of the OPM from (nearly) exact Kohn-Sham potentials, are discussed for Be and Ne.

7.1   Introduction

Local one-electron potentials are key ingredients of density functional theory (DFT) [5] The potential $v_s(\mathbf{r})$ of the familiar Kohn-Sham one-electron equations [11] (Hartree atomic units will be used throughout the paper)

$$(-\frac{1}{2}\nabla^2 + v_s([\rho];\mathbf{r}))\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{7.1}$$

and the effective potential $v_{eff}([\rho];\mathbf{r})$ of the Euler-Lagrange equation for the square root of the electron density [113, 114, 115, 116, 107]

$$(-\frac{1}{2}\nabla^2 + v_{eff}([\rho];\mathbf{r}))\sqrt{\rho(\mathbf{r})} = \mu\sqrt{\rho(\mathbf{r})} \tag{7.2}$$

accumulate various effects of the electron interaction in many-electron systems. They provide an informative picture of the electron interaction in terms of potential barriers and wells in physical

space. Evidently, the complete effect of electron correlation has been folded into these local potentials. In this paper we will analyze the relationship between the potentials and electron correlation. This analysis leads to a decomposition of the potentials in various components, which have been calculated within the (exchange only) optimized potential model [33] for the atoms Kr and Cd. For Be and Ne the correlation part of the exchange-correlation potential has been obtained from nearly exact v_s corresponding to the electron density as given by highly correlated wavefunctions. A distinct structure (peaks, steps) is observed in the various potentials, that can be rationalized in terms of aspects of the electron correlation. The observed structure will be very helpful when one tries to model the potentials accurately and efficiently by judiciously chosen functionals of the density and its derivatives.

7.2   Definition of local potentials

To analyze the structure of v_s and v_{eff}, we begin with the electronic energy expression of the constrained-search DFT [18] for a trial density ρ

$$E_v[\rho] = \int \rho(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r} + \min_{\psi \to \rho} \langle \psi | \hat{T} + \hat{W} | \psi \rangle \tag{7.3}$$

where v_{ext} is the external (usually, Coulombic) potential, $\hat{T}$ and $\hat{W}$ are operators of the kinetic energy and electron repulsion and the minimum is searched over all antisymmetric normalized wavefunctions ψ, which yield ρ. One can express E in terms of local potentials

$$E_v[\rho] = T_W[\rho] + (T[\rho] - T_W[\rho]) + V_{ext}[\rho] + W[\rho] = -\frac{1}{2} \int \sqrt{\rho(\mathbf{r})} \nabla^2 \sqrt{\rho(\mathbf{r})} d\mathbf{r} +$$

$$\int \rho(\mathbf{r}) v_{kin}([\rho]; \mathbf{r}) d\mathbf{r} + \int \rho(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \rho(\mathbf{r}) v_{cond}([\rho]; \mathbf{r}) d\mathbf{r} \tag{7.4}$$

Here $T_W[\rho]$ is the von Weizsäcker kinetic energy functional which is N times the kinetic energy of the normalized "density orbital" $\sqrt{\rho/N}$, and the explicit effect of the Fermi statistics and electron interaction on the true kinetic functional $T[\rho]$, i.e. the difference $T[\rho] - T_W[\rho]$, is represented with the local potential v_{kin}. According to [86], the latter potential can be defined via the conditional probability amplitude Φ [113]

$$\Phi(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1) = \frac{\psi_\rho(\vec{x}_1, ..., \vec{x}_N)}{\sqrt{\rho(\mathbf{r}_1)/N}} \tag{7.5}$$

of the (non-degenerate) function ψ_ρ, which minimizes the energy functional (7.3) for the prescribed density ρ ($\{\vec{x}_i\} = \{\vec{r}_i, s_i\}$, $\{\vec{r}_i\}$ are the space and $\{s_i\}$ are the spin variables)

$$v_{kin}([\rho]; \mathbf{r}_1) = \frac{1}{2} \int |\nabla_1 \Phi(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)|^2 ds_1 d\vec{x}_2 ... d\vec{x}_N . \tag{7.6}$$

$\Phi(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)$ embodies all effects of electron correlation (exchange as well as Coulomb) in that its square is the probability distribution of the remaining $N - 1$ electrons associated with positions $\vec{x}_2 ... \vec{x}_N$ when one electron is known to be at $\mathbf{r}_1$. v_{kin} can be interpreted as a measure of how strongly the motion of the reference electron at $\mathbf{r}_1$ is correlated with the other electrons in the system, in the sense that it reflects the magnitude of change in Φ with changing $\mathbf{r}_1$.

v_{cond} describes the effective Coulomb repulsion of the reference electron by the other electrons. It is expressed through the conditional density [86]

$$\rho_{cond}([\rho]; \mathbf{r}_1, \mathbf{r}_2) = g([\rho]; \mathbf{r}_1, \mathbf{r}_2)\rho(\mathbf{r}_2) =$$

$$= (N-1) \int \Phi^*(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)\Phi(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)ds_1 ds_2 d\vec{x}_3...d\vec{x}_N \tag{7.7}$$

$$v_{cond}([\rho]; \mathbf{r}_1) = \int \rho(\mathbf{r}_2)\frac{g([\rho]; \mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_2 \tag{7.8}$$

and can be considered as the potential of a interparticle Coulomb interaction screened due to exchange-correlation effects, with the pair-correlation function $g(\mathbf{r}_1, \mathbf{r}_2)$ being the screening factor. Evidently, v_{cond} can be broken up into the Hartree potential and a screening potential due to the exchange-correlation hole density $\rho(\mathbf{r}_2)(g(\mathbf{r}_1, \mathbf{r}_2)-1)$ surrounding the reference electron at $\mathbf{r}_1$:

$$v_{cond}([\rho]; \mathbf{r}_1) = \int \frac{\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_2 + \int \frac{\rho(\mathbf{r}_2)(g([\rho]; \mathbf{r}_1, \mathbf{r}_2) - 1)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_2 = v_H([\rho]; \mathbf{r}_1) + v_{scr}([\rho]; \mathbf{r}_1) \tag{7.9}$$

The potentials v_H and v_{scr} are long-ranged, having asymptotic N/r resp. $-1/r$ behaviour. Minimization of (7.4) with respect to ρ leads to the Euler-Lagrange equation (7.2) with the following expression for v_{eff}

$$v_{eff}([\rho], \mathbf{r}) = v_{ext} + v_H([\rho]; \mathbf{r}) + v_{scr}([\rho]; \mathbf{r}) + v_{kin}([\rho]; \mathbf{r})$$
$$+ v_{scr}^{resp}([\rho]; \mathbf{r}) + v_{kin}^{resp}([\rho]; \mathbf{r}) \tag{7.10}$$

Here the potential v_{scr}^{resp} is an integral of the linear "response" of g, $\frac{\delta g(\mathbf{r}_1, \mathbf{r}_2)}{\delta\rho(\mathbf{r}_3)}$

$$v_{scr}^{resp}([\rho]; \mathbf{r}_3) = \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\frac{\delta g([\rho]; \mathbf{r}_1, \mathbf{r}_2)}{\delta\rho(\mathbf{r}_3)}d\mathbf{r}_1 d\mathbf{r}_2 \tag{7.11}$$

This potential is a measure of the sensitivity of the pair-correlation function g to density variations. These density variations are to be understood in the following way. If we change the ground state density ρ to $\rho + \delta\rho$ (v-representable), then of course $E_v[\rho + \delta\rho]$ will be higher than $E_v[\rho]$ as we do not change the external potential v_{ext}. By the Hohenberg-Kohn theorem a wavefunction $\psi[\rho + \delta\rho]$ can be obtained as the ground state wavefunction corresponding to a unique external potential $v + \delta v$ (which will minimize $E_{v+\delta v}$). For this system we of course also have a corresponding Kohn-Sham system and a pair-correlation function $g([\rho + \delta\rho]; \mathbf{r}_1, \mathbf{r}_2)$. So the derivative of g in the above potential v_{scr}^{resp} may be regarded as linear response in the above sense of g to density changes $\delta\rho$ caused by potential changes δv.
v_{kin}^{resp} is the response of the potential v_{kin} to density variation

$$v_{kin}^{resp}([\rho]; \mathbf{r}_1) = \int \rho(\mathbf{r}_2)\frac{\delta v_{kin}([\rho]; \mathbf{r}_2)}{\delta\rho(\mathbf{r}_1)}d\mathbf{r}_2 \tag{7.12}$$

Like v_{scr} and v_{kin}, the potentials v_{scr}^{resp} and v_{kin}^{resp} can be related directly to the electron correlation in the ground state. Their sum can according to ref. [86] be expressed in terms of the conditional amplitude:

$$v_{scr}^{resp}([\rho]; \mathbf{r}_1) + v_{kin}^{resp}([\rho]; \mathbf{r}_1) = v^{N-1}([\rho]; \mathbf{r}_1)$$
$$= \int \Phi^*(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)H^{N-1}\Phi(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)ds_1 d\vec{x}_2 d\vec{x}_3...d\vec{x}_N - E_0^{N-1} \tag{7.13}$$

where H^{N-1} is the hamiltonian of electrons $2-N$ (interacting) in the given external potential. The value of v^{N-1} at position $\mathbf{r}$ is the energy expectation value of the system of $N-1$ electrons described by the conditional amplitude $\Phi(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r})$ minus the ground state energy of the $N-1$ electron system.

An expression for the Kohn-Sham potential v_s can be derived from equation (7.10) for v_{eff} and a suitable expression for the Pauli potential v_P [117, 115, 116, 107] , the difference between v_{eff} and v_s. The formula of [115] for the Pauli potential is useful. It may be interpreted with the help of the relation derived in [86] between the Pauli potential and the Kohn-Sham determinantal wavefunction ψ_s, which may be built from the Kohn-Sham orbitals $\phi_i(\mathbf{r})$ of eq.(7.1):

$$v_P([\rho]; \mathbf{r}) = v_{eff}([\rho]; \mathbf{r}) - v_s([\rho]; \mathbf{r}) = v_{s,kin}([\rho]; \mathbf{r}) + v_s^{N-1}([\rho]; \mathbf{r}) \tag{7.14}$$

where $v_{s,kin}$ and v_s^{N-1} are the analogues of (7.6) and (7.13) but now constructed with the conditional amplitude Φ_s of the determinant ψ_s. Due to the simple one-determinantal nature of ψ_s, $v_{s,kin}$ can be expressed straightforwardly in terms of the Kohn-Sham orbitals (cf. [115]):

$$v_{s,kin}([\rho]; \mathbf{r}_1) = \frac{1}{2} \int |\nabla_1 \Phi_s(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)|^2 ds_1 d\vec{x}_2 ... d\vec{x}_N = \frac{1}{2} \sum_{i=1}^{N} \left| \nabla_1 \frac{\phi_i(\mathbf{r}_1)}{\rho^{\frac{1}{2}}(\mathbf{r}_1)} \right|^2 \tag{7.15}$$

It has been argued [86] that from the relation between v_{kin} resp. $v_{s,kin}$ and the conditional amplitude Φ resp. Φ_s, one may expect large peaks in the kinetic potentials at shell boundaries due to the special behaviour of the exchange hole, a point to which we will return below.

The second contribution to v_P can also be obtained easily from the one-electron nature of H_s^{N-1} and the one-determinantal Φ_s:

$$\begin{aligned}
v_s^{N-1}([\rho]; \mathbf{r}_1) &= \int \Phi_s^*(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1) H_s^{N-1} \Phi_s(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1) ds_1 d\vec{x}_2 d\vec{x}_3 ... d\vec{x}_N - E_{s,0}^{N-1} \\
&= \mu - \sum_{i=1}^{N} \epsilon_i \frac{|\phi_i(\mathbf{r}_1)|^2}{\rho(\mathbf{r}_1)}
\end{aligned} \tag{7.16}$$

Here $E_{s,0}^{N-1}$ is the energy of the $N-1$ Kohn-Sham system in the following sense:

$$E_{s,0}^{N-1} = \int \psi_s^{N-1}(\vec{x}_2, \ldots, \vec{x}_N)^* H_s^{N-1}(\vec{x}_2, \ldots, \vec{x}_N) \psi_s^{N-1}(\vec{x}_2, \ldots, \vec{x}_N) d\vec{x}_2 \ldots d\vec{x}_N \tag{7.17}$$

where ψ_s^{N-1} is the $N-1$-particle Kohn-Sham determinant with one electron removed and where H_s^{N-1} is the Kohn-Sham Hamiltonian H_s with one electron removed (the Kohn-Sham potential $v_s([\rho]; \mathbf{r})$ is however still determined by the N-electron density ρ so this is not strictly the Kohn-Sham Hamiltonian of the ion):

$$H_s^{N-1}(\vec{x}_2, \ldots, \vec{x}_N) = \sum_{i=2}^{N} -\frac{1}{2} \nabla_i^2 + v_s([\rho]; \vec{r}_i). \tag{7.18}$$

$\mu = E_{s,0}^N - E_{s,0}^{N-1}$ is equal to the highest occupied one-electron Kohn-Sham energy, ϵ_M, which in turn is equal to the negative of the exact first ionisation energy, $\epsilon_M = E_0^N - E_0^{N-1}$.

The components of v_P are the Kohn-Sham analogues of the components v_{kin} and v_{kin}^{resp} of v_{eff} (eq.7.10), defined in terms of the exact wavefunction in eqs. (7.6) and (7.12). To demonstrate this we note that

$$T_s[\rho] = T_W[\rho] + \int \rho(\mathbf{r}) v_{s,kin}([\rho]; \mathbf{r}) d\mathbf{r} = T_W[\rho] + T_P[\rho] \tag{7.19}$$

and since $v_P = \delta T_P / \delta \rho$ [115]

$$v_P([\rho]; \mathbf{r}) = v_{s,kin}([\rho]; \mathbf{r}) + v_{s,kin}^{resp}([\rho]; \mathbf{r}) \tag{7.20}$$

where

$$v_{s,kin}^{resp}([\rho]; \mathbf{r}_1) = \int \rho(\mathbf{r}_2) \frac{\delta v_{s,kin}([\rho]; \mathbf{r}_2)}{\delta \rho(\mathbf{r}_1)} d\mathbf{r}_2$$

Comparison to eq. 7.14 identifies $v_{s,kin}^{resp}$ with v_s^{N-1}, and indeed v_s^{N-1} will not have a response term corresponding to the term v_{scr}^{resp} in v^{N-1} (eq.7.13) since the Kohn-Sham system of electrons is non-interacting.

Combining now equation 7.20 with equation 7.10 for v_{eff} one obtains for v_s

$$v_s([\rho]; \mathbf{r}) = v_{ext}(\mathbf{r}) + v_H([\rho]; \mathbf{r}) + v_{scr}([\rho]; \mathbf{r}) + v_{scr}^{resp}([\rho]; \mathbf{r})$$
$$+(v_{kin}([\rho], \mathbf{r}) - v_{s,kin}([\rho], \mathbf{r})) + (v_{kin}^{resp}([\rho], \mathbf{r}) - v_{s,kin}^{resp}([\rho], \mathbf{r})). \tag{7.21}$$

It is to be expected that the potentials v_{kin} and $v_{s,kin}$, and the potentials v_{kin}^{resp} and $v_{s,kin}^{resp}$, will be rather similar in atoms with predominantly exchange effects and only dynamical (no near-degeneracy) correlation, in which case the exchange-correlation potential $v_{xc} = v_s - v_{ext} - v_H$ would be dominated by $v_{scr} + v_{scr}^{resp}$. As a matter of fact, since the exchange-correlation energy $E_{xc}[\rho]$ can be written as the sum of a kinetic part

$$T_{xc}[\rho] = T[\rho] - T_s[\rho] = \int \rho(\mathbf{r})(v_{kin}([\rho]; \mathbf{r}) - v_{s,kin}([\rho]; \mathbf{r}))d\mathbf{r} \tag{7.22}$$

and an interparticle part

$$W_{xc}[\rho] = \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}(g([\rho]; \mathbf{r}_1, \mathbf{r}_2) - 1), \tag{7.23}$$

the exchange-correlation potential can be split into a kinetic part that corresponds to the second line of equation 7.21 and a two-particle interaction part that corresponds to $v_{scr} + v_{scr}^{resp}$:

$$v_{xc}([\rho]; \mathbf{r}_1) = \frac{\delta E_{xc}}{\delta \rho(\mathbf{r}_1)} = \frac{\delta T_{xc}}{\delta \rho(\mathbf{r}_1)} + \frac{\delta W_{xc}}{\delta \rho(\mathbf{r}_1)} = t_{xc}([\rho]; \mathbf{r}_1) + w_{xc}([\rho]; \mathbf{r}_1) \tag{7.24}$$

$$t_{xc}([\rho]; \mathbf{r}_1) = v_{kin}([\rho]; \mathbf{r}_1) - v_{s,kin}([\rho]; \mathbf{r}_1) + \int d\mathbf{r}_2 \rho(\mathbf{r}_2)(\frac{\delta v_{kin}([\rho]; \mathbf{r}_2)}{\delta \rho(\mathbf{r}_1)} - \frac{\delta v_{s,kin}([\rho]; \mathbf{r}_2)}{\delta \rho(\mathbf{r}_1)})$$

$$= v_{kin}([\rho]; \mathbf{r}_1) - v_{s,kin}([\rho]; \mathbf{r}_1) + v_{kin}^{resp}([\rho]; \mathbf{r}_1) - v_{s,kin}^{resp}([\rho]; \mathbf{r}_1) \tag{7.25}$$

$$w_{xc}([\rho]; \mathbf{r}_1) = \int d\mathbf{r}_2 \frac{\rho(\mathbf{r}_2)(g([\rho]; \mathbf{r}_1, \mathbf{r}_2) - 1)}{|\mathbf{r}_1 - \mathbf{r}_2|} + \frac{1}{2} \int d\mathbf{r}_2 d\mathbf{r}_3 \frac{\rho(\mathbf{r}_2)\rho(\mathbf{r}_3)}{|\mathbf{r}_2 - \mathbf{r}_3|} \frac{\delta g([\rho]; \mathbf{r}_2, \mathbf{r}_3)}{\delta \rho(\mathbf{r}_1)} \tag{7.26}$$

$$= v_{scr}([\rho]; \mathbf{r}_1) + v_{scr}^{resp}([\rho]; \mathbf{r}_1) \tag{7.27}$$

If the Kohn-Sham wavefunction ψ_s is regarded as a good approximation to the exact ground state wavefunction ψ_ρ then it follows from the above formula that t_{xc} is very small and the major part of the exchange-correlation potential then originates from the second term w_{xc}. In systems with strong near-degeneracy correlation, however, such as dissociating molecules, a one-determinantal wavefunction cannot be close to the exact wavefunction and in fact there are then regions (such as the bond-midpoint region [86]) where v_{kin} deviates strongly from $v_{s,kin}$.

Figure 7.1: Radial density $p(r) = 4\pi r^2 \rho(r)$ for Kr and Cd.

7.3   Structure of local potentials

We now proceed to investigate the relation between certain characteristics of the various local potentials defined above, and therefore of v_s and v_{eff}, and the atomic shell structure. The latter is described in a compact form with the radial Euler-Lagrange equation

$$(-\frac{1}{2}\frac{d^2}{dr^2} + v_{eff}(r))\sqrt{p(r)} = \mu\sqrt{p(r)} \tag{7.28}$$

where $p(r) = 4\pi r^2 \rho(r)$ is the radial density. Owing to (7.28), the atomic problem is effectively reduced to the one-dimensional problem of a single particle in the state $p^{\frac{1}{2}}(r)$ bound to the potential $v_{eff}(r)$ with the energy μ. $v_{eff} - v_{ext}$ represents the repulsive barrier of the electron potential, which prevents localization of the electron density in the inner region and produces the true atomic density distribution $p(r)$. Figure 1 displays the functions $p(r)$ for Kr and Cd with the characteristic peaks in the shell regions.

In Figures 2,3 the components of our partitioning of the potentials v_{eff} and v_s are presented for Kr and Cd. All local potentials are extracted from the optimized potential model (OPM) [33, 91]. The solution of the OPM is equivalent to that obtained by the minimization of (7.3) with respect to ρ with the restriction on the functions ψ to be Slater determinants. As a result, v_{eff} in the Euler-Lagrange equation for $p^{\frac{1}{2}}$ is approximated as

$$v_{eff}(r) \approx v_{ext}(r) + v_{OPM,cond}(r) + v_{OPM,kin}(r) + v_{OPM,kin}^{resp}(r) + v_{OPM,scr}^{resp}(r) \qquad , \tag{7.29}$$

where all local potentials are constructed as defined above with ψ_{OPM} being the optimal wavefunction of the OPM. The one-determinantal approach is well justified for atoms, so that in this case

Figure 7.2: $v_{eff} - v_{ext}$ and its components v_{cond}, v_{kin} and v_{kin}^{resp}. All potentials are derived from the optimized potential model, OPM [33]. a) Kr and b) Cd and c) Be

Figure 7.3: Demonstration of the step character of the screening-response potential v_{scr}^{resp}. a) Kr and b) Cd

(7.29) should be close to the exact expression (7.10). One can see from the Figures that partitioning (7.29) represents $v_{eff} - v_{ext}$ as a sum of the monotonous potential of Coulomb repulsion $v_{OPM,cond}$ and three additional potentials. It is remarkable, and also pleasing in view of the desirability of accurate modelling, that the potential v_{cond} is so smooth. The other potentials have distinct shell structure and have a very similar form for both represented atoms. Specifically, $v_{OPM,kin}$ has the peaks in the intershell regions that we anticipated, while $v_{OPM,kin}^{resp}$ is a step function, with the steps representing the energetical characteristics of the individual shells (see below). It is a quite intriguing observation that the potential $v_{OPM,scr}^{resp}$, which is particularly important for v_s, also exhibits very clear step function behaviour (see fig. 3).

The shell structure of the potentials finds its interpretation in the formulas for the local potentials presented in this paper. Considering first $v_{OPM,kin}$, it follows from eq. (7.15) that this potential is a measure of the change of the conditional probability amplitude as a function of position of the reference electron. The structure of the kinetic potential can readily be determined from physical considerations using the interpretation of the conditional probability amplitude. If the reference electron at $\mathbf{r}$ is positioned well within one atomic shell then as a consequence of the Pauli principle the probability will be large that the other electrons will have positions within the other atomic shells. If the position of the reference electron crosses an atomic shell boundary and moves into another atomic shell then the electrons in this new shell have to switch to another shell and the conditional probability amplitude will have to describe a large probability for a different distribution of the electrons over the atomic shells. So the changes in Φ_s as a function of the position $\mathbf{r}$ of the reference electron will be the largest at the atomic shell boundaries and we therefore expect $v_{s,kin}$ to have local maxima at the atomic shell boundaries which is indeed what we find. This explanation fits in with what is known about the behaviour of the exchange hole [87, 118, 119]. The exchange hole can be described by the Fermi amplitude or orbital, the square of which gives

Figure 7.4: The local exchange potential of the OPM as the sum of the screening or Slater (Fermi hole) potential v_{scr} and its response v_{scr}^{resp}. a) Kr and b) Cd

in a closed shell system the hole density [87]:

$$f(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i=1}^{N} \frac{\phi_i(\mathbf{r}_1)}{\rho^{\frac{1}{2}}(\mathbf{r}_1)} \phi_i(\mathbf{r}_2) \qquad g_s([\rho]; \mathbf{r}_1, \mathbf{r}_2) = 1 - \frac{1}{2} \frac{|f(\mathbf{r}_1, \mathbf{r}_2)|^2}{\rho(\mathbf{r}_2)} \qquad (7.30)$$

$$v_{s,kin}([\rho]; \mathbf{r}_1) = \frac{1}{2} \int |\nabla_1 \Phi_s(s_1, \vec{x}_2, ..., \vec{x}_N | \mathbf{r}_1)|^2 ds_1 d\vec{x}_2 ... d\vec{x}_N = \frac{1}{2} \int |\nabla_1 f(\mathbf{r}_1, \mathbf{r}_2)|^2 d\mathbf{r}_2 \qquad (7.31)$$

As was described in [87], for various positions $\mathbf{r}$ within an atomic shell, or generally within a localized orbital, $f(\mathbf{r}_1, \mathbf{r}_2)$, as a function of $\mathbf{r}_2$, is localized within the shell (localized orbital) region and insensitive to the reference electron position $\mathbf{r}_1$, but changing the reference position $\mathbf{r}_1$ through the small intershell interval produces a sharp change of $f(\mathbf{r}_1, \mathbf{r}_2)$, its "jump" from one localization region to another. This may already be seen from the fact that orbitals $\phi_i(\mathbf{r}_2)$ will get a large weight in the expansion of $f(\mathbf{r}_1, \mathbf{r}_2)$, when the reference position $\mathbf{r}_1$ is such that the coefficients $|\phi_i(\mathbf{r}_1)|/\sqrt{\rho(\mathbf{r}_1)}$ are large, i.e. when $\mathbf{r}_1$ is in the region where ϕ_i has a large amplitude (see [119]). Since the form of $v_{OPM,kin}$ is determined by the rate $\nabla_1 f(\mathbf{r}_1, \mathbf{r}_2)$ of the change of the Fermi orbital $f(\mathbf{r}_1, \mathbf{r}_2)$ with changing $\mathbf{r}_1$, the characteristic peaks of $v_{OPM,kin}$ in the intershell regions (See Figure 2) reflect the maximal mobility of $f(\mathbf{r}_1, \mathbf{r}_2)$ in these regions [86].
We now turn to the behaviour of $v_{s,kin}$ near the atomic nucleus. Some properties in this region can be deduced from equation 7.15. The electron density has the following behaviour near the atomic nucleus [120]:

$$\rho(\mathbf{r}) = \rho(\vec{0})(1 - 2Zr + O(r^2)) \quad (r \downarrow 0) \qquad (7.32)$$

where Z is the nuclear charge. The behaviour of the Kohn-Sham orbitals in the nuclear region can be deduced from the fact that close to the nucleus the only important potentials in the radial Kohn-Sham equations are the nuclear potential $-Z/r$ and the centrifugal potential $l(l+1)/2r^2$ where l is the orbital angular momentum quantum number of the orbital. Inserting a series expansion of the orbitals ϕ_i in the Kohn-Sham equation then yields:

$$\phi_i(\mathbf{r}) = c_l r^l + O(r^{l+1}) \quad (r \downarrow 0) \qquad (7.33)$$

and in particlar for s-orbitals:

$$\phi_i(\mathbf{r}) = \phi_i(\vec{0})(1 - Zr + O(r^2)) \quad (r \downarrow 0) \qquad (7.34)$$

By inserting the above expansions for the density of equation 7.32 and for the orbitals of equations 7.33 and 7.34 into equation 7.15 for $v_{s,kin}$ it follows that the only terms which give a finite contribution to $v_{s,kin}(\vec{0})$ are the terms containing orbitals with angular momentum quantum number $l = 1$,i.e. terms containing p-orbitals. So we can conclude that $v_{s,kin}$ is in general not zero at the atomic nucleus except for the case that no p-orbitals are occupied in the Kohn-Sham wavefunction in which case $v_{s,kin}(\vec{0}) = 0$. This has been illustrated in figure 2 where we find that $v_{s,kin}(0)$ is zero for the case of the beryllium atom (no p-orbitals occupied) whereas it is nonzero for the case of cadmium and krypton. For comparison we have drawn the plot for the beryllium atom on the same scale as the one given in reference [107].
We further note that the proof given in reference [86] that the kinetic potential v_{kin} of the exact ground state wave function is zero at the atomic nucleus is incorrect. The error is in the step from equation (A1) to (A2) in ref.[86]. The correct treatment of this step is given by Bingel [120].
For atoms $v_{OPM,kin}$ (cf. eq. 7.15) represents the dominant part of v_{kin}, so that going beyond the OPM, one can expect analogous intershell peaks of v_{kin} due to the specially large mobility of the

conditional amplitude Φ in these regions. It should however be remembered that in systems with strong nondynamical correlation, such as dissociating molecules, v_{kin} will show additional structure, not present in $v_{s,kin}$, notably in the bond region [86].

Turning now to the kinetic response potential $v^{resp}_{OPM,kin}$ (See Figure 2), we note that its stepped form follows from its expression (7.16). Writing this in the form

$$v^{resp}_{OPM,kin}(\mathbf{r}) = \sum_{i=1}^{N} (\mu - \epsilon_i) \frac{|\phi_i(\mathbf{r})|^2}{\rho(\mathbf{r})} \tag{7.35}$$

it is clear that for $\mathbf{r}$ within shell i, $|\phi_i(\mathbf{r})|^2$ (or the sum over $|\phi_i(\mathbf{r})|^2$ with the same ϵ_i) dominates over all other contributions, i.e. is approximately equal to $\rho(\mathbf{r})$, so eq.(7.35) is expected to describe a step function with step height in shell i equal to $\mu - \epsilon_i$, which is the energy needed to excite an electron to the "Fermi" level μ.

The sum $v_{OPM,kin} + v^{resp}_{OPM,kin}$, the Pauli potential v_P, may be interpreted as a contribution to the barrier v_{eff} due to the effect of the Fermi-Dirac statistics on the kinetic functional and its functional derivative [106]. v^{resp}_{scr} also brings a repulsive contribution to the barrier v_{eff}, but being of the order of a few a.u. for $r < 1$, this contribution is small as compared to those from v^{resp}_{kin}, v_{kin}, though not negligible. There is a clear correlation between the maxima in v_{eff} due to the peaks and steps of v_{kin} and v^{resp}_{kin} and the minima in the radial density, cf. [117]. As is immediately evident from equations (7.15) and (7.16), for all distances v_P is a non-negative function [115]. One can see from Figure 2 that v_{cond} is long-ranged (v_{cond} has an asymptotic $(N-1)/r$ behaviour), whereas both the kinetic potential and the kinetic response potential are short-ranged. At short radial distances $v_P(\mathbf{r})$ brings the dominant contribution to v_{eff}.

We now focus on the Kohn-Sham potential and its constituents (figures 3 and 4). In the OPM we are dealing with the situation that the Kohn-Sham potential does not have a kinetic contribution (cf. discussion following eq.(7.21)). If we subtract out from v_{OPM} the Hartree potential, the exchange potential of the OPM is obtained

$$v_{OPM,x} = v_{OPM} - v_{ext} - v_{OPM,H} = v_{OPM,scr} + v^{resp}_{OPM,scr} \tag{7.36}$$

The two contributions are displayed in figure 4. It is interesting to observe that the screening potential is quite smooth. In this case v_{scr} is purely due to the Fermi hole, there is of course no Coulomb hole in ψ_{OPM}. v_{scr} is then equal to the Slater potential [101], the average Hartree-Fock exchange potential (using the OPM orbitals). The small intershell peaks [33] that are present in v_x are evidently built in by the superposition of the stepped form of v^{resp}_{scr} on the smooth v_{scr}. This observation may be expected to considerably facilitate accurate modelling of v_x. The stepped form of v^{resp}_{scr} (figure 3) does not follow immediately from the expressions given in this paper. A detailed analysis of the response of the screening factor, $\delta g_s(\mathbf{r}_1, \mathbf{r}_2)/\delta\rho(\mathbf{r}_3)$, is called for but will clearly be more involved than the treatment of the other potentials and is outside the scope of the present paper.

Since we have restricted ourselves sofar to the OPM one-determinantal wavefunction, no information has been obtained on the effect of dynamical (Coulomb) correlation. For some light atoms such as Be and Ne extensive configuration interaction calculations [82, 83] have yielded highly accurate electron densities, from which corresponding Kohn-Sham potentials have been generated [79, 80, 32]. The difference $v_s - v_{ext} - v_H$ is the exchange-correlation potential $v_{xc} = \delta E_{xc}/\delta\rho$. The exchange

Figure 7.5: The correlation potential as the difference between the exchange-correlation potential and the local exchange potential of the OPM. (a) Be and (b) Ne

part may be defined as $v_{OPM,x}$, but an alternative and perhaps more consistent definition of the exchange part of v_s would be in terms of the exchange hole of the Kohn-Sham determinant ψ_s,

$$v_x([\rho]; \mathbf{r}) = \frac{\delta E_{s,x}}{\delta \rho(\mathbf{r})} = v_{s,scr}([\rho]; \mathbf{r}) + v_{s,scr}^{resp}([\rho]; \mathbf{r}) \qquad (7.37)$$

so that the correlation part of v_s would be

$$\begin{aligned} v_{corr}([\rho]; \mathbf{r}) &= v_s([\rho]; \mathbf{r}) - v_{ext}(\mathbf{r}) - v_H([\rho]; \mathbf{r}) - v_x([\rho]; \mathbf{r}) \\ &= v_{scr}([\rho]; \mathbf{r}) + v_{scr}^{resp}([\rho]; \mathbf{r}) + v_{kin}([\rho]; \mathbf{r}) + v_{kin}^{resp}([\rho]; \mathbf{r}) \\ &\quad - (v_{s,scr}([\rho]; \mathbf{r}) + v_{s,scr}^{resp}([\rho]; \mathbf{r}) + v_{s,kin}([\rho]; \mathbf{r}) + v_{s,kin}^{resp}([\rho]; \mathbf{r})) \end{aligned} \qquad (7.38)$$

This may also be written as a sum of the kinetic part of v_{xc} plus the correlation part w_c, defined as:

$$w_c([\rho]; \mathbf{r}) = w_{xc}([\rho]; \mathbf{r}) - v_x([\rho]; \mathbf{r}) \qquad (7.39)$$

Thus we have:

$$v_{corr}([\rho]; \mathbf{r}) = w_c([\rho]; \mathbf{r}) + t_{xc}([\rho]; \mathbf{r}) \qquad (7.40)$$

Probably v_x is very close to $v_{OPM,x}$. Figure 5 shows the correlation part of v_s as $v_s - v_{ext} - v_H - v_{OPM,x}$ for Be and Ne. Although we used a slightly different definition of the correlation part of v_s than Aryasetiawan and Stott [79], who took $v_x([\rho^{HF}]; \mathbf{r}) = v_s([\rho^{HF}]; \mathbf{r}) - v_{ext}(\mathbf{r}) - v_H([\rho^{HF}]; \mathbf{r})$ where ρ^{HF} is the Hartree-Fock density, for the exchange part of v_{xc} rather than $v_{OPM,x}$, our curves of v_{corr} are very similar to those of ref. [79]. It is notable that v_{corr} is quite small relative to v_x. This may be due to a cancellation of the contributions from different terms in v_{corr} (7.38). For instance, to the extent that the OPM determinantal wavefunction may be identified with the Hartree-Fock determinantal wavefunction, the correlation contribution to the kinetic energy will be $\int \rho(v_{kin} - v_{OPM,kin})d\vec{r}$. This is known to be positive and equal to the absolute value of the total correlation energy on account of the virial theorem, so the difference $v_{kin} - v_{OPM,kin}$ must be predominantly positive. On the other hand, $(1/2)\int \rho(v_{scr} - v_{OPM,scr})d\vec{r}$ is the contribution to the correlation energy due to the electron-electron interaction. For a system like Ne this is negative and close to twice the (negative) correlation energy [119], so v_{scr} will be predominantly more negative than $v_{OPM,scr}$. This is an expected effect of the Coulomb hole that affects v_{scr} but not $v_{OPM,scr}$. At the border region between the K and L shells, where the exchange hole is not so strongly localized around the reference position as when the latter is within a shell, the effect of the Coulomb hole is probably particularly noticable. This may explain the negative minimum in v_{corr} at the K/L shell boundary. It should be kept in mind that there are several contributions to v_{corr}, not only the negative Coulomb hole contribution $v_{scr} - v_{s,scr}$. Ref. [86] demonstrates that in the outer part of the $(1s)^2$ shell of He the positive contributions of the response potentials (v^{N-1}) and the kinetic potential outweigh the negative Coulomb hole contribution $v_{cond} - v^{HF} \approx v_{scr} - v_{s,scr}$. Only at the inner part of the shell the Coulomb hole is so strongly negative that it causes v_{corr} to turn negative. This pattern is repeated in the K and L shells of Be and Ne, although in Be the total v_{corr} never becomes positive. This may be related to the well-known strong near-degeneracy correlation and therefore more significant Coulomb hole contribution in Be.

The present paper analyzes the relationship between electron correlation and local one-electron potentials featuring in DFT, and establishes the connection with the atomic shell structure. The results of this analysis are also useful for the accurate and efficient approximation of v_s and v_{eff} in DFT [121, 32]. In this respect, the stepped shell structure found for v_{scr}^{resp} is of special interest in

our opinion. Further elucidation of the origin of this special structure will be undertaken, but it is clear that, now that it has been recognized, this simple step behaviour considerably facilitates the accurate modelling of the exchange-correlation Kohn-Sham potential. The corresponding work is in progress.

Chapter 8

Step structure in the atomic Kohn-Sham potential

In this work we analyze the exchange-correlation potential v_{xc} within the Kohn-Sham approach to density functional theory for the case of atomic systems. The exchange-correlation potential is written as the sum of two potentials. One of these potentials $v_{xc,scr}$ is the long-range Coulombic potential of the coupling constant integrated exchange-correlation hole which represents the screening of the two-particle interactions due to exchange-correlation effects. The other potential $v_{xc,scr}^{resp}$ contains the functional derivative with respect to the electron density of the coupling constant integrated pair-correlation function representing the sensitivity of this exchange-correlation screening to density variations. An explicit expression of the exchange-part of this functional derivative is derived using an approximation for the Greens function of the Kohn-Sham system and is shown to display a distinct atomic shell structure. The corresponding potential $v_{x,scr}^{resp}$ has a clear step structure and is constant within the atomic shells and changes rapidly at the atomic shell boundaries. Numerical examples are presented for the Be and Kr atoms using the Optimized Potential Model (OPM).

8.1   Introduction

The Kohn-Sham approach to density functional theory (DFT) [5] is an exact scheme to obtain the ground state properties of electronic many-particle systems by solving the problem of a system of noninteracting particles moving in the field of an effective Kohn-Sham potential $v_s([\rho]; \mathbf{r})$ which is a functional of the electron density ρ. An important constituent of this potential is the exchange-correlation potential $v_{xc}([\rho]; \mathbf{r})$ which is defined as the functional derivative of the exchange-correlation functional $E_{xc}[\rho]$. With the appearance of so-called generalized gradient approximations (GGA's) [44, 45, 43, 49] increasingly accurate approximations for this exchange-correlation functional have been applied with considerable succes in the calculations of properties of electronic systems ranging from atoms and molecules [71, 70, 122] to solids and surfaces [47]. However the corresponding exchange-correlation potential still has some deficiencies [49, 32].
In this paper we will perform an analysis of the exchange-correlation potential and in particular the exchange potential of the Kohn-Sham system. These potentials have two important characteristics. First of all the exchange potential v_x and the exchange-correlation potential v_{xc} have an asymptotic Coulombic $-1/r$ behaviour in finite systems such as atoms and molecules [31, 49, 32, 35]. Secondly, they display the atomic shell structure in atoms which is characterized by peaks at the atomic shell

boundaries [49, 35, 32]. In this paper we show that v_x and v_{xc} can be written as the sum of two potentials. One of these potentials is monotonous and has a long range Coulombic behaviour. The other is short range and displays a clear step structure being constant within atomic shells and changing rapidly at the atomic shell boundaries [108]. The step structure of the latter potential can be derived using an approximation of the Greens function of the Kohn-Sham system as we will demonstrate in the remainder of this paper.

The required features of v_{xc} and v_x, i.e. the the Coulombic asymptotics and the atomic shell structure are not well-presented by most of the approximate exchange and exchange-correlation potentials. The required asymptotic Coulombic $-1/r$ behaviour in finite systems such as atoms and molecules is not reproduced with the current local density (LDA) and generalized gradient approximations. As a result of this a too low absolute value for the highest occupied Kohn-Sham orbital, which should be equal to the ionisation energy of the system [31], is obtained. Furthermore also the required atomic shell structure is not correctly reproduced with the LDA and GGA approximations [121]. The GGA potentials even have a wrong Coulombic behaviour near the atomic nucleus [92, 32].

An other approach leading to better exchange-correlation potentials is the so-called weighted density approximation (WDA) [123, 124, 125] which is based on approximate pair-correlation functions. In this approach the exchange-correlation potential is split up as the sum of two potentials, one containing the coupling constant integrated pair-correlation function which we will call the screening potential $v_{xc,scr}$ and one containing the functional derivative of this function with respect to the electron density, which we will call the screening response potential $v_{xc,scr}^{resp}$. The exchange part of this last potential is the main subject of this paper. Due to the fact that the approximate pair-correlation function is required to satisfy the sum rule and must integrate to one electron the corresponding exchange-correlation potential has a Coulombic asymptotic behaviour. However as all current approximate pair-correlation functions are not symmetric under interchange of the electron coordinates the Coulombic behaviour is in general not $-1/r$ as required but $-c/r$ with c some constant.

Most of the approximate exchange and exchange-correlation potentials do not exhibit the atomic shell structure. Only a few attempts have been made to incorporate atomic shell structure in a WDA scheme [126, 127] employing the idea that electron interactions within one shell can be treated within a local density type of approximation but not the interaction between electrons in different atomic shells. This physical picture is reflected in the step potential, exchange and correlation effects do not change over regions which lie well within an atomic shell but they do change considerably when we move from one atomic shell to another.

A very good approximation to the exchange part v_x of v_{xc} has recently been proposed by Krieger et al. [35]. This approximation displays the atomic shell structure, has the correct asymptotics and also satisfies the requirement of integer preference. This approximation was derived using an approximation for the Green's function of the Kohn-Sham system [34, 35]. In this paper we will use a similar approximation to derive an explicit expression of the functional derivative with respect to the electron density pair-correlation function of the Kohn-Sham system. Using this approximation one can derive an expression for the screening response part $v_{x,scr}^{resp}$ of the exchange potential. This response part is shown to display a clear atomic step structure, being constant within the atomic shells and changing rapidly at the atomic shell boundaries.

The paper is divided as follows. In section II we will give a short account of the potentials we want to analyze. In section III we will derive the expression for the functional derivative of the Kohn-Sham pair-correlation function using an approximation for the Green's function of the Kohn-Sham system. In section IV we derive how the step structure in $v_{x,scr}^{resp}$ arises from this functional

derivative. In section V we will present some numerical results for some atoms and in section VI we present our conclusions and discuss implications for future approximations.

8.2 The exchange-correlation potential: separation into a long range and a short range part

Within the constrained search approach to spinpolarized density functional theory the energy functional is defined as [5, 128, 67]

$$E_v[\{\rho_\sigma\}] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_L[\{\rho_\sigma\}] \tag{8.1}$$

where the total electron density $\rho = \sum_\sigma \rho_\sigma$ is the sum of the spin densities and the universal functional F_L is defined as:

$$F_L[\{\rho_\sigma\}] = \inf_{\hat{D}\to\{\rho_\sigma\}} tr\{\hat{D}(\hat{T}+\hat{W})\} \tag{8.2}$$

In this functional the infimum of the expectation value of the kinetic energy operator $\hat{T}$ and the interparticle interaction operator $\hat{W}$ is searched over all N-particle density matrices $\hat{D}$ which integrate to the prescribed spin densities $\{\rho_\sigma\}$. The functional F_L is usually split up as:

$$F_L[\{\rho_\sigma\}] = T_L[\{\rho_\sigma\}] + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\{\rho_\sigma\}] \tag{8.3}$$

where T_L is defined as:

$$T_L[\{\rho_\sigma\}] = \inf_{\hat{D}\to\{\rho_\sigma\}} tr\{\hat{D}\hat{T}\} \tag{8.4}$$

As all functionals except the exchange-correlation functional E_{xc} now are defined equation 8.3 actually defines E_{xc}. Minimalization of the energy functional E_v leads to the well-known Kohn-Sham equations:

$$(-\frac{1}{2}\nabla^2 + v_{s\sigma}([\{\rho_\sigma\}];\mathbf{r}))\phi_{i\sigma}(\mathbf{r}) = \epsilon_{i\sigma}\phi_{i\sigma}(\mathbf{r}) \tag{8.5}$$

where the spin density is given by:

$$\rho_\sigma(\mathbf{r}) = \sum_i^{N_\sigma} f_{i\sigma}|\phi_{i\sigma}(\mathbf{r})|^2 \tag{8.6}$$

and the Kohn-Sham potential $v_{s\sigma}$ is split up as:

$$v_{s\sigma}([\{\rho_\sigma\}];\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r}-\mathbf{r}_1|} + v_{xc\sigma}([\{\rho_\sigma\}];\mathbf{r}) \tag{8.7}$$

where the exchange-correlation potential $v_{xc\sigma}$ is defined as:

$$v_{xc\sigma}([\{\rho_\sigma\}];\mathbf{r}) = \frac{\delta E_{xc}[\{\rho_\sigma\}]}{\delta\rho_\sigma(\mathbf{r})} \tag{8.8}$$

The exchange-correlation potential can be further analyzed in terms of the coupling constant integrated pair-correlation function $\bar{g}_{\sigma_1\sigma_2}$ [127, 129] defined as:

$$\bar{g}_{\sigma_1\sigma_2}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2) = \int_0^1 g^\lambda_{\sigma_1\sigma_2}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2)d\lambda \tag{8.9}$$

where $g^\lambda_{\sigma_1\sigma_2}$ is defined as:

$$g^\lambda_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2) = \frac{\Gamma^\lambda_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2)}{\rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)} \tag{8.10}$$

with:

$$\Gamma^\lambda_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2) = \sum_{\sigma_3,\ldots,\sigma_N} \int D^\lambda_{\{\rho_\sigma\}}(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N) d\mathbf{r}_3 \ldots d\mathbf{r}_N \tag{8.11}$$

Here $D^\lambda_{\{\rho_\sigma\}}$ is the diagonal part of the density matrix which minimizes the functional F_L for the prescribed spin densities $\{\rho_\sigma\}$ and in which the interparticle operator $\hat{W}$ is multiplied by the constant λ. The case $\lambda = 1$ corresponds to the ground state density matrix of the fully interacting system and the case $\lambda = 0$ corresponds to the density matrix of the Kohn-Sham system. A useful relation relating the paircorrelation function at coupling strength λ to the pair correlation function of the fully interacting system is proven by Levy [27]:

$$g^\lambda_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2) = g_{\sigma_1\sigma_2}([\{\rho_{\sigma,1/\lambda}\}], \lambda\mathbf{r}_1, \lambda\mathbf{r}_2) \tag{8.12}$$

where:

$$\rho_{\sigma,1/\lambda}(\mathbf{r}) = \lambda^{-3}\rho_\sigma(\lambda^{-1}\mathbf{r}) \tag{8.13}$$

The exchange-correlation energy in terms of $\bar{g}$ is given by:

$$E_{xc}[\{\rho_\sigma\}] = \frac{1}{2} \sum_{\sigma_1\sigma_2} \int \frac{\rho_{\sigma_1}(\mathbf{r}_1)\rho_{\sigma_2}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} (\bar{g}_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2) - 1) d\mathbf{r}_1 d\mathbf{r}_2 \tag{8.14}$$

By functional differentiation it follows that $v_{xc\sigma}$ can be split up into two terms:

$$v_{xc\sigma}([\{\rho_\sigma\}]; \mathbf{r}_1) = v_{xc,scr\sigma}([\{\rho_\sigma\}]; \mathbf{r}) + v^{resp}_{xc,scr\sigma}([\{\rho_\sigma\}]; \mathbf{r}) \tag{8.15}$$

where the screening potential is defined as:

$$v_{xc,scr\sigma_1}([\{\rho_\sigma\}]; \mathbf{r}_1) = \sum_{\sigma_2} \int \frac{\rho_{\sigma_2}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} (\bar{g}_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2) - 1) d\mathbf{r}_2 \tag{8.16}$$

and the screening response potential as:

$$v^{resp}_{xc,scr\sigma_1}([\{\rho_\sigma\}]; \mathbf{r}_1) = \frac{1}{2} \sum_{\sigma_2\sigma_3} \int \frac{\rho_{\sigma_2}(\mathbf{r}_2)\rho_{\sigma_3}(\mathbf{r}_3)}{|\mathbf{r}_2 - \mathbf{r}_3|} \frac{\delta\bar{g}_{\sigma_2\sigma_3}([\{\rho_\sigma\}]; \mathbf{r}_2, \mathbf{r}_3)}{\delta\rho_{\sigma_1}(\mathbf{r}_1)} d\mathbf{r}_2 d\mathbf{r}_3 \tag{8.17}$$

The screening potential is equal to the potential of the coupling constant integrated exchange-correlation hole and represents the screening of the interparticle Coulomb potential between electrons at $\mathbf{r}_1$ and $\mathbf{r}_2$ with spin σ_1 and σ_2 by exchange-correlation effects with screening factor equal to $\bar{g}_{\sigma_1\sigma_2}([\{\rho_\sigma\}]; \mathbf{r}_1, \mathbf{r}_2)$. Due to the fact that the exchange-correlation hole integrates to one electron the screening potential has a long-range Coulombic behaviour:

$$v_{xc,scr}([\{\rho_\sigma\}]; \mathbf{r}) \sim -\frac{1}{r} \qquad (r \to \infty) \tag{8.18}$$

Physically this means that if we move one electron away from a finite system such as an atom it just experiences the potential of the ion it leaves behind.

We will carry out an analysis of the exchange part of $v_{xc,scr}$ and $v_{xc,scr\sigma}^{resp}$. We split up $\bar{g}$ into an exchange and a correlation part:

$$\bar{g}_{\sigma_1\sigma_2}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2) = g_{s\sigma_1}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2)\delta_{\sigma_1\sigma_2} + \bar{g}_{c\sigma_1\sigma_2}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2) \tag{8.19}$$

where $g_{s\sigma}$ is the pair correlation of the Kohn-Sham system:

$$g_{s\sigma}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2) = 1 - \frac{|\gamma_{s\sigma}(\mathbf{r}_1,\mathbf{r}_2)|^2}{\rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2)} \tag{8.20}$$

where

$$\gamma_{s\sigma}(\mathbf{r}_1,\mathbf{r}_2) = \sum_i^{N_\sigma} f_{i\sigma}\phi_{i\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_2) \tag{8.21}$$

the one-particle density matrix for the Kohn-Sham system. The pair correlation function g_s is unaffected by the coupling strength integration as it satisfies the following scaling property:

$$g_{s\sigma}([\{\rho_\sigma\}];\mathbf{r}_1,\mathbf{r}_2) = g_{s\sigma}([\{\rho_{\sigma,1/\lambda}\}];\lambda\mathbf{r}_1,\lambda\mathbf{r}_2) \tag{8.22}$$

This is also apparent from the fact that the coupling constant integration keeps the density constant and therefore does not influence the Kohn-Sham noninteracting system. The exchange-correlation potential $v_{xc} = v_x + v_c$ can now be written as an exchange and a correlation part with corresponding screening and screening response potentials:

$$v_{x\sigma}([\{\rho_\sigma\}];\mathbf{r}) = v_{x,scr\sigma}([\{\rho_\sigma\}];\mathbf{r}) + v_{x,scr\sigma}^{resp}([\{\rho_\sigma\}];\mathbf{r}) \tag{8.23}$$

and

$$v_{c\sigma}([\{\rho_\sigma\}];\mathbf{r}) = v_{c,scr\sigma}([\{\rho_\sigma\}];\mathbf{r}) + v_{c,scr\sigma}^{resp}([\{\rho_\sigma\}];\mathbf{r}) \tag{8.24}$$

where the potentials $v_{x,scr}$ and $v_{c,scr}$ and their responses are defined as in equations 8.16 and 8.17 with $\bar{g}$ replaced by g_s and $\bar{g}_c$. The potential $v_{x,scr}$ has a Coulombic asymptotic behaviour due to the fact that the exchange hole integrates to one electron. There is no Coulombic term in $v_{c,scr}$ as coupling constant integrated Coulomb hole integrates to zero electrons.

In figure 1 we plot some of the described potentials for the case of the beryllium atom. The exact exchange-correlation potential v_{xc} is calculated from an accurate CI (Configuration Interaction) density using the procedure described in reference [32]. The potentials $v_x, v_{x,scr}$ and $v_{x,scr}^{resp}$ are calculated within the optimized potential model (OPM) [33, 35, 36] and are probably very close to their exact values which can be obtained from the solution for v_x of the OPM integral equation [33, 35, 36] by insertion of the exact Kohn-Sham orbitals instead of the OPM orbitals. We further plotted $v_{xc} - v_{x,scr}^{OPM}$ which can be regarded as an approximation to $v_{xc,scr}^{resp} + v_{c,scr}$. Note the very clear step structure in $v_{x,scr}^{resp,OPM}$. This potential is almost constant within the 1s-shell and drops rapidly to zero at the atomic shell boundary between the 1s and the 2s shell at a radial distance of about 1 bohr. As can be seen from the graph of $v_{xc} - v_{x,scr}^{OPM}$ this step structure is somewhat smoothed by correlation effects but as these effects are less important than the exchange effects the step structure is still clearly visible.

The atomic structure of the potential $v_{x,scr}^{resp}$ is the main subject of this paper. This requires an analysis of the functional derivative $\delta g_{s\sigma'}(\mathbf{r}_1,\mathbf{r}_2)/\delta\rho_\sigma(\mathbf{r}_3)$ of the Kohn-Sham pair-correlation function. Such an analysis will be carried out in the next section.

Figure 8.1: The exchange-correlation potential and the OPM exchange potential with corresponding screening and screening response parts for the beryllium atom

8.3 Functional derivative of the Kohn-Sham pair-correlation function

In order to understand the structure of the $v_{x,scr\sigma}^{resp}$ potential we must calculate the functional derivative of the Kohn-Sham pair-correlation function. This function describes the sensitivity of the exchange screening between two electrons at $\mathbf{r}_1$ and $\mathbf{r}_2$ to density changes at point $\mathbf{r}_3$. One property of this function is readily derived. As $g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_1) = 0$ for any electron density it immediately follows that:

$$\frac{\delta g_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_1)}{\delta \rho_\sigma(\mathbf{r}_3)} = 0 \tag{8.25}$$

This puts a constraint on approximate functional derivatives of $g_{s\sigma}$. In general from the definition of $g_{s\sigma}$ it follows that:

$$\frac{\delta g_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} = -\left(\gamma_{s\sigma'}^*(\mathbf{r}_1, \mathbf{r}_2)\frac{\delta \gamma_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} + c.c.\right)\frac{1}{\rho_{\sigma'}(\mathbf{r}_1)\rho_{\sigma'}(\mathbf{r}_2)}$$

$$-\left(\frac{\delta(\mathbf{r}_1 - \mathbf{r}_3)}{\rho_\sigma(\mathbf{r}_1)} + \frac{\delta(\mathbf{r}_2 - \mathbf{r}_3)}{\rho_\sigma(\mathbf{r}_2)}\right)(g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) - 1)\delta_{\sigma\sigma'} \tag{8.26}$$

From constraint 8.25 it follows that the functional derivatives of $\gamma_{s\sigma}$ must contain deltafunctions in order to cancel the deltafunctions in the second part of the above equation for $\mathbf{r}_1$ equal to $\mathbf{r}_2$.

We now must calculate:

$$\frac{\delta \gamma_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} = \sum_i^{N_{\sigma'}} f_{i\sigma'} \left(\frac{\delta \phi_{i\sigma'}(\mathbf{r}_1)}{\delta \rho_\sigma(\mathbf{r}_3)} \phi_{i\sigma'}^*(\mathbf{r}_2) + \phi_{i\sigma'}(\mathbf{r}_1) \frac{\delta \phi_{i\sigma'}^*(\mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} \right) \tag{8.27}$$

We therefore have to calculate the functional derivative of the Kohn-Sham orbitals with respect to the density. This can be done using:

$$\frac{\delta \phi_{i\sigma'}(\mathbf{r}_1)}{\delta \rho_\sigma(\mathbf{r}_3)} = \sum_{\sigma''} \int \frac{\delta \phi_{i\sigma'}(\mathbf{r}_1)}{\delta v_{s\sigma''}(\mathbf{r}_4)} \frac{\delta v_{s\sigma''}(\mathbf{r}_4)}{\delta \rho_\sigma(\mathbf{r}_3)} d\mathbf{r}_4 \tag{8.28}$$

where $v_{s\sigma}$ is the Kohn-Sham potential. The derivative of $\phi_{i\sigma}$ with respect to the Kohn-Sham potential is known in terms of the Kohn-Sham orbitals and one-electron energies:

$$\frac{\delta \phi_{i\sigma'}(\mathbf{r}_1)}{\delta v_{s\sigma}(\mathbf{r}_4)} = -G_{i\sigma}(\mathbf{r}_1, \mathbf{r}_4) \phi_{i\sigma}(\mathbf{r}_4) \delta_{\sigma\sigma'} \tag{8.29}$$

where $G_{i\sigma}$ is the following Greens function:

$$G_{i\sigma}(\mathbf{r}_1, \mathbf{r}_4) = \sum_{j \neq i} \frac{\phi_{j\sigma}(\mathbf{r}_1) \phi_{j\sigma}^*(\mathbf{r}_4)}{\epsilon_{j\sigma} - \epsilon_{i\sigma}} \tag{8.30}$$

The other functional derivative is equal to the inverse density response function which we only need to know for equal spins:

$$\frac{\delta v_{s\sigma}(\mathbf{r}_4)}{\delta \rho_\sigma(\mathbf{r}_3)} = \chi_{s\sigma}^{-1}(\mathbf{r}_4, \mathbf{r}_3) \tag{8.31}$$

So:

$$\delta(\mathbf{r}_1 - \mathbf{r}_2) = \int \chi_{s\sigma}(\mathbf{r}_1, \mathbf{r}_4) \chi_{s\sigma}^{-1}(\mathbf{r}_4, \mathbf{r}_2) d\mathbf{r}_4 \tag{8.32}$$

where $\chi_{s\sigma}$ is the density response function:

$$\chi_{s\sigma}(\mathbf{r}_2, \mathbf{r}_4) = \frac{\delta \rho_\sigma(\mathbf{r}_2)}{\delta v_{s\sigma}(\mathbf{r}_4)} = -\sum_i^{N_\sigma} f_{i\sigma} \phi_{i\sigma}^*(\mathbf{r}_2) G_{i\sigma}(\mathbf{r}_2, \mathbf{r}_4) \phi_{i\sigma}(\mathbf{r}_4) + c.c. \tag{8.33}$$

However as a density variation $\delta \rho_\sigma(\mathbf{r})$ determines the potential variation $\delta v_{s\sigma}(\mathbf{r})$ only up to constant the derivative $\delta v_{s\sigma}/\delta \rho_\sigma$ and hence $\chi_{s\sigma}^{-1}$ is only defined up to constant. This arbitrariness can be lifted for instance by specifying $\chi_{s\sigma}^{-1}$ as an integral operator acting in the space of functions orthogonal to the constant function. The constant function is an eigenfunction of the integral operator $\chi_{s\sigma}$ with zero eigenvalue. This follows from the relation:

$$\delta \rho_{s\sigma}(\mathbf{r}_1) = \int \chi_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) \delta v_{s\sigma}(\mathbf{r}_2) d\mathbf{r}_2 \tag{8.34}$$

and from the fact that a constant potential variation $\delta v_{s\sigma}(\mathbf{r}) = C$ does not produce a density variation. Thus we have:

$$\int \chi_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = 0 \tag{8.35}$$

This also immediately follows from equation 8.33 using the fact that $G_{i\sigma}$ projects on the space orthogonal to $\phi_{i\sigma}$. From the above relation 8.35 it clearly follows that equation 8.32 has no unique solution for $\chi_{s\sigma}^{-1}$ as adding a constant to $\chi_{s\sigma}^{-1}$ gives another solution for equation 8.32. However

choosing a specific constant (which amounts to choosing a gauge for the potential) cannot influence the calculations. Keeping this in mind, we may write:

$$\frac{\delta\phi_{i\sigma'}(\mathbf{r}_1)}{\delta\rho_\sigma(\mathbf{r}_3)} = -\delta_{\sigma\sigma'}\int G_{i\sigma}(\mathbf{r}_1,\mathbf{r}_4)\phi_{i\sigma}(\mathbf{r}_4)\chi_{s\sigma}^{-1}(\mathbf{r}_4,\mathbf{r}_3)d\mathbf{r}_4 \tag{8.36}$$

and it follows that:

$$\frac{\delta\gamma_{s\sigma}(\mathbf{r}_1,\mathbf{r}_2)}{\delta\rho_\sigma(\mathbf{r}_3)} = \sum_i^{N_\sigma} -f_{i\sigma}\phi_{i\sigma}^*(\mathbf{r}_2)\int G_{i\sigma}(\mathbf{r}_1,\mathbf{r}_4)\phi_{i\sigma}(\mathbf{r}_4)\chi_{s\sigma}^{-1}(\mathbf{r}_4,\mathbf{r}_3)d\mathbf{r}_4 + (1\leftrightarrow 2)^* \tag{8.37}$$

As one can readily verify the left hand sides of the above equations indeed do not change when adding a constant to $\chi_{s\sigma}^{-1}$. To find an explicit expression for the above functional derivatives we must find an expression for the inverse density response function $\chi_{s\sigma}^{-1}$. In order to do this we make the following approximation to the Greens function (see Sharp and Horton [34], Krieger et al. [35]):

$$G_{i\sigma}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{\Delta\tilde{\epsilon}_{i\sigma}}(\delta(\mathbf{r}_1-\mathbf{r}_2) - \phi_{i\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_2)) \tag{8.38}$$

where $\Delta\tilde{\epsilon}_{i\sigma}$ is some mean energy difference. Just as the exact Greens function of equation 8.30 the above approximate Greens function projects orbital $\phi_{i\sigma}$ to zero which garantees condition 8.35. This approximate Greens function yields the following expression for the density response function $\chi_{s\sigma}$:

$$\chi_{s\sigma}(\mathbf{r}_1,\mathbf{r}_2) = a_\sigma(\mathbf{r}_1)\delta(\mathbf{r}_1-\mathbf{r}_2) - b_\sigma(\mathbf{r}_1,\mathbf{r}_2) \tag{8.39}$$

where

$$a_\sigma(\mathbf{r}_1) = \sum_i^{N_\sigma} -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}|\phi_{i\sigma}(\mathbf{r}_1)|^2 \tag{8.40}$$

and

$$b_\sigma(\mathbf{r}_1,\mathbf{r}_2) = \sum_i^{N_\sigma} -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}|\phi_{i\sigma}(\mathbf{r}_1)|^2|\phi_{i\sigma}(\mathbf{r}_2)|^2 \tag{8.41}$$

The inverse of this function is derived in the appendix using an approximation which fixes the arbitrary constant (i.e. the gauge of potential $v_{s\sigma}$) so that $v_{s\sigma} \to 0$ in infinity. It is given by:

$$\chi_{s\sigma}^{-1}(\mathbf{r}_2,\mathbf{r}_3) = \frac{\delta(\mathbf{r}_2-\mathbf{r}_3)}{a_\sigma(\mathbf{r}_2)} + \sum_{ik}^{N_\sigma-1} \eta_{ik}^\sigma \frac{|\phi_{i\sigma}(\mathbf{r}_2)|^2|\phi_{k\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_2)a_\sigma(\mathbf{r}_3)} \tag{8.42}$$

where the η_{ik}^σ are given by:

$$\eta_{ik}^\sigma = -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}(\bar{I}-\bar{N}^\sigma)_{ik}^{-1} \tag{8.43}$$

for $i,k = 1,\ldots,N_\sigma-1$ and matrix $\bar{N}^\sigma$ is given by:

$$N_{ik}^\sigma = -\frac{2f_{k\sigma}}{\Delta\tilde{\epsilon}_{k\sigma}}\int \frac{|\phi_{i\sigma}(\mathbf{r}_2)|^2|\phi_{k\sigma}(\mathbf{r}_2)|^2}{a_\sigma(\mathbf{r}_2)}d\mathbf{r}_2 \tag{8.44}$$

for $i, k = 1, \ldots, N_\sigma - 1$. With the above equation for the inverse density response function and the approximate equation for the Greens function we find using equation 8.36 that:

$$\frac{\delta \phi_{i\sigma'}(\mathbf{r}_1)}{\delta \rho_\sigma(\mathbf{r}_3)} = -\frac{\delta_{\sigma\sigma'}}{\Delta \tilde{\epsilon}_{i\sigma}} \int \left(\delta(\mathbf{r}_1 - \mathbf{r}_4) - \phi_{i\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_4) \right) \phi_{i\sigma}(\mathbf{r}_4)\chi_{s\sigma}^{-1}(\mathbf{r}_4, \mathbf{r}_3)d\mathbf{r}_4$$

$$= -\frac{\delta_{\sigma\sigma'}}{\Delta \tilde{\epsilon}_{i\sigma}}\phi_{i\sigma}(\mathbf{r}_1)\left(\chi_{s\sigma}^{-1}(\mathbf{r}_1, \mathbf{r}_3) - \int |\phi_{i\sigma}(\mathbf{r}_4)|^2 \chi_{s\sigma}^{-1}(\mathbf{r}_4, \mathbf{r}_3)d\mathbf{r}_4 \right) \tag{8.45}$$

The last integral can be worked out as:

$$\int |\phi_{i\sigma}(\mathbf{r}_4)|^2 \chi_{s\sigma}^{-1}(\mathbf{r}_4, \mathbf{r}_3)d\mathbf{r}_4 = \frac{|\phi_{i\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)} + \sum_{jk}^{N_\sigma - 1} \eta_{jk}^\sigma \int \frac{|\phi_{i\sigma}(\mathbf{r}_4)|^2 |\phi_{j\sigma}(\mathbf{r}_4)|^2}{a_\sigma(\mathbf{r}_4)}d\mathbf{r}_4 \frac{|\phi_k(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)}$$

$$= \sum_k^{N_\sigma - 1} \xi_{ik}^\sigma \frac{|\phi_{k\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)} \tag{8.46}$$

where

$$\xi_{ik}^\sigma = (\bar{I} - \bar{N}^\sigma)_{ik}^{-1} \tag{8.47}$$

We now define:

$$K_{i\sigma}(\mathbf{r}_1, \mathbf{r}_3) = -\frac{1}{\Delta \tilde{\epsilon}_{i\sigma}}\left(\frac{\delta(\mathbf{r}_1 - \mathbf{r}_3)}{a_\sigma(\mathbf{r}_3)} - \sum_k^{N_\sigma} \xi_{ik}^\sigma \frac{|\phi_{k\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)} + \sum_{jk}^{N_\sigma} \eta_{jk}^\sigma \frac{|\phi_{j\sigma}(\mathbf{r}_1)|^2 |\phi_{k\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_1)a_\sigma(\mathbf{r}_3)} \right) \tag{8.48}$$

where we for notational convenience define ξ_{ik}^σ and η_{ik}^σ to be zero for i or k equal to N_σ. This is a real function ($K_{i\sigma}^* = K_{i\sigma}$). We then have:

$$\frac{\delta \phi_{i\sigma'}(\mathbf{r}_1)}{\delta \rho_\sigma(\mathbf{r}_3)} = \phi_{i\sigma}(\mathbf{r}_1)K_{i\sigma}(\mathbf{r}_1, \mathbf{r}_3)\delta_{\sigma\sigma'} \tag{8.49}$$

and similarly:

$$\frac{\delta \phi_{i\sigma'}^*(\mathbf{r}_1)}{\delta \rho_\sigma(\mathbf{r}_3)} = \phi_{i\sigma}^*(\mathbf{r}_1)K_{i\sigma}(\mathbf{r}_1, \mathbf{r}_3)\delta_{\sigma\sigma'} \tag{8.50}$$

This formula can be used to calculate the functional derivative of all explicitly orbital dependent functionals. Using the above expressions we find:

$$\frac{\delta \gamma_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} = \delta_{\sigma\sigma'} \sum_i^{N_\sigma} f_{i\sigma}\phi_{i\sigma}(\mathbf{r}_1)(K_{i\sigma}(\mathbf{r}_1, \mathbf{r}_3) + K_{i\sigma}(\mathbf{r}_2, \mathbf{r}_3))\phi_{i\sigma}^*(\mathbf{r}_2) \tag{8.51}$$

We thus have:

$$\frac{\delta |\gamma_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)|^2}{\delta \rho_\sigma(\mathbf{r}_3)} = 2\delta_{\sigma\sigma'} \sum_{ij}^{N_\sigma} f_{i\sigma}f_{j\sigma}\phi_{j\sigma}^*(\mathbf{r}_1)\phi_{i\sigma}(\mathbf{r}_1)(K_{i\sigma}(\mathbf{r}_1, \mathbf{r}_3) + K_{i\sigma}(\mathbf{r}_2, \mathbf{r}_3))\phi_{i\sigma}^*(\mathbf{r}_2)\phi_{j\sigma}(\mathbf{r}_2) \tag{8.52}$$

Due to the $\delta_{\sigma\sigma'}$ term we have:

$$\frac{\delta g_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} = \frac{\delta g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)}\delta_{\sigma\sigma'} \tag{8.53}$$

where

$$\frac{\delta g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} = -2 \sum_{ij}^{N_\sigma} f_{i\sigma} f_{j\sigma} \frac{\phi_{j\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_1)(K_{i\sigma}(\mathbf{r}_1, \mathbf{r}_3) + K_{i\sigma}(\mathbf{r}_2, \mathbf{r}_3))\phi_{i\sigma}(\mathbf{r}_2)\phi_{j\sigma}^*(\mathbf{r}_2)}{\rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2)}$$

$$-\left(\frac{\delta(\mathbf{r}_1 - \mathbf{r}_3)}{\rho_\sigma(\mathbf{r}_1)} + \frac{\delta(\mathbf{r}_2 - \mathbf{r}_3)}{\rho_\sigma(\mathbf{r}_2)}\right)(g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) - 1) \tag{8.54}$$

If we insert the expression 8.48 for the function $K_{i\sigma}$ in the above equations we obtain:

$$\frac{\delta g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} = S_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + D_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \tag{8.55}$$

where S_σ is a part we will call, for reasons to be explained in the next section, the steplike part and D_σ the deltafunction part given by:

$$S_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \sum_{ij}^{N_\sigma} -f_{i\sigma} f_{j\sigma} \frac{2}{\Delta\tilde{\epsilon}_{i\sigma}} \frac{\phi_{j\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_1)}{\rho_\sigma(\mathbf{r}_1)} \frac{\phi_{i\sigma}(\mathbf{r}_2)\phi_{j\sigma}^*(\mathbf{r}_2)}{\rho_\sigma(\mathbf{r}_2)} \sum_k^{N_\sigma} 2\xi_{ik}^\sigma \frac{|\phi_{k\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)}$$

$$+ \sum_{ij}^{N_\sigma} f_{i\sigma} f_{j\sigma} \frac{2}{\Delta\tilde{\epsilon}_{i\sigma}} \frac{\phi_{j\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_1)}{\rho_\sigma(\mathbf{r}_1)} \frac{\phi_{i\sigma}(\mathbf{r}_2)\phi_{j\sigma}^*(\mathbf{r}_2)}{\rho_\sigma(\mathbf{r}_2)} \sum_{kl}^{N_\sigma} \eta_{kl}^\sigma \left[\frac{|\phi_{k\sigma}(\mathbf{r}_1)|^2}{a_\sigma(\mathbf{r}_1)} + \frac{|\phi_{k\sigma}(\mathbf{r}_2)|^2}{a_\sigma(\mathbf{r}_2)}\right] \frac{|\phi_{l\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)} \tag{8.56}$$

and:

$$D_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \delta(\mathbf{r}_1 - \mathbf{r}_3)\left(h_\sigma(\mathbf{r}_1, \mathbf{r}_3) - \frac{g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) - 1}{\rho_\sigma(\mathbf{r}_1)}\right) + \delta(\mathbf{r}_2 - \mathbf{r}_3)\left(h_\sigma(\mathbf{r}_2, \mathbf{r}_3) - \frac{g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) - 1}{\rho_\sigma(\mathbf{r}_2)}\right) \tag{8.57}$$

where h_σ is defined as:

$$h_\sigma(\mathbf{r}_1, \mathbf{r}_2) = \frac{2}{a_\sigma(\mathbf{r}_1)} \sum_{ij}^{N_\sigma} \frac{f_{i\sigma} f_{j\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}} \frac{\phi_{j\sigma}(\mathbf{r}_1)\phi_{i\sigma}^*(\mathbf{r}_1)}{\rho_\sigma(\mathbf{r}_1)} \frac{\phi_{i\sigma}(\mathbf{r}_2)\phi_{j\sigma}^*(\mathbf{r}_2)}{\rho_\sigma(\mathbf{r}_2)} \tag{8.58}$$

As we made an approximation for the Greens function in the derivation of $\delta g_{s\sigma'}(\mathbf{r}_1, \mathbf{r}_2)/\delta\rho_\sigma(\mathbf{r}_3)$ one might ask whether this function satisfies constraint 8.25. One can easily verify that $S_\sigma(\mathbf{r}_1, \mathbf{r}_1, \mathbf{r}_3)$ is not equal to zero unless one uses the additional approximation of Krieger et al. [35] that $\Delta\tilde{\epsilon}_{i\sigma} = \Delta\tilde{\epsilon}_\sigma$ independent of i. In this approximation the expressions for the functions h_σ and S_σ simplify to:

$$h_\sigma(\mathbf{r}_1, \mathbf{r}_2) = -\frac{1}{\rho_\sigma(\mathbf{r}_1)} \frac{|\gamma_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)|^2}{\rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2)} = \frac{g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) - 1}{\rho_\sigma(\mathbf{r}_1)} \tag{8.59}$$

and

$$S_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{\gamma_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)}{\rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2)} \sum_i^{N_\sigma} \phi_{i\sigma}^*(\mathbf{r}_1)\phi_{i\sigma}(\mathbf{r}_2) \sum_k^{N_\sigma} 2\zeta_{ik}^\sigma \frac{|\phi_{k\sigma}(\mathbf{r}_3)|^2}{\rho_\sigma(\mathbf{r}_3)}$$

$$-\frac{|\gamma_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)|^2}{\rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2)} \sum_{kl}^{N_\sigma} \zeta_{kl}^\sigma \left[\frac{|\phi_{k\sigma}(\mathbf{r}_1)|^2}{\rho_\sigma(\mathbf{r}_1)} + \frac{|\phi_{k\sigma}(\mathbf{r}_2)|^2}{\rho_\sigma(\mathbf{r}_2)}\right] \frac{|\phi_{l\sigma}(\mathbf{r}_3)|^2}{\rho_\sigma(\mathbf{r}_3)} \tag{8.60}$$

where:

$$\zeta_{ik}^\sigma = (\bar{I} - \bar{M}^\sigma)^{-1} \tag{8.61}$$

for $i, k = 1, \ldots, N_\sigma - 1$ and zero otherwise. Matrix $\bar{M}^\sigma$ is defined as:

$$M_{ik}^\sigma = f_{k\sigma} \int \frac{|\phi_{i\sigma}(\mathbf{r})|^2 |\phi_{k\sigma}(\mathbf{r})|^2}{\rho_\sigma(\mathbf{r})} d\mathbf{r} \tag{8.62}$$

for $i, k = 1, \ldots, N_\sigma - 1$. The deltafunction part in this approximation D_σ unlike S_σ does not satisfy $D_\sigma(\mathbf{r}_1, \mathbf{r}_1, \mathbf{r}_3) = 0$, however it satisfies this constraint in an integral sense:

$$\int D_\sigma(\mathbf{r}_1, \mathbf{r}_1, \mathbf{r}_3) d\mathbf{r}_1 = 0 \tag{8.63}$$

Some properties of the function S_σ are readily derived from equation 8.60. Suppose that electrons at $\mathbf{r}_1$ and $\mathbf{r}_2$ are well-separated in different atomic shells with small overlap. In that case the one-particle density matrix $\gamma_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)$ will be small and consequently the function S_σ will also be small. If on the other hand the two electrons are close together within the same atomic shell then, because S_σ is exactly zero for $\mathbf{r}_1 = \mathbf{r}_2$, the function S_σ will also be very small. We can therefore expect that the largest contribution to $S_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ for a fixed position $\mathbf{r}_3$ is obtained if one of the electrons is well within one atomic shell and the other electron is at an atomic shell boundary. The behaviour of S_σ as a function of $\mathbf{r}_3$ is determined by factors of the form $|\phi_{k\sigma}(\mathbf{r}_3)|^2 / \rho_\sigma(\mathbf{r}_3)$ which is approximately constant if $\mathbf{r}_3$ is within atomic shell k. The contribution of this factor to the total function S_σ is determined by the constants ζ_{ik}^σ describing the coupling of the density perturbation in shell k with an electron in shell i. These constants are the largest if $i = k$. Plots of the function S_σ will be presented in section V for the case of the beryllium atom.

8.4 The step structure in the exchange potential

We now will derive an expression for the potential $v_{x,scr\sigma}^{resp}$ using the expression for the functional derivative of g_s derived in the previous section. We will for generality not yet make the additional approximation $\Delta \tilde{\epsilon}_{i\sigma} = \Delta \tilde{\epsilon}_\sigma$. We have:

$$v_{x,scr\sigma}^{resp}(\mathbf{r}_3) = \frac{1}{2} \int \frac{\rho_\sigma(\mathbf{r}_1) \rho_\sigma(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \frac{\delta g_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho_\sigma(\mathbf{r}_3)} d\mathbf{r}_1 d\mathbf{r}_2 \tag{8.64}$$

Inserting the expression for the functional derivative of g_s we find:

$$v_{x,scr\sigma}^{resp}(\mathbf{r}_3) = \frac{1}{2} \int \frac{\rho_\sigma(\mathbf{r}_1) \rho_\sigma(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} S_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) d\mathbf{r}_1 d\mathbf{r}_2 + \frac{1}{2} \int \frac{\rho_\sigma(\mathbf{r}_1) \rho_\sigma(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} D_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) d\mathbf{r}_1 d\mathbf{r}_2 \tag{8.65}$$

The part containing S_σ yields:

$$\frac{1}{2} \int \frac{\rho_\sigma(\mathbf{r}_1) \rho_\sigma(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} S_\sigma(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) d\mathbf{r}_1 d\mathbf{r}_2 = \sum_l^{N_\sigma} \frac{w_{l\sigma} |\phi_{l\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_3)} \tag{8.66}$$

where

$$w_{l\sigma} = \sum_{ij}^{N_\sigma} \int \frac{f_{i\sigma} f_{j\sigma}}{\Delta \tilde{\epsilon}_{i\sigma}} \frac{\phi_{j\sigma}(\mathbf{r}_1) \phi_{i\sigma}^*(\mathbf{r}_1) \phi_{i\sigma}(\mathbf{r}_2) \phi_{j\sigma}^*(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \left[-2\xi_{il}^\sigma + \sum_k^{N_\sigma} \eta_{kl}^\sigma \left(\frac{|\phi_{k\sigma}(\mathbf{r}_1)|^2}{a_\sigma(\mathbf{r}_1)} + \frac{|\phi_{k\sigma}(\mathbf{r}_2)|^2}{a_\sigma(\mathbf{r}_2)} \right) \right] d\mathbf{r}_1 d\mathbf{r}_2 \tag{8.67}$$

So the S_σ term leads to a steplike function given by equation 8.66. Within shell l this function is almost constant and equal to the value w_l given by equation 8.67. If we now denote:

$$v_{i\sigma}(\mathbf{r}_1) = \frac{1}{f_{i\sigma} \phi_{i\sigma}^*(\mathbf{r}_1)} \frac{\delta E_x}{\delta \phi_{i\sigma}(\mathbf{r}_1)} = -\frac{1}{\phi_{i\sigma}^*(\mathbf{r}_1)} \sum_k^{N_\sigma} \int f_{k\sigma} \frac{\phi_{i\sigma}^*(\mathbf{r}_2) \phi_{k\sigma}(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_2 \phi_{k\sigma}^*(\mathbf{r}_1) \tag{8.68}$$

which is equal to the orbital dependent potential within the Hartree-Fock approximation (except for the fact that we do not use Hartree-Fock orbitals) and further define

$$W_\sigma(\mathbf{r}_1) = \sum_i^{N_\sigma} -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}} \frac{v_{i\sigma}(\mathbf{r}_1)|\phi_{i\sigma}(\mathbf{r}_1)|^2}{a_\sigma(\mathbf{r}_1)} = \frac{\sum_i^{N_\sigma} \frac{1}{\Delta\tilde{\epsilon}_{i\sigma}} v_{i\sigma}(\mathbf{r}_1)\rho_{i\sigma}(\mathbf{r}_1)}{\sum_i^{N_\sigma} \frac{1}{\Delta\tilde{\epsilon}_{i\sigma}}\rho_{i\sigma}(\mathbf{r}_1)} \tag{8.69}$$

which is similar to the Slater potential [101] except for the energies $\Delta\tilde{\epsilon}_{i\sigma}$ and where

$$\rho_{i\sigma}(\mathbf{r}_1) = f_{i\sigma}|\phi_i(\mathbf{r}_1)|^2 \tag{8.70}$$

then we can write:

$$w_{l\sigma} = \sum_i^{N_\sigma} \frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}\xi_{il}^\sigma \bar{v}_{i\sigma} + \sum_i^{N_\sigma} \eta_{il}^\sigma \bar{W}_{i\sigma} = \sum_i^{N_\sigma-1} -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}(\bar{I} - \bar{N}^\sigma)_{il}^{-1}(\bar{W}_{i\sigma} - \bar{v}_{i\sigma}) \tag{8.71}$$

where

$$\bar{v}_{i\sigma} = \int v_{i\sigma}(\mathbf{r}_1)|\phi_{i\sigma}(\mathbf{r}_1)|^2 d\mathbf{r}_1 \tag{8.72}$$

and

$$\bar{W}_{i\sigma} = \int W_\sigma(\mathbf{r}_1)|\phi_{i\sigma}(\mathbf{r}_1)|^2 d\mathbf{r}_1 \tag{8.73}$$

Now the deltafunction part of $v_{x,scr\sigma}^{resp}$ gives:

$$\frac{1}{2}\int \frac{\rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}D_\sigma(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)d\mathbf{r}_1 d\mathbf{r}_2 = W_\sigma(\mathbf{r}_3) - v_{x,scr\sigma}(\mathbf{r}_3) \tag{8.74}$$

This is the difference of two long range potentials both having a Coulombic $-1/r$ behaviour and it therefore decays faster than Coulombic. In the approximation $\Delta\tilde{\epsilon}_{i\sigma} = \Delta\tilde{\epsilon}_\sigma$ this difference is exactly zero. So we find:

$$v_{x,scr\sigma}^{resp}(\mathbf{r}_1) = \sum_i^{N_\sigma-1} \frac{w_{i\sigma}|\phi_{i\sigma}(\mathbf{r}_1)|^2}{a_\sigma(\mathbf{r}_1)} + W_\sigma(\mathbf{r}_1) - v_{x,scr\sigma}(\mathbf{r}_1) \tag{8.75}$$

This gives for the exchange potential:

$$v_{x\sigma}(\mathbf{r}_1) = v_{x,scr}(\mathbf{r}_1) + v_{x,scr\sigma}^{resp}(\mathbf{r}_1) = W_\sigma(\mathbf{r}_1) + \sum_i^{N_\sigma-1} \frac{w_{i\sigma}|\phi_{i\sigma}(\mathbf{r}_1)|^2}{a_\sigma(\mathbf{r}_1)} \tag{8.76}$$

So we find from this equation:

$$\bar{v}_{x\sigma i} - \bar{W}_{i\sigma} = \sum_l^{N_\sigma-1} w_{l\sigma}\int \frac{|\phi_{i\sigma}(\mathbf{r}_1)|^2|\phi_{l\sigma}(\mathbf{r}_1)|^2}{a_\sigma(\mathbf{r}_1)}d\mathbf{r}_1 = -\frac{1}{2}\frac{\Delta\tilde{\epsilon}_{i\sigma}}{f_{i\sigma}}\sum_l^{N_\sigma} w_{l\sigma}N_{li}^\sigma \tag{8.77}$$

On the other hand it follows from equation 8.71 that

$$\sum_l^{N_\sigma-1} N_{lk}^\sigma w_{l\sigma} = \frac{2f_{k\sigma}}{\Delta\tilde{\epsilon}_{k\sigma}}(\bar{W}_{k\sigma} - \bar{v}_{k\sigma}) + w_{k\sigma} \tag{8.78}$$

So we obtain:

$$\bar{v}_{x\sigma i} - \bar{W}_{i\sigma} = -(\bar{W}_{i\sigma} - \bar{v}_{i\sigma}) - \frac{1}{2}\frac{\Delta\tilde{\epsilon}_{i\sigma}}{f_{i\sigma}}w_{i\sigma} \tag{8.79}$$

and we have:

$$w_{i\sigma} = -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}(\bar{v}_{x\sigma i} - \bar{v}_{i\sigma}) \tag{8.80}$$

So our final exchange potential is:

$$v_{x\sigma}(\mathbf{r}_1) = W_\sigma(\mathbf{r}_1) + \frac{\sum_i^{N_\sigma-1}\frac{1}{\Delta\tilde{\epsilon}_{i\sigma}}(\bar{v}_{x\sigma i} - \bar{v}_{i\sigma})\rho_{i\sigma}(\mathbf{r}_1)}{\sum_i^{N_\sigma}\frac{1}{\Delta\tilde{\epsilon}_{i\sigma}}\rho_{i\sigma}(\mathbf{r}_1)} \tag{8.81}$$

The first term in this equation is the term derived by Sharp and Horton [34]. The second term was derived by Krieger et al. [35] in the additional approximation $\Delta\tilde{\epsilon}_{i\sigma} = \Delta\tilde{\epsilon}_\sigma$. This additional approximation then leads to the exchange potential of Krieger, Li and Iafrate [35]:

$$v_{x\sigma}^{KLI}(\mathbf{r}_1) = \sum_i^{N_\sigma}\frac{v_{i\sigma}(\mathbf{r}_1)\rho_{i\sigma}(\mathbf{r}_1)}{\rho_\sigma(\mathbf{r}_1)} + \sum_i^{N_\sigma-1}\frac{(\bar{v}_{x\sigma i} - \bar{v}_{i\sigma})\rho_{i\sigma}(\mathbf{r}_1)}{\rho_\sigma(\mathbf{r}_1)} \tag{8.82}$$

If we multiply equation 8.81 by $\sum_i^{N_\sigma}(\Delta\tilde{\epsilon}_{i\sigma})^{-1}\rho_{i\sigma}(\mathbf{r}_1)$ and integrate over $\mathbf{r}_1$ one can easily show that our approximate $v_{x\sigma}$ satisfies:

$$\bar{v}_{x\sigma m} = \bar{v}_{m\sigma} \tag{8.83}$$

where $m = N_\sigma$ corresponds to the highest occupied Kohn-Sham orbital $\phi_{m\sigma}$. This equation is exactly valid within the so-called optimized potential model (OPM) [33, 36] exchange potential and also for the approximate exchange potential $v_{x\sigma}^{KLI}$ [35] which follows directly from our derivation by putting $\Delta\tilde{\epsilon}_{i\sigma} = \Delta\tilde{\epsilon}_\sigma$.

8.5 Results

In this section we present some numerical results obtained from the optimized potential model (OPM) [33, 35, 36]. The OPM-orbitals and exchange potential $v_{x\sigma}^{OPM}([\{\rho_\sigma\}];\mathbf{r}) = v_{x\sigma}([\{\rho_\sigma^{OPM}\}];\mathbf{r})$ are obtained by putting the correlation functional equal to zero and minimizing the exchange-only Kohn-Sham energy functional. This leads to an integral equation for the exchange potential which is equal to the functional derivative of the exact exchange functional evaluated at the OPM electron density ρ_σ^{OPM}.

In figure 2 we plot the function $g_{s\sigma}^{OPM}(r_1, r_2) - 1$ for the case of the beryllium atom. As in the Be atom only s-shells are occupied this function only depends on the radial distance $r_1 = |\mathbf{r}_1|$ and $r_2 = |\mathbf{r}_2|$ of electrons 1 and 2 from the atomic nucleus and not on the angle between vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. This is a convenient feature for analysis. As Be is a closed shell atom $g_{s\sigma}$ is equal for up and down spin $g_{s\uparrow} = g_{s\downarrow}$. From figure 2 we can see that $g_{s\sigma}(r_1, r_2) - 1$ is close to -1 if r_1 are r_2 within the same atomic shell. This is an effect of the Pauli-principle, the probability that two electrons of the same spin are close together is small. If the two electrons are in different atomic shells (the boundary between the 1s-shell and 2s-shell is at a radial distance of about one bohr) then $g_{s\sigma}(r_1, r_2) - 1$ is close to zero.

In figure 3 we plot the S_σ part of equation 8.60 of the functional derivative $\delta g_{s\sigma}(r_1, r_2)/\delta\rho_\sigma(r_3)$ as a function of r_2 and $r_3 = |\mathbf{r}_3|$ for the Be atom where $r_1 = 0.1$ bohr which is well within the 1s-shell. As we can see from this figure S_σ is small when the other electron at radial distance r_2 is either situated in the 1s-shell or in the 2s-shell. This function only becomes large when electron 2 crosses the boundary between the 1s- and the 2s-shell at a radial distance from the atomic nucleus of about 1 bohr. However this function is then only large when r_3 is also located within the 1s-shell

Figure 8.2: The screening factor $g_{s\sigma}(r_1, r_2) - 1$ as a function of the radial distance r_1 and r_2 to the atomic nucleus of electron 1 and 2 for the beryllium atom

Figure 8.3: The functional derivative $\delta g_{s\sigma}(r_1, r_2)/\delta \rho_\sigma(r_3)$ as a function of the radial distance r_2 and r_3 to the atomic nucleus. Electron 1 is located at a distance $r_1 = 0.1$ bohr from the atomic nucleus

Figure 8.4: The exchange potential with corresponding screening and screening response parts as calculated within the optimized potential model (OPM) and within the Krieger-Li-Iafrate (KLI) approximation for the beryllium atom

Figure 8.5: The exchange potential with corresponding screening and screening response parts as calculated within the optimized potential model (OPM) and within the Krieger-Li-Iafrate (KLI) approximation for the krypton atom

and within this shell S_σ is almost constant as a function of r_3 clearly displaying the step structure.

We now turn to the exchange potential. In figure 4 we plot $v_{x,scr\sigma}^{OPM}$ and $v_{x,scr\sigma}^{resp,OPM}$ and their sum and the corresponding ones $v_{x,scr\sigma}^{KLI}$ and $v_{x,scr\sigma}^{resp,KLI}$ from selfconsistent solution of equation 8.82 for the beryllium atom. The terms $v_{x,scr\sigma}^{KLI}$ and $v_{x,scr\sigma}^{resp,KLI}$ are defined as the first and the second term of equation 8.82. In figure 5 we present the same quantities for the krypton atom. As we can see from these figures the $v_{x,scr\sigma}^{OPM}$ and $v_{x,scr\sigma}^{KLI}$ are so close that they can not be distinguised on the scale presented. Also the step potentials $v_{x,scr\sigma}^{resp,OPM}$ and $v_{x,scr\sigma}^{resp,KLI}$ are very close, the largest difference being constant within the 1s-shell. The most striking difference between the OPM and KLI exchange potential is the smoothing of the intershell peak at the atomic shell boundaries in the KLI exchange potential. This difference is not the most important difference from an energetical point of view, as for the energy the atomic core region is the most important. This is most easily seen from the Levy-Perdew relation [22]:

$$E_x[\{\rho_\sigma\}] = \sum_\sigma \int v_{x\sigma}(\mathbf{r})(3\rho_\sigma(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho_\sigma(\mathbf{r}))d\mathbf{r} \qquad (8.84)$$

which relates the exchange potential to the exchange energy. This relation is not exactly satisfied by the KLI exchange potential as an approximation is made to the functional derivative of E_x but the error is within 1% [35]. For approximate exchange potentials based on separate approximations for the $v_{x,scr\sigma}$ and the $v_{x,scr\sigma}^{resp}$ part it is therefore important to obtain accurate values for the steps in $v_{x,scr\sigma}^{resp}$ in the atomic core region.

8.6   Conclusions

In this paper we analyzed the structure of the atomic Kohn-Sham potential. The exchange-correlation potential was written as the sum of two terms, one term containing the coupling constant integrated pair-correlation function which represents the long-range potential of the exchange-correlation hole, and one term, containing the functional derivative of the coupling constant integrated pair-correlation function, which is short ranged and displays a distinct atomic shell structure. An explicit expression for the exchange part of this functional derivative was derived using an approximation for the Greens function of the Kohn-Sham system. Properties of this function are analyzed and plots are presented for some atomic systems. It is shown that $\delta g_{s\sigma}(\mathbf{r}_1,\mathbf{r}_2)/\delta\rho_\sigma(\mathbf{r}_3)$ is small when electrons at $\mathbf{r}_1$ and $\mathbf{r}_2$ are close together or when they are in the middle of the same or different atomic shells. This function is the largest when one of the electrons at $\mathbf{r}_1$ or $\mathbf{r}_2$ crosses an atomic shell boundary and as a function of $\mathbf{r}_3$ it is proportional to $\rho_{i\sigma}(\mathbf{r}_3)/\rho_\sigma(\mathbf{r}_3)$ which is the electron density of atomic shell i in which point $\mathbf{r}_3$ is located divided by the total electron density. As a function of $\mathbf{r}_3$ we therefore see a steplike behaviour, $\delta g_{s\sigma}/\delta\rho_\sigma(\mathbf{r}_3)$ is constant within the atomic shells and changes rapidly at the shell boundaries. This behaviour is induced in the short range part of the exchange potential. The other part of the exchange potential is monotonous, has a Coulombic long range behaviour and does not show any distinct atomic shell structure [108]. The natural splitting of the exchange and exchange-correlation potential in these two parts has important implications for obtaining accurate approximations for the Kohn-Sham potential. Different strategies can be developed to approximate the long and the short range part separately. This can improve one-electron energies, the ionisation energy and density dependent quantities such as dipole moments. A full discussion of the this is deferred to a subsequent paper [130].

8.7 Appendix:The inverse density response function

We will derive the expression for the inverse density response function $\chi_{s\sigma}^{-1}$. We first split up $\chi_{s\sigma}^{-1}$ as:

$$\chi_{s\sigma}^{-1}(\mathbf{r}_2, \mathbf{r}_3) = \frac{\delta(\mathbf{r}_2 - \mathbf{r}_3)}{a_\sigma(\mathbf{r}_2)} + c_\sigma(\mathbf{r}_2, \mathbf{r}_3) \tag{8.85}$$

where c_σ is a function to be determined. For this function we find using equation 8.39 the following equation:

$$\delta(\mathbf{r}_1 - \mathbf{r}_3) = \int \chi_{s\sigma}(\mathbf{r}_1, \mathbf{r}_2) \chi_{s\sigma}^{-1}(\mathbf{r}_2, \mathbf{r}_3) d\mathbf{r}_2 =$$

$$\delta(\mathbf{r}_1 - \mathbf{r}_3) - \frac{b_\sigma(\mathbf{r}_1, \mathbf{r}_3)}{a_\sigma(\mathbf{r}_3)} + a_\sigma(\mathbf{r}_1)c_\sigma(\mathbf{r}_1, \mathbf{r}_3) - \int b_\sigma(\mathbf{r}_1, \mathbf{r}_2)c_\sigma(\mathbf{r}_2, \mathbf{r}_3)d\mathbf{r}_2 \tag{8.86}$$

So we find:

$$c_\sigma(\mathbf{r}_1, \mathbf{r}_3) = \frac{b_\sigma(\mathbf{r}_1, \mathbf{r}_3)}{a_\sigma(\mathbf{r}_1)a_\sigma(\mathbf{r}_3)} + \frac{1}{a_\sigma(\mathbf{r}_1)} \int b_\sigma(\mathbf{r}_1, \mathbf{r}_2)c_\sigma(\mathbf{r}_2, \mathbf{r}_3)d\mathbf{r}_2 \tag{8.87}$$

If we further define:

$$d_\sigma(\mathbf{r}_1, \mathbf{r}_3) = a_\sigma(\mathbf{r}_1)a_\sigma(\mathbf{r}_3)c_\sigma(\mathbf{r}_1, \mathbf{r}_3) \tag{8.88}$$

we have

$$d_\sigma(\mathbf{r}_1, \mathbf{r}_3) = b_\sigma(\mathbf{r}_1, \mathbf{r}_3) + \int \frac{b_\sigma(\mathbf{r}_1, \mathbf{r}_2)d_\sigma(\mathbf{r}_2, \mathbf{r}_3)}{a_\sigma(\mathbf{r}_2)}d\mathbf{r}_2 \tag{8.89}$$

If we insert the expression 8.41 for b_σ we find:

$$d_\sigma(\mathbf{r}_1, \mathbf{r}_3) = \sum_i^{N_\sigma} -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}|\phi_{i\sigma}(\mathbf{r}_1)|^2|\phi_{i\sigma}(\mathbf{r}_3)|^2 + \sum_i^{N_\sigma} -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}|\phi_{i\sigma}(\mathbf{r}_1)|^2\alpha_{i\sigma}(\mathbf{r}_3) \tag{8.90}$$

where

$$\alpha_{i\sigma}(\mathbf{r}_3) = \int \frac{|\phi_{i\sigma}(\mathbf{r}_2)|^2 d_\sigma(\mathbf{r}_2, \mathbf{r}_3)}{a_\sigma(\mathbf{r}_2)}d\mathbf{r}_2 \tag{8.91}$$

If we insert into the above formula the expression 8.90 for the function d_σ we obtain a system of N_σ equations for the functions $\alpha_{i\sigma}$:

$$\alpha_{i\sigma}(\mathbf{r}_3) = \sum_k^{N_\sigma} N_{ik}^\sigma(|\phi_{k\sigma}(\mathbf{r}_3)|^2 + \alpha_{k\sigma}(\mathbf{r}_3)) \tag{8.92}$$

where

$$N_{ik}^\sigma = -\frac{2f_{k\sigma}}{\Delta\tilde{\epsilon}_{k\sigma}} \int \frac{|\phi_{i\sigma}(\mathbf{r}_2)|^2|\phi_{k\sigma}(\mathbf{r}_2)|^2}{a_\sigma(\mathbf{r}_2)}d\mathbf{r}_2 \tag{8.93}$$

A general solution to the equations 8.92 can now be found. As the functions $|\phi_{i\sigma}|^2$ are only finite in number the most general expression for the $\alpha_{i\sigma}$ is a linear combination of N_σ of these functions:

$$\alpha_{i\sigma}(\mathbf{r}_3) = \sum_k^{N_\sigma} \beta_{ik}^\sigma |\phi_{k\sigma}(\mathbf{r}_3)|^2 \tag{8.94}$$

This immediately gives an equation for the matrix β_{ik}^σ:

$$\beta_{ik}^\sigma = N_{ik}^\sigma + \sum_l^{N_\sigma} N_{il}^\sigma \beta_{lk}^\sigma \tag{8.95}$$

or if we denote the matrices by $\bar{\beta}^\sigma$ and $\bar{N}^\sigma$:

$$\bar{\beta}^\sigma = \bar{N}^\sigma (\bar{I} - \bar{N}^\sigma)^{-1} = (\bar{I} - \bar{N}^\sigma)^{-1} \bar{N}^\sigma = (\bar{I} - \bar{N}^\sigma)^{-1} - \bar{I} \tag{8.96}$$

The above equations assume that the inverse of $\bar{I} - \bar{N}^\sigma$ exists. This is necessary to find a unique solution for the coefficients β_{ik}^σ. However as discussed the inverse $\chi_{s\sigma}^{-1}$ is only defined up to a constant and, as discussed below $(\bar{I} - \bar{N}^\sigma)^{-1}$ does not exist, we cannot specify $\chi_{s\sigma}^{-1}$ uniquely unless we make a special choice for this arbitrary constant. We return to this point after we have obtained our final expression for $\chi_{s\sigma}^{-1}$. So we find:

$$d_\sigma(\mathbf{r}_1, \mathbf{r}_3) = \sum_{ik}^{N_\sigma} \eta_{ik}^\sigma |\phi_{i\sigma}(\mathbf{r}_1)|^2 |\phi_{k\sigma}(\mathbf{r}_3)|^2 \tag{8.97}$$

where the η_{ik}^σ are given by:

$$\eta_{ik}^\sigma = -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}(\delta_{ik} + \beta_{ik}^\sigma) \tag{8.98}$$

where β_{ik}^σ is one of the solutions of equation 8.95. We then have our final expression for the inverse density response function which is however still not unique:

$$\chi_{s\sigma}^{-1}(\mathbf{r}_2, \mathbf{r}_3) = \frac{\delta(\mathbf{r}_2 - \mathbf{r}_3)}{a_\sigma(\mathbf{r}_2)} + \sum_{ik}^{N_\sigma} \eta_{ik}^\sigma \frac{|\phi_{i\sigma}(\mathbf{r}_2)|^2 |\phi_{k\sigma}(\mathbf{r}_3)|^2}{a_\sigma(\mathbf{r}_2) a_\sigma(\mathbf{r}_3)} \tag{8.99}$$

Now we adress the question of the invertability of $\bar{I} - \bar{N}^\sigma$. From equation 8.93 it follows that:

$$\sum_k^{N_\sigma} N_{ik}^\sigma = 1 \tag{8.100}$$

so:

$$\sum_k^{N_\sigma} (\bar{I} - \bar{N}^\sigma)_{ik} = 0 \tag{8.101}$$

From this equation it follows that the columns of matrix $\bar{I} - \bar{N}^\sigma$ are linearly dependent and so this matrix has no inverse. Consequently the approximate density response function $\chi_{s\sigma}$ has no unique inverse. As discussed this is due to the fact that inverse is only determined up to an arbitrary constant. From the equation of the inverse density response function 8.99 one can see that $\chi_{s\sigma}^{-1}(\mathbf{r}_1, \mathbf{r}_2)$ goes to a constant if $\mathbf{r}_1$ or $\mathbf{r}_2$ goes to infinity. This constant is determined by the orbital density of the highest occupied Kohn-Sham orbital $|\phi_{N_\sigma}|^2$ as it has the slowest exponential decay of all orbital densities. We can therefore fix this constant by requiring that:

$$\lim_{|\mathbf{r}_1| \to \infty} \chi_{s\sigma}^{-1}(\mathbf{r}_1, \mathbf{r}_2) = 0 \tag{8.102}$$

This property can be satisfied by restricting the sum over the orbital densities in the equation 8.41 for the function b_σ to the first $N_\sigma - 1$ orbitals. In this approximation $\chi_{s\sigma}$ has an inverse and the

derivation of this appendix can be carried out in a similar way. One finds for $\chi_{s\sigma}^{-1}$ expression 8.99 in which N_σ is replaced with $N_\sigma - 1$ with matrix η_{ik}^σ given by:

$$\eta_{ik}^\sigma = -\frac{2f_{i\sigma}}{\Delta\tilde{\epsilon}_{i\sigma}}(\bar{I} - \bar{N}^\sigma)_{ik}^{-1} \tag{8.103}$$

which is a matrix of dimension $N_\sigma - 1$. Note that the approximation used here is equivalent to the approximation used by Krieger for the second part of the exchange potential in equation 8.82 where the summation is restricted to $N_\sigma - 1$ orbitals. This then leads to the fulfilment of the exact constraint 8.83 by the approximate exchange potential.

Chapter 9

A self-consistent approximation to the Kohn-Sham exchange potential

A scheme of approximation of the Kohn-Sham exchange potential v_x has been proposed, making use of a partitioning of v_x into the long-range Slater v_S and the short-range response v_{resp} components. The model v_{resp}^{mod} has been derived from the dimensional arguments, which possesses the proper short-range behaviour and the characteristic atomic-shell stepped structure. When combined with the accurate v_S, v_{resp}^{mod} provides an excellent approximation to v_x^{OPM}. With the GGA approximation to v_S, v_{resp}^{mod} provides an efficient DFT approach which, for the first time, fits closely the form of the accurate exchange potential and yields reasonably accurate exchange and total energies as well as the energies of the highest occupied orbital.

9.1   Introduction

Efficient approximation of the Kohn-Sham exchange potential v_x remains one of the important problems of the density functional theory (DFT) [35, 49, 131, 32, 121]. v_x is defined in the DFT as a functional derivative of the exchange energy $E_x[\rho]$ with respect to the density $\rho(\mathbf{r})$

$$v_x(\mathbf{r}) = \frac{\delta E_x[\rho]}{\delta \rho(\mathbf{r})} \tag{9.1}$$

In the one-electron Kohn-Sham equations (Hartree atomic units will be used throughout the paper)

$$\left[-\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_x(\mathbf{r}) + v_c(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{9.2}$$

v_x represents the local effect of exchange, the dominant part of the electron correlation, while the external v_{ext} and the Hartree v_H potentials represent the electrostatic interaction and v_c is the effective potential of the electron Coulomb correlation.

Figure 1 illustrates the typical behaviour of v_x with the exchange potential of the optimized potential model (OPM) [33, 91, 37] for Mg. v_x is finite at the nucleus position

$$v_x(r) = \text{const} \qquad (r = 0) \tag{9.3}$$

it has the Coulombic asymptotics [33, 35, 57, 31]

$$v_x(\mathbf{r}) \to -\frac{1}{|\mathbf{r}|} \qquad (|\mathbf{r}| \to \infty) \tag{9.4}$$

Figure 9.1: The OPM exchange potential and its components for Mg

and it possesses the shell structure, i.e. a non-monotonous dependence on r with the small peaks in the regions between the atomic shells [35, 33, 50, 93]. Taken together, these features of v_x present considerable difficulties for its efficient approximation.

Recent progress of the DFT is due to the generalized gradient approximation (GGA) [50, 95, 44, 43, 47, 96] that gives a remarkable improvement to the exchange-correlation energies of the local density approximation (LDA) [11, 97, 98, 99]. It was shown [49, 32, 121, 92, 94], however, that the standard GGA exchange potentials do not possess the proper asymptotic behaviour and shell structure. The corresponding gradient corrections add little to the LDA exchange potential.

For the analysis [108] and approximation [35, 121] of v_x it appears very useful to separate the Slater potential v_S [101]

$$v_S(\vec{r}_1) = \int \frac{\rho(\vec{r}_2)[g_x([\rho];\vec{r}_1,\vec{r}_2)-1]}{|\vec{r}_1-\vec{r}_2|}d\vec{r}_2 \tag{9.5}$$

as an individual part of v_x. In equation 9.5 g_x is the exchange pair-correlation function, which can be expressed in terms of the occupied Kohn-Sham orbitals $\phi_i(\mathbf{r})$

$$g_x([\rho];\mathbf{r}_1,\mathbf{r}_2) = 1 - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\phi_i(\mathbf{r}_1)\phi_i^*(\mathbf{r}_2)\phi_j^*(\mathbf{r}_1)\phi_j(\mathbf{r}_2)}{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)} \tag{9.6}$$

The separation of v_S follows naturally from the expression of E_x in terms of v_S

$$E_x = \frac{1}{2} \int \rho(\mathbf{r}) v_S(\mathbf{r}) d\mathbf{r} \tag{9.7}$$

Differentiation of equation 9.7 with a proper account of 9.5 represents v_x as a sum of v_S and the additional potential v_{resp}, an integral of the linear "response" of g, $\frac{\delta g(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho(\mathbf{r}_3)}$ [6]

$$v_x(\mathbf{r}) = v_S(\mathbf{r}) + v_{resp}(\mathbf{r}) \tag{9.8}$$

The potentials v_S and v_{resp} have rather distinct, characteristic behaviour [121, 108] and Figure 1 clearly illustrates this with the v_S^{OPM} and v_{resp}^{OPM} components of the OPM exchange potential v_x^{OPM} [33, 91, 37]. v_S is an attractive potential, which accumulates the Coulombic asymptotics (eqn. 9.4) of the total potential v_x

$$v_S(\mathbf{r}) \to -\frac{1}{|\mathbf{r}|} \qquad (|\mathbf{r}| \to \infty) \tag{9.9}$$

It is rather smooth and does not display the pronounced shell structure.

Contrary to this, v_{resp} is repulsive and short-range. Remarkably enough, it exhibits very clear step function behaviour [108]. As it has been shown in [132], the steps of v_{resp} originate from the corresponding stepped structure of the "response" $\frac{\delta g(\mathbf{r}_1, \mathbf{r}_2)}{\delta \rho(\mathbf{r}_3)}$ as a function of $\mathbf{r}_3$. One can see from Figure 1 that the abovementioned small intershell peaks of v_x are evidently built in by the superposition of the stepped form of v_{resp} on the smooth v_S.

This specialized behaviour makes desirable the modelling of v_x with the direct individual approximation of v_S and v_{resp}. In this paper a model v_{resp}^{mod} is derived from the dimensional arguments, which represents v_{resp} as the statistical average of the orbital energy contributions. v_{resp}^{mod} possesses the properties and closely reproduces the behaviour of the accurate potential. When combined with the accurate functional v_S, v_{resp}^{mod} provides an excellent approximation to v_x^{OPM}. A more practical approximation is obtained with v_{resp}^{mod} and the GGA [44] approximation to v_S. Contrary to the standard GGA schemes, the latter combination provides both the proper form and eigenvalues of the one-electron potential and the reasonable estimate of atomic E_x and the total energies E_{tot}.

9.2 A model potential v_{resp}^{mod}

As a starting point for our model we use an approximate equation of Krieger, Li and Iafrate (KLI) [35] for the OPM exchange potential v_x^{OPM}. OPM neglects the electron Coulomb correlation and its Kohn-Sham equations have the form [33, 35]

$$\left[-\frac{1}{2} \nabla^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_x(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{9.10}$$

The ground-state OPM wavefunction is the Slater determinant of the eigenfunctions ϕ_i of 9.10, so that the components v_S^{OPM} and v_{resp}^{OPM} of v_x^{OPM} have the form 9.5 and 9.8 with g_x built from these functions. In [35] the very accurate approximation to v_x^{OPM} was defined with the following equation

$$v_x^{KLI}(\mathbf{r}) = v_S(\mathbf{r}) + \sum_{i=1}^{N} w_i \frac{|\phi_i(\mathbf{r})|^2}{\rho(\mathbf{r})} \tag{9.11}$$

In equation 9.11 v_S is the exact functional 9.5,9.6 and the second term is the statistical average of the orbital contributions w_i, the latter being the difference between the expectation values of the potential 9.11 and the Hartree-Fock exchange operator v_{xi} for the orbital ϕ_i

$$w_i = \int |\phi_i(\mathbf{r})|^2 [v_x^{OPM}(\mathbf{r}) - v_{xi}(\mathbf{r})] d\mathbf{r} \tag{9.12}$$

$$v_{xi}(\vec{r}_1) = -\frac{1}{\phi_i(\vec{r}_1)} \sum_{j=1}^{N} \phi_j^*(\vec{r}_1) \int \frac{\phi_i^*(\vec{r}_2)\phi_j(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d\vec{r}_2 \tag{9.13}$$

Note, that for the highest occupied orbital ϕ_N the expectation values of v_x^{OPM} and v_{xi} are equal [35], thus providing the zero value of the corresponging parameter w_N

$$w_N = 0 \tag{9.14}$$

The second term of equation 9.11 offers a promising form for the model v_{resp}^{mod}. Because of the integral kernel 9.13, straightforward evaluation of 9.12 requires laborious calculation of the two-electron integrals with the orbitals ϕ_i. However, with the suitable approximation for w_i one can develop an efficient model

$$v_{resp}^{mod}(\mathbf{r}) = \sum_{i=1}^{N} w_i \frac{|\phi_i(\mathbf{r})|^2}{\rho(\mathbf{r})} \tag{9.15}$$

w_i can be calculated, for example, as the orbital expectation values of some local potential v_w

$$w_i = \int |\phi_i(\mathbf{r})|^2 v_w(\mathbf{r}) d\mathbf{r} \tag{9.16}$$

chosen as the best local approximation of the difference $[v_x(\mathbf{r}) - v_{xi}(\mathbf{r})]$. In this paper an alternative approach is presented. We propose to use the model 9.15 with the orbital contributions w_i being approximated by a function of the orbital energies of eq.2. The form of this function is chosen to provide the gauge invariance, the proper scaling and the short-range behaviour of v_{resp}^{mod}.

According to the gauge invariance requirement, addition of a constant to the eigenvalues ϵ_i should not alter w_i values. To satisfy this requirement, we choose w_i to be a function of the difference $(\mu - \epsilon_i)$

$$w_i = f(\mu - \epsilon_i) \tag{9.17}$$

where μ is the Fermi level of a given system, which is equal to the one-electron energy of the highest occupied orbital, $\mu = \epsilon_N$.

The exchange potential 9.1 and its components v_S and v_{resp} have the following scaling property

$$v_x([\rho_\lambda]; \mathbf{r}) = \lambda v_x([\rho]; \lambda \mathbf{r}) \tag{9.18}$$

where

$$\rho_\lambda(\mathbf{r}) = \lambda^3 \rho(\lambda \mathbf{r}) \tag{9.19}$$

while ϵ_i has the scaling property

$$\epsilon_i[\rho_\lambda] = \lambda^2 \epsilon_i[\rho(\mathbf{r})] \tag{9.20}$$

To provide 9.18, the function f from eq.24 should scale as follows

$$f(\lambda^2(\mu - \epsilon_i)) = \lambda f(\mu - \epsilon_i) \tag{9.21}$$

and so we find the square root of $(\mu - \epsilon_i)$ to be the proper scaling function f

$$w_i = f(\mu - \epsilon_i) = K[\rho]\sqrt{\mu - \epsilon_i} \tag{9.22}$$

By the definition, 9.22 satisfies the condition 9.14. Owing to this, the highest occupied orbital ϕ_N does not contribute to the numerator of 9.15, thus providing the short-range behaviour of v_{resp}^{mod}

$$v_{resp}^{mod}(\mathbf{r}) = K[\rho] \sum_{i=1}^{N} \sqrt{\mu - \epsilon_i} \frac{|\phi_i(\mathbf{r})|^2}{\rho(\mathbf{r})} \tag{9.23}$$

As a result, our model potential 9.23 possesses the gauge invariance, the proper scaling and the short-range behaviour.

$K[\rho]$ in equation 9.23 is a numerical coefficient, which can be determined from the homogeneous electron gas model. For the gas of a density ρ the exact v_{resp} of eq.9 has the form

$$v_{resp} = \frac{k_F}{2\pi} \tag{9.24}$$

where k_F is the Fermi wavevector

$$k_F = (3\pi^2 \rho)^{\frac{1}{3}} \tag{9.25}$$

Putting v_{resp}^{mod} of eq.30 to be equal to 9.24, one can calculate K_g. For the homogeneous electron gas the Kohn-Sham orbitals and eigenvalues of eq.2 are given by

$$\phi_{\vec{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\vec{k}\mathbf{r}} \tag{9.26}$$

where V is a volume of the system and

$$\epsilon_{\vec{k}} = \frac{k^2}{2} + v_x[\rho] + v_c[\rho] \tag{9.27}$$

The Fermi level is given by

$$\mu = \frac{k_F^2}{2} + v_x[\rho] + v_c[\rho] \tag{9.28}$$

Inserting the above expression in equation 9.23, we obtain

$$v_{resp}^{mod} = \frac{K_g}{\sqrt{2}\rho V} \sum_{|\vec{k}|<k_F} \sqrt{k_F^2 - k^2} \tag{9.29}$$

A replacement of the sum in equation 9.29 by an integral yields

$$v_{resp}^{mod}(\mathbf{r}) = \frac{K_g}{\sqrt{2}(2\pi)^3 \rho} \int_0^{k_F} \sqrt{k_F^2 - k^2} \, 4\pi k^2 dk =$$

$$= \frac{K_g k_F^4}{2\sqrt{2}(2\pi)^2 \rho} \int_0^1 \sqrt{1-x^2} x^2 dx = \frac{3\pi K_g}{16\sqrt{2}} k_F \tag{9.30}$$

From equations 9.24 and 9.30 the K value is defined by

$$K_g = \frac{8\sqrt{2}}{3\pi^2} \approx 0.382 \tag{9.31}$$

which is valid for the homogeneous electron gas of an arbitrary density.

9.3   A self-consistent scheme with v_{resp}^{mod}

We propose to use the model 9.23 within the self-consistent scheme, in which the Kohn-Sahm equations 9.10 are solved with the following approximate exchange potential

$$v_x^{mod}(\mathbf{r}) = v_S(\mathbf{r}) + v_{resp}^{mod}(K, \mathbf{r}) \tag{9.32}$$

where v_S is the Slater potential 9.5 or its suitable approximation. The resulting Kohn-Sham orbitals are used to calculate the total energy, with the exchange energy being calculated from equation 9.7 with the self-consistent potential v_S. The electron-gas value K_g of equation 9.31 can be chosen as the universal parameter of v_{resp}^{mod} for all systems. Another option is to determine K self-consistently from the requirement, that the Levy-Perdew relation [22]

$$E_x = \int [3\rho(\mathbf{r}) + \mathbf{r}\nabla\rho(\mathbf{r})] v_x^{mod}(\mathbf{r}) d\mathbf{r} \tag{9.33}$$

should yield the same value of E_x as eq. 7 with the potential v_S

$$E_x = \int [3\rho(\mathbf{r}) + \mathbf{r}\nabla\rho(\mathbf{r})][v_S(\mathbf{r}) + K_{sc}[\rho]R(\mathbf{r})]d\mathbf{r} = \frac{1}{2} \int \rho(\mathbf{r})v_S(\mathbf{r})d\mathbf{r} \tag{9.34}$$

$$R(\mathbf{r}) = \sum_{i=1}^{N} \sqrt{\mu - \epsilon_i} \frac{|\phi_i(\mathbf{r})|^2}{\rho(\mathbf{r})} \tag{9.35}$$

From the requirement 9.34 it follows the expression for $K[\rho]$

$$K_{sc}[\rho] = -\frac{I_1}{I_2} \tag{9.36}$$

$$I_1 = \int [\frac{5}{2}\rho(\mathbf{r}) + \mathbf{r}\nabla\rho(\mathbf{r})]v_S d\mathbf{r} \tag{9.37}$$

$$I_2 = \int [3\rho(\mathbf{r}) + \mathbf{r}\nabla\rho(\mathbf{r})]R(\mathbf{r})]d\mathbf{r} \tag{9.38}$$

In order to develop an efficient DFT scheme with 9.32, an appropriate approximation to v_S of eq.9.5 is needed. One can use, for example, the weighted-density approximation (WDA) [125, 133, 123, 126] for the pair-correlation function g_x, which guarantees the correct asymptotics 9.9 of v_S. In this paper we use a more practical model of v_S, which is obtained from the GGA of Becke [44, 121]. In [44] the exchange energy density $\epsilon_x(\rho; \mathbf{r})$

$$E_x = \int \epsilon_x(\rho; \mathbf{r})d\mathbf{r} \tag{9.39}$$

of the second order of gradient expansion

$$\epsilon_x^{approx}(\rho; \mathbf{r}) = -\rho^{\frac{4}{3}}(\mathbf{r})[\frac{3}{2}(\frac{3}{8\pi})^{\frac{1}{3}} + a_0\xi] \tag{9.40}$$

$$\xi = 2^{\frac{2}{3}}(\frac{\nabla\rho(\mathbf{r})}{\rho^{\frac{4}{3}}(\mathbf{r})})^2 \tag{9.41}$$

is modified with the correction function f of the argument $\xi^{\frac{1}{2}}\sinh^{-1}\xi^{\frac{1}{2}}$

$$\epsilon_x^{GGA}(\rho; \mathbf{r}) = -\rho^{\frac{4}{3}}(\mathbf{r})[\frac{3}{2}(\frac{3}{8\pi})^{\frac{1}{3}} + a_0 f(\xi^{\frac{1}{2}}\sinh^{-1}\xi^{\frac{1}{2}})\xi] \tag{9.42}$$

ATOM	Be	Ne	Mg	Ar	Ca	Zn	Kr	Sr	Cd	Xe
$v_x^{mod(S)}$	0.305	0.342	0.384	0.365	0.389	0.381	0.381	0.397	0.388	0.386
$v_x^{mod(GGA)}$	0.803	0.518	0.536	0.463	0.478	0.440	0.440	0.453	0.432	0.428

Table 9.1: Values of the parameter K_{sc} obtained within the self-consistent scheme for potentials $v_x^{mod(S)}$ and $v_x^{mod(GGA)}$

$$f(\xi^{\frac{1}{2}}\sinh^{-1}\xi^{\frac{1}{2}}) = \frac{1}{1 + 6a_0\xi^{\frac{1}{2}}\sinh^{-1}\xi^{\frac{1}{2}}} \tag{9.43}$$

Using equations 9.7 and 9.39, one can derive from relation 9.42 the corresponding approximation for v_S

$$v_S^{GGA}(\mathbf{r}) = \frac{2\epsilon_x^{GGA}(\rho;\mathbf{r})}{\rho(\mathbf{r})} = -\rho^{\frac{4}{3}}(\mathbf{r})[3(\frac{3}{8\pi})^{\frac{1}{3}} + 2a_0 f(\xi^{\frac{1}{2}}\sinh^{-1}\xi^{\frac{1}{2}})\xi] \tag{9.44}$$

This potential has the proper scaling 9.18, the correct asymptotics 9.9 and for the fitted value $a_0 = 0.0042$ it yields rather accurate estimate of E_x via the integral 9.7. The results of the self-consistent calculations within the proposed scheme will be discussed in the next sections.

9.4   A comparison of the self-consistent potentials

In order to test the proposed self-consistent scheme, the comparative exchange-only atomic calculations have been performed for the closed-shell atoms Be, Ne, Mg, Ar, Ca, Zn, Kr, Sr, Cd, Xe. The exchange potentials v_x compared can be subdivided into two groups, depending on whether the accurate functional v_S 9.5 or its GGA approximation v_S^{GGA} 9.44 is used as the Slater potential in 9.32. The first group includes the single v_S (i.e. in this case v_{resp} is neglected in 9.32), $v_x^{mod(S)}$, which is a sum of v_S and v_{resp}^{mod} calculated in both variants with K_{sc} and K_g, and v_x^{OPM}. The second group includes the single v_S^{GGA}, $v_x^{mod(GGA)}$, which is a sum of v_S^{GGA} and v_{resp}^{mod} calculated with either K_{sc} or K_g, and the total exchange potential v_x^{GGA} of the standard GGA approximation [92, 44]. The results are also compared with those obtained with the KLI potential 9.11.

To investigate a quality of the GGA approximation 9.44 to v_S, in Figure 2 $v_S^{\widetilde{GGA}}$ calculated self-consistently as a part of v_x^{GGA} and v_S^{GGA} obtained with the neglect of v_{resp} are compared with v_S^{OPM} calculated as a part of v_x^{OPM} and v_S obtained with the neglect of v_{resp} for Ne and Mg. It is interesting to note, that the neglect of v_{resp}^{OPM} does not influence on the Slater part of the exchange potential, so that v_S^{OPM} and v_S are hardly distinguished from each other on the scale presented. Contrary to this, the neglect of v_{resp}^{GGA} makes v_S^{GGA} visibly more attractive than the corresponding part $v_S^{\widetilde{GGA}}$ of v_x^{GGA}.

One can also see from Figure 2 the appreciable local deviations of v_S^{GGA} from v_S. For both Ne and Mg the former is more attractive than the latter within the regions of $1s-$ and $2s-$ shells and less attractive in the intershell region. At longer distances (not shown here) all the potentials presented have the same Coulombic asymptotics 9.9.

The local deviations of v_S^{GGA} from v_S clearly manifest themselves in Table 1 where the atomic K_{sc} values for $v_x^{mod(S)}$ and $v_x^{mod(GGA)}$ are presented. For all atoms (with the exception of the lightest Be and Ne) K_{sc} values obtained with the accurate functional 9.5 are rather close to the electron-gas constant $K_g = 0.382$. However, substitution of v_S for v_S^{GGA} leads to a great overestimation of the

Figure 9.2: Slater potentials calculated self-consistently as the parts of the OPM and GGA exchange potentials and also obtained with the neglect of v_{resp}. a) Ne and b) Mg

ATOM	$-E_{tot}^{OPM}$	v_{KLI}^{resp}	$v_{mod}^{resp}(K_{sc})$	$v_{mod}^{resp}(K_g)$	neglect of v_{resp}
Be	14.572	0	0	0	11
Ne	128.545	0	1	0	44
Mg	199.612	1	2	2	79
Ar	526.812	2	3	4	109
Ca	676.752	2	4	4	146
Zn	1777.834	4	6	6	258
Kr	2752.043	4	5	5	288
Sr	3131.533	4	7	7	324
Cd	5465.114	6	6	6	419
Xe	7232.121	7	12	11	450

Table 9.2: Comparison of overestimates (in mHartrees) of the OPM total energies E_{tot}^{OPM} [40, 36] (the latter are given in Ha calculated self-consistently with the exact functional v_S and various approximations to v_{resp}

integral I_1 of 9.37. As a result, K_{sc} value for Be calculated with v_S^{GGA} is 2.6 times as large as that calculated with v_S. The overestimation decreases rapidly with the increasing atomic number. Still, even for Xe K_{sc} for v_S^{GGA} is about 1.1 times as large as that for v_S.

The analysis of the contributions to I_1 shows, that the major part of the abovementioned overestimation comes from the region close to the nucleus where v_S^{GGA} exhibits the largest deviations from v_S (See Figure 2). The success of the GGA approximation is due to the fact, that the exchange energy integrals 9.7 with v_S^{GGA} are very close to those with v_S, the typical error is only about 0.1show that not a high local quality of GGA approximation, but an incredibly precise cancellation of local errors for the integral 9.7 provides this success. For the integrals of the type 9.37 associated with the Levy-Perdew relation this balance is destroyed, which leads to the overestimated K_{sc} values (See Table 1).

To analyse a local quality of the proposed model 9.23, $v_{resp}^{mod(S)}$ with the parameter K_{sc} and $v_{resp}^{mod(GGA)}$ with K_g are compared for Ne and Mg in Figure 3 with v_{resp}^{OPM}, v_{resp}^{KLI} with the parameters w_i from 9.12 and also with the corresponding potential v_{resp}^{GGA}, the latter has been obtained from the total exchange potential v_x^{GGA} of the GGA [92, 44] by the subtraction of its Slater part v_S^{GGA} (eq.9.44). In this case GGA gives considerably worse approximation than in the case of the Slater potentials discussed above. Due to the inclusion of the uncompensated Laplacian terms, v_{resp}^{GGA} has incorrect Coulombic divergence at nucleus, being proportional to $-\frac{1}{r}$ [32, 92]. At large distances it has incorrect Coulombic decay, being proportional to $\frac{1}{r}$. Furthermore, v_{resp}^{GGA} does not display the shell structure at intermediate distances, thus exhibiting large local deviations from v_{resp}^{OPM}.

Contrary to this, the simple model 9.23 provides a good fit to v_{resp}^{OPM} and to a more complicated approximation v_{resp}^{KLI}. Both potentials $v_{resp}^{mod(S)}$ and $v_{resp}^{mod(GGA)}$ have the proper short-range behaviour and they reproduce well a height and a width of the individual shell steps of v_{resp}^{OPM}, the largest differences being constants within $1s$-shell.

In Figure 4 various self-consistent approximate exchange potentials are compared with v_x^{OPM}. The Slater potential v_S is everywhere more attractive than v_x^{OPM} and the former can be considered as a satisfactory approximation to the latter only at larger distances where both potentials have the same Coulombic asymptotics. At $r < 1$ a.u. the neglect of the repulsive potential v_{resp} leads to the

Figure 9.3: Comparison of v_{resp}^{OPM} with various approximate response potentials. a) Ne and b) Mg

substantial overestimation of the exchange effect.

The GGA [44] offers a more balanced approximation to v_x, though the abovementioned defects of its components v_S^{GGA} and v_{resp}^{GGA} clearly manifest themselves in the total potential v_x^{GGA}. In particular, it has incorrect long-range asymptotics $-\frac{1}{r^2}$ and it is not attractive enough in the outer region. However, due to the divergence of its component v_{resp}^{GGA} at the nucleus and the overattractive character of v_S^{GGA} in the region close to the nucleus, v_x^{GGA} strongly overestimates the exchange effect in this region. At intermediate distances v_x^{GGA} smooths away the clear shell structure of v_x^{OPM}.

The self-consistent scheme of Section produces potentials with a higher local quality of approximation to v_x^{OPM}. One can see from Figure 4 that $v_x^{mod(S)}$ with the components v_S and v_{resp}^{mod} is an excellent approximation to v_x^{OPM}. $v_x^{mod(S)}$ is very close to v_x^{OPM} at any distance and in the major intervals they are even hardly distinguished from each other. The main difference is that the approximation 9.23 smooths away the small peaks of v_x^{OPM}. However, this difference does not seem the very important one, since these peaks appears to be a minor detail of the stepped shell structure of v_x^{OPM} (See Figures 1,4).

$v_x^{mod(GGA)}$ with the components v_S^{GGA} and v_{resp}^{mod} shows a worse local quality of approximation, mainly due to the defects of v_S^{GGA} discussed above. Nevertheless, the replacement of v_{resp}^{GGA} for v_{resp}^{mod} improves the short- and long-range asymptotics of the approximate potential and makes more clear its shell structure. As a result, $v_x^{mod(GGA)}$ is finite at the nucleus and has the proper Coulombic asymptotics at longer distances (See Figure 4).

9.5 Calculations with the accurate Slater potential

Tables 2,3 present the total E_{tot} and exchange E_x atomic energies calculated self-consistently with v_S and various potentials v_{resp}, namely, with v_{resp}^{OPM}, v_{resp}^{KLI}, v_{resp}^{mod}, the latter has been calculated in both variants with K_{sc} and K_g, and also with the neglect of v_{resp}. In all these approaches the energies are calculated with the same functional 9.5-9.7, so that the corresponding energy differences are caused exclusively by the differences in orbitals and densities generated with various exchange potentials. It is of no surprise, that the OPM E_{tot} values [40, 36, 94] are always the least ones presented in Table 2. The OPM potential is defined within the variational method and so it provides the true minimum of E_{tot} calculated within the one-determinantal approach with the exchange functional 9.5-9.7. The inclusion of the proper v_{resp} is of importance for a quality of calculated E_{tot}. The neglect of v_{resp} leads to the considerable errors, the error of E_{tot} obtained with the single v_S increases with the increasing atomic number and reaches 0.45 a.u. for Xe.

One can see from the Table, that the model 9.23 provides an excellent approximation to v_{resp}^{OPM} as regards the total energy. It is not too sensitive to variations of the parameter K and calculations with either K_{sc} or K_g yield the same energies as a more complicated KLI approximation. E_{tot} values obtained with v_S and v_{resp}^{mod} are only by a few milliHartrees higher than those of the OPM. The corresponding error increases (though non-monotonously) with the increasing atomic number, but even for Xe it is about 0.01 a.u. It means that the addition of 9.23 considerably improves a quality of the Kohn-Sham orbitals and density, which becomes very close to those of the rigorous OPM.

Because of the overattractive character of v_S, calculations with the neglect of v_{resp} yield too negative E_x (See Table 3). Addition of the approximate repulsive potentials v_{resp} overcompensates this error and makes the self-consistent E_x values more positive than E_x^{OPM} [37] (the only exception is

Figure 9.4: Comparison of v_x^{OPM} with various approximate exchange potentials. a) Ne and b) Mg

ATOM	$-E_x^{OPM}$	v_{KLI}^{resp}	$v_{mod}^{resp}(K_{sc})$	$v_{mod}^{resp}(K_g)$	neglect of v_{resp}
Be	2.666	-1	2	14	-49
Ne	12.107	9	5	30	-200
Mg	15.992	10	11	9	-305
Ar	30.182	8	41	63	-436
Ca	35.209	5	40	30	-525
Zn	69.647	78	55	60	-1030
Kr	93.875	65	88	90	-1101
Sr	101.974	58	59	14	-1177
Cd	148.963	123	150	123	-1536
Xe	179.173	119	218	199	-1591

Table 9.3: Comparison of differences (in mHartrees) of the OPM exchange energies E_x^{OPM} [37] (the latter are given in Hartrees) an self-consistently with the exact functional v_S and various approximations to v_{resp}

ATOM	$-\epsilon_N^{OPM}$	v_{KLI}^{resp}	$v_{mod}^{resp}(K_{sc})$	$v_{mod}^{resp}(K_g)$	neglect of v_{resp}
Be	309	0	1	6	-17
Ne	851	2	21	30	-61
Mg	253	1	5	5	-31
Ar	591	2	18	21	-47
Ca	196	1	10	10	-29
Zn	293	1	-14	-14	-64
Kr	523	1	20	20	-44
Sr	179	1	12	10	-30
Cd	265	0	2	1	-65
Xe	456	1	23	22	-40

Table 9.4: Comparison of differences of the OPM energies of the highest occupied orbital ϵ_N^{OPM} (in mHartrees) [40, 36] and those, calculated self-consistently with the exact functional v_S and various approximations to v_{resp}

E_x for Be obtained with v_{resp}^{KLI}) and much more close to the latter. A comparison of Tables 2 and 3 reveals another type of compensation. In all cases the errors of E_x are compensated with those of the opposite sign in other parts of the total energy, so that E_{tot} errors are much smaller than those of E_x. In most cases the E_x errors of the model 9.23 are somewhat larger (though of the same magnitude) than those of KLI.

The most important one-electron energetical characteristic of the Kohn-Sham theory is the energy ϵ_N of the highest occupied orbital ϕ_N. For the exact Kohn-Sham potential ϵ_N is equal to minus the ionisation energy I_p of the system [134, 31], and in the exchange-only case ϵ_N has the same meaning through the Koopmans's theorem [135, 35]. Table 4 represents ϵ_N values obtained with v_S and various v_{resp}. It follows from the Table, that in the case of the orbital energies ϵ_N the qualitative trends are the same as in the case of the exchange energies discussed above. In spite of its correct asymptotics 9.9, the single v_S always has too negative ϵ_N values.

Addition of v_{resp} compensates this error. By the construction, both v_{resp}^{KLI} and v_{resp}^{mod} decay exponentially in the region of ϕ_N location, because ϕ_N does not contribute into the numerators of 9.23 and the second term of 9.11, while contributing into the density ρ in the denominators. The resulting exponential tails of v_{resp}^{mod} and v_{resp}^{KLI} produce the compensating repulsive contribution to ϵ_N. Addition of v_{resp}^{mod} even has an overcompensating effect, because the approximation 9.23 always (with the only exception of Zn) overestimates a value of the parameter w_{N-1} for ϕ_{N-1}, the next to the highest occupied orbital. As a result, v_{resp}^{mod} becomes more positive than v_{resp}^{KLI} in the outer valence region (See Figure 3) and yields more positive ϵ_N values than those of KLI (the only exception is ϵ_N of Zn), the latter are virtually the same as the OPM ϵ_N. However, the corresponding errors are not large and in both variants with K_{sc} and K_g vary within 0.01 - 0.03 a.u.

The present results show, that the self-consistent scheme with v_S and v_{resp}^{mod} can be used as a very good approach to the OPM. The variants with K_{sc} and K_g yield results of the same quality, so one can use a more simple variant with the universal electron-gas parameter K_g for all systems. Because of the exact functional v_S, the scheme requires calculation of the two-electron integrals with the orbitals ϕ_i and the required computational time per iteration is approximately the same as in the case of KLI. However, the replacement of v_{resp}^{KLI} for v_{resp}^{mod} greatly accelerates the convergence of the self-consistent procedure. The ratio of iterations before the convergence in KLI and in the present scheme varies within 6-12 for the noble-gas atoms from Ar to Xe and within 12-18 for the alkaline-earth atoms from Mg to sr. As a result, the proposed scheme takes about an order of magnitude of the computational time as small as KLI.

Still, to develop a practical DFT scheme, one should approximate not only v_{resp}, but also v_S. The results of calculations with the GGA approximation to v_S will be presented in the next section.

9.6   Calculations with the GGA approximation to v_S

Tables 5,6 represent E_{tot} and E_x values obtained with the GGA approximation 9.44 to v_S with and without the response potentials v_{resp}^{GGA} and v_{resp}^{mod}. All these energies are calculated with the same standard GGA energy functional, for which the variationally derived v_{resp}^{GGA} yields the minimal E_{tot}. However, because of the approximate nature of this functional, E_{tot}^{GGA} are not bound to E_{tot}^{OPM} from below and in the most cases the former are too negative (See Table 5). On the other hand, similarly to Table 2, calculations with the neglect of v_{resp} yield too positive E_{tot} values. In the latter case the corresponding error increases monotonously with the increasing atomic number, the only exception is Zn.

ATOM	$-E_{tot}^{OPM}$	v_{GGA}^{resp}	$v_{mod}^{resp}(K_{sc})$	$v_{mod}^{resp}(K_g)$	neglect of v_{resp}
Be	14.572	6	34	12	21
Ne	128.545	-45	-30	-29	35
Mg	199.612	-21	3	-5	78
Ar	526.812	11	33	26	141
Ca	676.752	-2	27	15	149
Zn	1777.834	-287	-272	-268	33
Kr	2752.043	-60	-39	-41	269
Sr	3131.533	-48	-19	-23	298
Cd	5465.114	-79	57	60	364
Xe	7232.121	2	35	25	463

Table 9.5: Comparison of differences (in mHartrees) between the OPM total energies E_{tot}^{OPM} [40, 36] and those, calculated self- v_S^{GGA} and various approximations to v_{resp}

ATOM	$-E_x^{OPM}$	v_{GGA}^{resp}	$v_{mod}^{resp}(K_{sc})$	$v_{mod}^{resp}(K_g)$	neglect of v_{resp}
Be	2.666	14	44	-10	-73
Ne	12.107	20	-36	-122	-359
Mg	15.992	20	25	-92	-406
Ar	30.182	59	93	-13	-520
Ca	35.209	27	97	-39	-606
Zn	69.646	-76	-138	-303	-1375
Kr	93.876	76	95	-83	-1277
Sr	101.974	61	91	-128	-1332
Cd	148.963	126	189	-27	-1688
Xe	179.174	167	314	100	-1714

Table 9.6: Comparison of differences (in mHartrees) between the OPM total energies E_x^{OPM} [37] and those, calculated self-consistent v_S^{GGA} and various approximations to v_{resp}

ATOM	$-\epsilon_N^{OPM}$	v_{GGA}^{resp}	$v_{mod}^{resp}(K_{sc})$	$v_{mod}^{resp}(K_g)$	neglect of v_{resp}
Be	309	128	29	8	-15
Ne	851	396	128	98	9
Mg	253	104	19	6	-30
Ar	591	249	57	43	-23
Ca	196	80	15	6	-34
Zn	293	102	-19	-25	-74
Kr	523	218	42	33	-31
Sr	179	72	12	5	-38
Cd	265	95	-6	-13	-82
Xe	456	187	35	28	-34

Table 9.7: Comparison of differences of the OPM energies of the highest occupied orbital ϵ_N^{OPM} (in mHartrees) [40, 36] and those, calculated self-consistently with v_S^{GGA} and various approximations to v_{resp}

Addition of v_{resp}^{mod} considerably reduces the error, the only exceptions are Zn in both variants with K_{sc} and K_g and Be in the variant with K_{sc}. E_{tot} for Be obtained with K_g is much closer to E_{tot}^{OPM} than that obtained with K_{sc}, because of the overestimation of K_{sc} in the GGA discussed above. In general, however, both variants yield the very similar E_{tot} values, which are of the same accuracy as those of the standard GGA. The absolute error, induced mainly by the GGA approximation to v_S, vary largely for different atoms, the typical value is of a few centiHartrees and the maximal (and the exceptionally large) one is 0.3 a.u. for Zn.

In the complete analogy with the case of v_S, calculations with the single v_S^{GGA} yield too negative E_x values (See Table 6). Addition of the approximate v_{resp} considerably compensates the corresponding error. In the case of v_{resp}^{GGA} and v_{resp}^{mod} with K_{sc} this leads even to overcompensation and for the most atoms E_x are too positive, while in the case of v_{resp}^{mod} with K_g the compensation is not enough, thus producing too negative E_x (the only exception in the latter case is Xe). In general, all the schemes with the approximate v_{resp} yield E_x values of a comparable accuracy.

The self-consistent scheme with v_S^{GGA} and v_{resp}^{mod} shows a definite advantage over the standard GGA in calculation of ϵ_N (See Table 7). As it was indicated in [36, 32] and one can see this from the Table, the GGA greatly underestimates the absolute magnitude of ϵ_N. Due to the incorrect asymptotics of its response part (See Figure 3), $|\epsilon_N|$ values of v_x^{GGA} are about twice as small as compared with the OPM values and are very close to the LDA ones. On the other hand, the single v_S^{GGA} overestimates $|\epsilon_N|$ (the only exception is Ne), though the corresponding errors are considerably smaller.

Addition of v_{resp}^{mod} with K_{sc} to v_S overcompensates this effect and produces comparable errors of the opposite sign. The ϵ_N values of the potential $(v_S^{GGA} + v_{resp}^{mod})$ obtained with K_g are in most cases the best approximate ones and the closest to ϵ_N^{OPM}.

It follows from the above analysis that the self-consistent scheme with v_S^{GGA} and v_{resp}^{mod} provides the same accuracy for the total and exchange energies as the standard GGA scheme [44] and considerably improves the form and the eigenvalue ϵ_N of the one-electron potential. Bearing in mind high quality of the presented results, we propose v_x^{mod} with the components v_S^{GGA} and v_{resp}^{mod} as an efficient approximation to v_x. The variant with K_g provides a better fit to the accurate v_{resp} and it is also simpler than that with K_{sc}, so the electron-gas value K_g can be recommended as the universal parameter for the many-electron calculations.

9.7 Conclusions

In this paper a scheme of approximation of the Kohn-Sham exchange potential v_x has been proposed, making use of a partitioning of v_x into the Slater and response potentials. The model v_{resp}^{mod} has been derived from the dimensional arguments, which possesses the proper short-range behaviour and the characteristic atomic-shell stepped structure. When combined with the accurate v_S, v_{resp}^{mod} provides an excellent approximation to v_x^{OPM}. With the GGA approximation to v_S, v_{resp}^{mod} provides an efficient DFT approach which, for the first time, fits closely the form of the accurate exchange potential and yields reasonably accurate exchange and total energies as well as the energies of the highest occupied orbital.

Still, we have to mention appreciable errors, which are introduced with the GGA approximation to v_S. For example, the typical error of the calculated total energies is increased from milliHartrees to santiHartrees with the replacement of v_S for v_S^{GGA}. So, in order to provide a better practical DFT scheme, one should improve, first of all, a quality of v_S approximation. In particular, the present GGA approximation shows considerable deviations from the accurate v_S in the region close to the nucleus. A promising way of v_S^{mod} refinement is to construct it not as a function of ρ and $\nabla\rho$, but as some function of the orbital densities $|\phi_i|^2$ and their gradients.

To test a quality of the developed v_x^{mod} approximation, the exchange-only atomic calculations have been performed in this work. Our main goal, however,is to apply this approximation for the exchange-correlation molecular calculations. For this purpose one can use v_x^{mod} either as an independent exchange part of the approximate exchange-correlation Kohn-Sham potential v_{xc}^{appr} or as the basic functional form for approximation of the total v_{xc}. Within the former approach v_x^{mod} is inserted in the Kohn-Sham equations 9.2 together with some approximation for the Coulomb correlation potential v_c, the latter is to be constructed independently. Then, the exchange energy is calculated via 9.7 with the self-consistent density and v_S^{mod}, while the correlation energy is calculated with the independent functional, which corresponds to v_c.

An alternative approach is based on the fact, that v_x is the dominant part of v_{xc} and the exchange energy density ϵ_x is the dominant part of ϵ_{xc}. To approximate the total v_{xc} and ϵ_{xc}, one can use the same models v_S^{mod} and v_{resp}^{mod} with other parameters chosen to fit the available accurate exchange-correlation Kohn-Sham potentials [32, 74, 78, 79, 80]. The work along both abovementioned lines is in progress.

Chapter 10

An analysis of nonlocal density functionals in chemical bonding

In this work we carry out an analysis of the gradient corrected density functionals in molecules that are used in the Kohn-Sham density functional approach. We concentrate on the special features of the exchange and correlation energy densities and exchange and correlation potentials in the bond region. By comparing to the exact Kohn-Sham potential it is shown that the gradient corrected potentials build in the required peak in the bond midplane, but not completely correctly. The gradient corrected potentials also exhibit wrong asymptotic behaviour. Contributions from different regions of space (notably bond and outer regions) to nonlocal bonding energy contributions are investigated by integrating the exchange and correlation energy densities in various spatial regions. This provides an explanation why the gradient corrections reduce the LDA overbinding of molecules. It explains the success of the presently used nonlocal corrections, although it is possible that there is a cancellation of errors, too much repulsion being derived from the bond region and too little from the outer region.

10.1 Introduction

In the last few years considerable progress has been achieved within density functional theory [5] in the calculation of molecular [70, 71, 122] and solid state [47, 136, 137] properties by the use of nonlocal density functionals. The use of these nonlocal corrections gives an overall improvement of the results obtained by the local density approximation (LDA). This is especially true for energy quantities such as bonding energies in molecules. The successes are not restricted to simple covalent molecules (cf. the G1 and G2 sets [70, 122]) but also occur for more complicated cases where the conventional ab-initio Hartree-Fock method fails notoriously such as in multiply bonded molecules and particularly in the field of transition metal chemistry [71]. Although numerical calculations clearly show that the nonlocal corrections do work, not much understanding has been gained in why and how they work. There are several questions left unanswered. Why do the nonlocal corrections always reduce the bonding energies of molecules? What do molecular Kohn-Sham potentials look like and do the nonlocal corrections give a good representation of these potentials? How can the correlation effects (left-right correlation) in dissociating molecules be incorporated in gradient corrected density functionals and their potentials? Why does the Becke correction for exchange, that has been devised to yield good (Hartree-Fock) exchange energies in atoms, change LDA bond energies not towards Hartree-Fock but (close) to exact bond energies? An understanding of such

167

questions might lead to the construction of still more improved density functionals which is not only of theoretical but also of practical importance.

In the literature little attention has been paid to the question why the nonlocal corrections improve the energies. It has been argued [70, 138] that the improved bond energies due to the Becke functional [44] are due to the improvement of the asymptotic behaviour of the exchange energy density. However in a clear paper by Engel et al. [49] it was shown for the case of atoms that the correct behaviour of the Becke exchange energy density is only reached at very large distances where it has almost no effect on the energy. Moreover these authors constructed a functional which did not satisfy the correct asymptotic property but which yielded even better atomic exchange energies. Whether the correct asymptotic behaviour of the Becke functional is responsible for the improved bond energies is therefore an open problem which we will investigate more closely in this paper.

The paper is divided in the following way. In section 10.2 we discuss the properties of the exact molecular Kohn-Sham potential based on calculations of the hydrogen molecule. We then discuss how the left-right correlation shows up in the potential. In section 10.3 we discuss the properties of the potentials of two widely used nonlocal functionals namely those of Becke [44, 139] and of Perdew [45]. It is shown that these potentials are characterised by a peak in the bond midpoint region. In section 10.4 we analyse the exchange and correlation energy densities of the Becke and the Perdew functionals and give an explanation for the fact that the nonlocal corrections reduce the bonding energy in molecules. In section 10.5 we then give a numerical analysis of the bonding contributions of the nonlocal functionals. This is done by integrating the exchange and correlation energy densities of Becke and Perdew over specific spatial regions of some diatomic molecule. We finally present a summary and conclusions in section 10.6.

10.2 Left-right correlation and the exact Kohn-Sham potential

In order to improve on existing density functionals it is very helpful to know how molecular Kohn-Sham potentials look like. Especially the knowledge on the correlation potential is very limited. There is however one molecular system for which the correlation potential can be accurately calculated and in which correlation is very important. This sytem is the dissociating hydrogen molecule. For a general molecular system with N electrons the Kohn-Sham equations are given by:

$$(-\frac{1}{2}\nabla^2 + v^{KS}([\rho]; \mathbf{r}))\psi_i(\mathbf{r}) = \epsilon\psi_i(\mathbf{r}) \tag{10.1}$$

where the electron density ρ is given by the sum of orbital densities:

$$\rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \tag{10.2}$$

The Kohn-Sham potential v^{KS} is usually split up in the following way:

$$v^{KS}([\rho]; \mathbf{r}) = v_{nuc}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}([\rho]; \mathbf{r}) \tag{10.3}$$

The first term in this equation is the Coulomb attraction of the atomic nuclei and the second term is the repulsion of the electronic cloud. The last term, the one we are interested in, is the exchange-correlation potential which is the functional derivative of the exchange-correlation functional. For a

spin restricted two-electron system such as the hydrogen molecule in its ground state the occupied orbitals can explicitly be expressed in the electron density as $\psi_i = \sqrt{\rho/2}$. Substituting this in the Kohn-Sham equations then yields an explicit expression of the Kohn-Sham potential expressed in the exact ground state density:

$$v^{KS}(\mathbf{r}) = -\frac{1}{8}\frac{(\nabla\rho)^2}{\rho^2} + \frac{1}{4}\frac{\nabla^2\rho}{\rho} + \epsilon \tag{10.4}$$

For a two-electron system ϵ is the highest occupied Kohn-Sham orbital and therefore equal to the ionisation energy of the system. The above equation gives us the possibility to calculate the Kohn-Sham potential from a correlated electron density which can for instance be obtained from an accurate configuration interaction (CI) calculation. This has been done by Smith et al. [140] for the helium atom. A different approach to obtain the Kohn-Sham potential has been used by Buijse et al. [86] who carried out a thorough investigation of this potential for the hydrogen molecule at various distances. For comparison in the next section we stress some points of their results. The basic quantity in the analysis of Buijse et al. is the conditional probability amplitude $\Phi(\mathbf{r}_2,\ldots,\mathbf{r}_N|\mathbf{r}_1)$. defined by:

$$\Phi(\mathbf{r}_2,\ldots,\mathbf{r}_N|\mathbf{r}_1) = \frac{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\sqrt{\rho(\mathbf{r}_1)/N}} \tag{10.5}$$

where Ψ is the ground state wavefunction of the system. The amplitude gives a description of the $(N-1)$-electron system when one of the electrons is known to be at position $\mathbf{r}_1$. The amplitude squared gives the probability that the other electrons are at positions $\mathbf{r}_2,\ldots,\mathbf{r}_N$ if one electron (the reference electron) is known to be at position $\mathbf{r}_1$. All the correlation effects of the system are contained in this quantity. These correlation effects especially show up in the case of the hydrogen molecule when we move the position of the reference electron along the bond axis from one atom to the other atom. Due to the Coulomb correlation of the electrons we have that if one electron is known to be at a given atom then the probability amplitude is large that the other electron will be at the other atom. This changes rapidly when the reference electron crosses the bond midpoint region because in that case the other electron has to switch quickly from one atom to the other atom. A quantity which measures this change in conditional amplitude (the so-called left-right correlation effect) as we move our reference electron is the so-called kinetic potential defined by:

$$v_{kin}(\mathbf{r}_1) = \frac{1}{2}\int|\nabla_1\Phi|^2 d\mathbf{r}_2\ldots d\mathbf{r}_N \tag{10.6}$$

This potential is clearly positive definite. The origin of the name of this potential stems from the fact that it contributes to the kinetic energy of the system which is given by:

$$T = \frac{1}{8}\int\frac{(\nabla\rho)^2}{\rho}d\mathbf{r} + \int\rho(\mathbf{r})v_{kin}(\mathbf{r})d\mathbf{r} \tag{10.7}$$

This follows directly from the definition of the kinetic energy and formula 10.5. Note that the first term in the above formula is the von Weiszäcker kinetic energy $T_W[\rho]$ which for a two-electron system is equal to the Kohn-Sham kinetic energy $T_s[\rho]$. The term with v_{kin} therefore contributes to the Kohn-Sham exchange-correlation energy of the system. As has been shown in reference [86] v_{kin} is also an important constituent term of the Kohn-Sham potential. By numerical calculation the following properties of the kinetic potential were observed [86].

First of all this potential is peaked on the bond midpoint. As a result of this the total exchange-correlation potential is also peaked in this region. This peak in v_{kin} is not surprising from its definition. The changes in the conditional amplitude as we move our reference electron is the largest

in the bond midpoint region as explained before. Secondly the peak in the kinetic potential gets higher when we dissociate the molecule. This is also what one expects since the left-right correlation effects grow stronger as we dissociate. This can also be seen from the well-known Hartree-Fock error for a dissociating system. The Hartree-Fock one-determinant wavefunction upon dissociation no longer gives a right description of the division of the electrons over the atomic fragments by putting too much weight on ionic terms [17].

Knowing the behaviour of the Kohn-Sham potential in the bonding region of molecules one might wonder whether these peculiar properties are well represented by the presently used nonlocal gradient corrections. We will show in the next section that both the Becke and Perdew potentials exhibit peaks in this region in cases of strong left-right correlation. The Becke potential exhibits a positive peak in this region although the Becke functional is purely derived as an exchange correction and not a correlation correction. The Perdew potential however, which is derived as a correlation correction, exhibits a negative peak although not large enough to cancel against the peak in the Becke potential. The peaks in both potentials are also shown to grow upon dissociation of the molecule.

10.3   Analysis of Beckes and Perdews molecular potentials

In this section we will analyse the properties of both Beckes exchange potential and Perdews correlation potential. We will start with Beckes exchange potential. The Becke gradient correction to the exchange energy which belongs to the family of Generalised Gradient Approximations (GGA) [43, 49] is defined as:

$$E_x^{GGA}[\rho] = \sum_\sigma \int \rho_\sigma^{\frac{4}{3}} f(x_\sigma) d\mathbf{r} \tag{10.8}$$

in which σ is a spin index and $x_\sigma = |\nabla \rho_\sigma|/\rho_\sigma^{\frac{4}{3}}$. The function f is in the case of the Becke functional defined as:

$$f(x) = -\beta \frac{x^2}{1 + 6\beta x \sinh^{-1} x} \tag{10.9}$$

The constant β is given by $\beta = 0.0042$. This function is chosen so as to satisfy some scaling, asymptotic and weak inhomogeneity properties [32]. The potential corresponding to E_x^{GGA} is given by (for convenience we leave the spin index out, it simply can be added to the final result):

$$v_x^{GGA}(\mathbf{r}) = \frac{\delta E_x}{\delta \rho(\mathbf{r})} = \frac{4}{3}\rho^{\frac{1}{3}}(f(x) - x\frac{df}{dx} + x^2\frac{d^2 f}{dx^2}) +$$

$$\sum_{i,j} \frac{\partial_i \rho \partial_i \partial_j \rho \partial_j \rho}{|\nabla \rho|^3}(\frac{df}{dx} - x\frac{d^2 f}{dx^2}) - \frac{\nabla^2 \rho}{|\nabla \rho|}\frac{df}{dx} \tag{10.10}$$

We will now analyse this potential in the vicinity of the bond midpoint region. In that case the gradients are small and our parameter x is therefore also small. We therefore have:

$$f(x) \sim -\beta x^2 + O(x^4) \qquad (x \downarrow 0) \tag{10.11}$$

hence:

$$\frac{df}{dx} \sim -2\beta x + O(x^3) \qquad (x \downarrow 0) \tag{10.12}$$

and

$$\frac{d^2 f}{dx^2} \sim -2\beta + O(x^2) \qquad (x \downarrow 0) \tag{10.13}$$

Inserting this in the potential formula 10.10 yields:

$$v_x^{GGA}(\mathbf{r}) = 2\beta \frac{\nabla^2 \rho}{\rho^{\frac{4}{3}}} + O(|\nabla \rho|^2) \tag{10.14}$$

This yields in the bond midpoint $\mathbf{r}_M$, where $\nabla \rho = 0$, for the Becke potential:

$$v_x^B(\mathbf{r}_M) = 0.0106 \frac{\nabla^2 \rho}{\rho^{\frac{4}{3}}} \tag{10.15}$$

for the spin unpolarised case ($\rho_\sigma = \frac{\rho}{2}$). In this derivation we only used the property $f(x) \sim x^2$ ($x \downarrow 0$). This requirement originates from the gradient expansion for slowly varying densities of the exchange energy, for which the gradient correction in the so-called Gradient Expansion Approximation (GEA) is given by:

$$E_x^{GEA}[\rho] = -\beta \int \frac{(\nabla \rho)^2}{\rho^{\frac{4}{3}}} d\mathbf{r} = -\beta \int \rho^{\frac{4}{3}} x^2 d\mathbf{r} \tag{10.16}$$

This is the gradient correction used by Herman et al. in the $X\alpha\beta$-method [76]. The corresponding potential is given by:

$$v_x^{GEA}(\mathbf{r}) = 2\beta \frac{\nabla^2 \rho}{\rho^{\frac{4}{3}}} - \frac{4}{3}\beta \frac{(\nabla \rho)^2}{\rho^{\frac{7}{3}}} \tag{10.17}$$

which is in accordance with equation 10.14. We will analyse the $\nabla^2 \rho / \rho^{\frac{4}{3}}$ term a bit further. In order to do this we consider the case of a homonuclear diatomic molecule. In the dissociation limit the molecular density can be well approximated by the sum of atomic densities $\rho = \rho_A + \rho_B$ of atom A and B. If we use the asymptotic relation $\rho_{A,B} \sim Ne^{-\alpha r_{A,B}}$ ($r \to \infty$) where $r_{A,B}$ is the distance to atom A, B we obtain

$$\nabla^2 \rho = \alpha^2 \rho - \frac{2\alpha}{r_A} \rho_A - \frac{2\alpha}{r_B} \rho_B \approx \alpha^2 \rho \quad (r_{A,B} \to \infty) \tag{10.18}$$

and for the height of the peak in the potential in the bond midpoint at large internuclear distance R_{AB} of the atoms:

$$v_x^{GGA}(\mathbf{r}_M) \sim \frac{K}{\rho^{\frac{1}{3}}} \to \infty \quad (R_{AB} \to \infty) \tag{10.19}$$

where K is a positive constant. We thus conclude that the peak increases at large distances and even has the unphysical behaviour of becoming infinite in the limit of infinite separation of the atoms. However at these distances the density decreases faster ($\rho \sim e^{-\alpha r}$) than the peak increases so the effect of this erroneous behaviour of the GGA exchange potential in the bond midpoint region on the molecular density may still be very small at large bond distances.

At shorter bond distances we cannot make any definite conclusions on the behaviour of the Becke (and Perdew) potential from the above analysis. For instance from formula 10.14 we can deduce that the sign of the peak depends on the sign of the Laplacian. In the weak bonding limit this sign is positive, however at smaller distances this sign depends strongly on the type of bond considered

which has extensively been studied in reference [102].

We will next discuss the bond midpoint properties of Perdews correlation potential. The Perdew correlation energy functional is defined by [45] :

$$E_c^{GGA}[\rho] = \int \frac{1}{d(\zeta)} C(\rho) \rho^{\frac{4}{3}} x^2 e^{-\Phi(\rho, |\nabla\rho|)} d\mathbf{r} \tag{10.20}$$

where $\zeta = \frac{\rho_\sigma - \rho_{\sigma'}}{\rho}$ is the spin polarisation and:

$$d(\zeta) = 2^{\frac{1}{3}} \left(\left(\frac{1+\zeta}{2}\right)^{\frac{5}{3}} + \left(\frac{1-\zeta}{2}\right)^{\frac{5}{3}} \right)^{\frac{1}{2}} \tag{10.21}$$

The function $C(\rho)$ has been obtained from the gradient expansion for slowly varying densities of the correlation energy of the electron gas [42] and is usually expressed in the Wigner-Seitz radius $r_s = \left(\frac{3}{4\pi\rho}\right)^{\frac{1}{3}}$ as:

$$C(\rho) = 0.001667 + \frac{0.002568 + \alpha r_s + \beta r_s^2}{1 + \gamma r_s + \delta r_s^2 + 10^4 \beta r_s^3} \tag{10.22}$$

The constants $\alpha, \beta, \gamma, \delta$ are given by $0.023266, 7.389 \times 10^{-6}, 8.723, 0.472$. We finally have:

$$\Phi(\rho, |\nabla\rho|) = 1.745 \tilde{f} \frac{C(\infty)}{C(\rho)} \frac{|\nabla\rho|}{\rho^{\frac{7}{6}}} \tag{10.23}$$

where $\tilde{f} = 0.11$. The form of this function is obtained from the wavevector analysis of Langreth and Mehl [50]. For slowly varying densities the Perdew functional reduces to the gradient expansion result for the correlation energy (we consider the spin unpolarised case for which $d(\zeta) = 1$):

$$E_c^{GEA}[\rho] = \int C(\rho) \rho^{\frac{4}{3}} x^2 d\mathbf{r} \tag{10.24}$$

Both this functional and the Perdew functional have the same values for their functional derivatives in the bond midpoint region where $\nabla\rho = 0$. We have:

$$v_c^{GGA}(\mathbf{r}_M) = -2C(\rho) \frac{\nabla^2 \rho}{\rho^{\frac{4}{3}}} \tag{10.25}$$

In the dissociation limit where the bond midpoint densities are small we have $C(\rho) \approx C(0) = 0.001667$ and thus:

$$v_c^{GGA}(\mathbf{r}_M) = -0.0033 \frac{\nabla^2 \rho}{\rho^{\frac{4}{3}}} \tag{10.26}$$

This has the same form as Beckes potential in the bond midpoint region, however with a different sign. We therefore find a negative peak in the Perdew potential which increases in absolute value when the molecule dissociates. However, the total nonlocal potential which is the sum of Beckes and Perdews potentials still has a positive peak in the bond region since the negative peak in Perdews potential is not large enough to cancel the positive peak in Beckes potential. We have for the total nonlocal exchange-correlation potential:

$$v_{xc}^{NL}(\mathbf{r}_M) = v_x^{Becke}(\mathbf{r}_M) + v_c^{Perdew}(\mathbf{r}_M) = 0.0073 \frac{\nabla^2 \rho}{\rho^{\frac{4}{3}}} \tag{10.27}$$

So in the bond region in the dissociation limit we have a repulsive Becke potential and an attractive Perdew potential.

We can now ask the question how this behaviour of the Becke-Perdew potential compares with the exact potential. Let us compare equation 10.27 in the case of a two-electron system with the exact Kohn-Sham potential of equation 10.4 in the bond midpoint region. In the bond midpoint we have:

$$v^{KS}(\mathbf{r}_M) = \frac{1}{4}\frac{\nabla^2 \rho}{\rho} + \epsilon \tag{10.28}$$

In contrast to the potential in equation 10.27 the bond midpoint value of the exact Kohn-Sham potential does not go to infinity if $(R_{AB} \to \infty)$ but to the positive constant $-\epsilon$ (see also [86]) which can easily be derived from formula 10.18 using $\alpha = 2\sqrt{-2\epsilon}$. This shows that the Becke-Perdew potential cannot give a good description of the bond midpoint behaviour of the exact potential at very large bond distances. However for intermediate bond distances it can give an improvement of the LDA potential but this improvement is only of qualitative nature as the height of the peak in the Kohn-Sham potential in the LDA+Becke-Perdew approximation has a different functional dependence on the density than the exact Kohn-Sham potential. In section 10.5 we will present some numerical examples of the molecular Becke and Perdew potentials. However we will first address the question of bond energies and analyse the exchange and correlation energy densities in the bonding region.

10.4   Gradient corrected energy densities

The positive peak in the potential already indicates that adding the Becke-Perdew nonlocal corrections to the local density calculations may in weak bonding cases reduce the bonding energy of molecules, i.e. give a positive contribution to ΔE defined as the molecular energy minus the atomic energies. We will try to identify the origin of the positive contribution due to the nonlocal corrections by considering the difference between the molecular energy density and the sum of atomic energy densities.

The exchange and correlation energy densities can be defined as:

$$\epsilon_x(\mathbf{r}) = \frac{1}{2}\rho(\mathbf{r})V_x(\mathbf{r}) \tag{10.29}$$

and

$$\epsilon_c(\mathbf{r}) = \frac{1}{2}\rho(\mathbf{r})V_c(\mathbf{r}) \tag{10.30}$$

where V_x and V_c are the potentials due to the coupling constant integrated Fermi and Coulomb holes [24] which should not be confused with the potentials v_x and v_c which are the functional derivatives of the exchange and correlation functionals. There are however other definitions possible of V_x and V_c. One could for instance define V_x to be the potential due to the Fermi hole of the optimised potential model (OPM) [33, 36] and define V_c as the difference between this potential and the potential due to the coupling constant integrated exchange-correlation hole. In either definition the potential V_x goes asymptotically like $-\frac{1}{r}$ due to the fact that the Fermi hole integrates to one electron. The exchange potential potential v_x has a similar asymptotic decay. The contribution of the nonlocal corrections to ϵ_x and ϵ_c we will in the following denote by e_x and e_c.

The contribution to the bonding energy ΔE of a molecule of the nonlocal corrections to the exchange and correlation functionals can be written as:

$$\Delta E_x^{NL} = \int \delta e_x(\mathbf{r})d\mathbf{r} \tag{10.31}$$

and

$$\Delta E_c^{NL} = \int \delta e_c(\mathbf{r})d\mathbf{r} \tag{10.32}$$

where

$$\delta e_x(\mathbf{r}) = e_x^M(\mathbf{r}) - \sum_A e_x^A(\mathbf{r}) \tag{10.33}$$

and

$$\delta e_c(\mathbf{r}) = e_c^M(\mathbf{r}) - \sum_A e_c^A(\mathbf{r}) \tag{10.34}$$

in which e_x^M and e_c^M denote the gradient corrections to the molecular exchange and correlation energy densities. The functions e_x^A and e_c^A denote the gradient corrections to the atomic exchange and correlation energy densities where the atoms are situated at their molecular positions. In equations 10.33 and 10.34 the sum is taken over all atoms in the molecule. We will analyse the energy density difference functions δe_x and δe_c for the Becke and Perdew functional. In the bond midpoint we have that the gradient of the molecular density is zero, $\nabla \rho^M = 0$, and in that case the Becke and Perdew energy density functions are given by:

$$\delta e_x(\mathbf{r}_M) = \sum_A \beta \rho^{A\frac{4}{3}} \frac{x^{A2}}{1 + 6\beta x^A \sinh^{-1} x^A} > 0 \tag{10.35}$$

and

$$\delta e_c(\mathbf{r}_M) = -\sum_A \frac{C(\rho^A)}{d(\zeta^A)} \rho^{A\frac{4}{3}} x^{A2} e^{-\Phi(\rho^A, |\rho^A|)} < 0 \tag{10.36}$$

We will show in the next section for the example of N_2 that δe_x does have a positive peak in the bond midpoint region and δe_c has a negative peak in the same region. Suppose that the main contributions to the functions δe_x and δe_c are found in the bond midpoint region of the molecule. In that case we would have:

$$\Delta E_x^{NL} = \int \delta e_x(\mathbf{r})d\mathbf{r} > 0 \tag{10.37}$$

and

$$\Delta E_c^{NL} = \int \delta e_c(\mathbf{r})d\mathbf{r} < 0 \tag{10.38}$$

So we can infer that in the bond midpoint region the Becke functional gives a positive correction to the bond energy and the Perdew functional gives a negative contribution. The main reason for the improvement of the bonding energy of molecules by the nonlocal corrections is then the replacement of a region of nonzero gradients in the sum of atomic exchange-correlation densities by a region of zero gradients in the exchange-correlation energy density of the molecule.

There is also another region from where we may expect a positive contribution from gradient corrections to the LDA bond energy, namely the outer region where the density decays exponentially to zero. It is well known that the LDA exchange-correlation energy density also decays exponentially to zero, whereas it should behave asymptotically like $-\rho/2r$. From any gradient correction that builds in the correct asymptotic behaviour, as does the Becke correction, one may expect a positive contribution to the bond energy from the asymptotic region. In that region, where

$$e_x(\mathbf{r}) \sim -\frac{\rho}{2r} \qquad (r \to \infty) \tag{10.39}$$

we have for δe_x that:

$$\delta e_x(\mathbf{r}) \sim -\frac{\rho^M}{2r} + \sum_A \frac{\rho^A}{2r} = -\frac{\delta\rho}{2r} \qquad (r \to \infty) \tag{10.40}$$

in which $\delta\rho = \rho^M - \sum_A \rho^A$ is the deformation density of the molecule. If the deformation density $\delta\rho$ is negative in the outer regions of the molecule then:

$$\delta e_x(\mathbf{r}) = -\frac{\delta\rho}{2r} > 0 \qquad (r \to \infty) \tag{10.41}$$

This then contributes positively to ΔE_x. As may be inferred from elementary considerations regarding the relation between the exponential decay of molecular and atomic densities and the first ionisation energies, $\delta\rho$ is usually negative in the outer region of the molecule, but it is also quite small in that region. On the other hand the region is very large, so it is difficult to make a quantitative estimate of this effect. If the effect is significant we expect the Becke functional to provide a positive contribution from the outer region, but maybe not large enough since it attains the asymptotic behaviour at too large distances [49, 32].

To analyse the properties of the functions δe_x and δe_c and determine from which region in the molecule the bonding energy correction originates (bond midpoint region or outer asymptotic region) we numerically integrated the functions δe_x and δe_c in different spatial regions. The procedure and results will be discussed in the next section.

10.5   Numerical procedure and results

In this section we will carry out a numerical investigation of the energy density difference functions δe_x and δe_c. In order to do this we performed several self-consistent density functional calculations on some dissociating diatomic molecules using the Becke and Perdew nonlocal corrections. The bonding contribution analysis is carried out using integration schemes in two different types of grid. The first scheme uses an integration in prolate spheroidal coordinates in which we integrated ellipsoidal regions around the molecular bond midpoint and in which the atomic positions coincide with the foci of the ellipsoid. If the atoms A and B are situated on the z-axis at positions $z = \pm a$ then we can define the coordinates u, v, ϕ by:

$$\begin{aligned}
x &= a\sinh(u)\sin(v)\cos(\phi) \\
y &= a\sinh(u)\sin(v)\sin(\phi) \\
z &= a\cosh(u)\cos(v)
\end{aligned} \tag{10.42}$$

As the functions δe_x and δe_c are invariant for rotations around the z-axis the ϕ-integration just yields a factor 2π. We define further a variable $p = \cos(\phi)$. The Jacobian in this coordinate system is given by:

$$J(u, p) = a^3 \sinh(u)(\cosh^2(u) - p^2) \tag{10.43}$$

We then want to calculate the following functions:

$$\delta I_k(R) = 2\pi \int_0^U du \int_{-1}^1 dp J(u, p) \delta e_k(u, p) \tag{10.44}$$

where k stands for exchange $k = x$ and correlation $k = c$ and in which $R = a \sinh(U)$ is the minor axis of the ellipsoid. These functions which represent the bonding energy contribution of the exchange and correlation energies within an ellipsoid around the bonding axis are calculated using Gauss-Legendre integration in the variables u and p.

The second integration scheme uses a cylindrical coordinate system defined by:

$$\begin{aligned} x &= r\cos(\phi) \\ y &= r\sin(\phi) \\ z &= z \end{aligned} \tag{10.45}$$

If the atoms are again situated at positions $z = \pm a$ on the z-axis then r represents the distance to the line through the bonding axis which is the z-axis and ϕ describes the rotation angle around this axis. For this coordinate system we want to calculate the following functions:

$$\delta I_k(Z) = \int_{-Z}^Z dz \int_0^\infty dr \delta e_k(r, z) J(r, z) \tag{10.46}$$

The subindex k again stands for exchange and correlation $k = x, c$ and the Jacobian J is given by $J(r, z) = 2\pi r$. The functions δI_k are again obtained using Gauss-Legendre integration. Using these two integration schemes it is possible to locate the regions in the exchange- and correlation energy density differences δe_x and δe_c which have the largest contributions to the bonding energy. Using the elliptic integration we can compare the contribution of the region outside some ellipsoid to the contribution of the inside and thus determine the importance of the outer asymptotic region. Using the integration in cylindrical coordinates we can determine the importance of the bond midpoint region.

In the following we carry out an analysis of the Becke and Perdew functionals in the nitrogen molecule N_2. The nitrogen molecule is a typical case of a molecule where the Becke correction gives a lowering of the bonding energy and where the Perdew correction gives a smaller increase in the bonding energy. Most molecules (for instance all first and second row diatomic molecules) exhibit this behaviour. A notable exception is the hydrogen molecule H_2 for which both the Becke and Perdew corrections have a different sign. For the nitrogen molecule an LDA calculation at the experimental bond distance of 2.07 bohr yields a bonding energy of 11.5 eV which is to be compared to the experimental bonding energy of 9.9 eV. A self-consistent calculation at the same bond distance using the Becke correction gives a bonding energy of 10.0 eV and a Becke-Perdew calculation yields 10.2 eV. We therefore have $\Delta E_x^{NL} = 1.5$eV and $\Delta E_c^{NL} = -0.2$ eV at equilibrium distance. In these calculations we used a triple-zeta Slater Type Orbital (STO) basis set plus polarisation and as atomic reference we used spin unrestricted nitrogen atoms in which we occupied (according to Hunds rule) the $2p$-shell with three electrons with the same spin. The atomic exchange energy density is then given by a sum of different contributions from α- and β-spin electrons. Our results

Figure 10.1: Becke and Perdew potential around the bond midpoint region for the dissociating nitrogen molecule at bondlengths of 3.0,5.0,7.0 and 9.0 bohr

are in substantial agreement with the completely numerical results of Becke [141].

In figures 10.1a and 10.1b we display the Becke and Perdew potential of the nitrogen molecule along the bond axis around the bond midpoint region for several bond distances. In figures 10.1c and 10.1d we display the same potentials along an axis perpendicular to the bond axis through the bond midpoint. As we can see from these figures the Becke peak is larger than the Perdew peak at each distance and the peaks are increasing in absolute value as we dissociate the molecule. We can also see that the width of the peak in the direction perpendicular to the bond axis is much larger than the width of the peak along the bond axis. The total potential (Becke+Perdew) therefore builds a repulsive wall between the atoms.

In figures 10.2a an 10.2b we display the exchange and correlation energy density difference functions δe_x and δe_c for the Becke and Perdew functional along the bond axis for the equilibrium distance of 2.07 bohr and the larger distance of 3.0 bohr. In figures 10.2c and 10.2d we display the same functions along an axis through the bond midpoint perpendicular to the bond axis. As we can see from figures 10.2a and 10.2c the function δe_x is positive in the bond midpoint region as was expected from equation 10.35. This positive region arises because the gradient of the molecular density is zero at the bond midpoint. As the gradient of the sum of atomic densities is also zero in the bond midpoint this positive region in δe_x is not the result of density changes due to bond formation but merely a result of the topological properties of the molecular density. In general δe_x is positive whenever the gradient of the molecular density vanishes. Due to the A1-symmetry of the molecular density this often happens in points which belong to the invariant manifold of the molecular symmetry group. In figures 10.2a and 10.2b we can also see peaks in δe_x around the nuclei. As the gradient of the sum of atomic densities in the region of one atom is close to the gradient of the density of the atom itself in this region these peaks can only be explained by density changes in the atomic region due to bond formation. This point is illustrated in figure 10.3a which displays the deformation density due to bond formation along the bond axis. We can see from this figure that the gradients of the deformation density are the largest around the nuclei. We have therefore identified two possible contributions to the Becke correction to the bonding energy. There is a contribution from the region around the nuclei which arises from density changes induced by bond formation and there is a contribution from the bond midpoint region which arises from the shape of the molecular density. The relative importance of these regions will be discussed below. Having discussed the properties of δe_x it is clear that the shape of δe_c as displayed in figures 10.2b and 10.2d can be explained in a similar way. The negative peak at the bond midpoint is due to the zero gradient in the bond midpoint (see equation 10.36) and the peaks closer to the nuclei are caused by density changes in the atomic regions.

After having discussed the local properties of δe_x and δe_c in different regions of space we will now investigate how much the various regions contribute to the bonding energy at equilibrium bondlength using the integration schemes described in the beginning of this section. In figure 10.4a we display the functions δI_x and δI_c of equation 10.44 as a function of the length R of the minor axis of the ellisoid which has the atoms positioned at the foci. As we can see from this figure half of the nonlocal correction to the bonding energy due to exchange is obtained for $R = 1$ bohr. The major axis then has a length of 1.44 bohr which therefore extends 0.40 bohr beyond the atoms. More than about 80% of the bonding energy contribution is obtained at $R = 1.5$ bohr with major axis of length 1.82 bohr which extends 0.79 bohr beyond the atoms. The ellipsoid is then just enclosing the valence region of the atoms. We therefore conclude that it is mainly from the inner region of the molecule and not from the asymptotic region that the Becke correction originates, the precise ratio of the contributions depending of course on the admittedly somewhat arbitrary

Figure 10.2: The exchange and correlation energy density differences of Becke and Perdew around the bond midpoint region

Figure 10.3: The deformation density of the nitrogen molecule at the equilibrium bondlength of 2.07 bohr

definition of these regions.

Let us first consider the asymptotic region somewhat more closely. The relatively small contribution from the asymptotic region, where the expression of equation 10.41 for δe_x holds, may be related to figure 10.3b which displays the deformation density along an axis through the bond midpoint perpendicular to the bond axis. From this figure we see that when the deformation density $\delta\rho$ is negative (which is the case in the outer asymptotic region) it is very small and therefore leads to a small asymptotic δe_x. However, the asymptotic contribution is not negligible and we may even wonder whether the division of the exchange energy over the inner and outer region of the molecule given by the Becke functional reflects the true situation as there are clear indications that the Becke functional underestimates the importance of the outer region. It has been demonstrated by Engel et al. [49] for the case of atoms that the exchange energy density of Becke reaches its Coulombic behaviour much further from the atomic nucleus (beyond 10 bohr) than the exact exchange energy density which reaches this behaviour just past the outermost orbital. Also the Becke exchange potential has a much too small contribution in the asymptotic region [49, 32] which is reflected in the fact that the LDA one-electron energies are not improved by the Becke potential[32]. The consequences of the too fast decay of the Becke exchange potential for the asymptotic contribution to the exchange energy can be inferred from the following definition of exchange energy density. We define the Levy-Perdew exchange energy density $e_x^{Langreth Perdew:wavevector}$ [22] in terms of the exchange potential v_x and the density ρ by:

$$e_x^{LP}(\mathbf{r}) = v_x(\mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{10.47}$$

This exchange energy density integrates to the exchange energy, which follows from a scaling property of the exchange functional [22].

$$E_x[\rho] = \int e_x^{LP}(\mathbf{r})d\mathbf{r} \tag{10.48}$$

As the LDA+Becke exchange potential is much smaller in the outer region of the molecule than the exact exchange potential it gives a too low contribution from this region. As a corollary, the

Figure 10.4: Contributions of the Becke and Perdew functional to the bonding energy in the elliptic (a) and cylindrical (b) coordinate system as a function of the minor elliptic axis (a) and the Z-coordinate (b) at the equilibrium bondlength of 2.07 bohr

roughly 20% of the total positive nonlocal contribution to the bond energy that comes from the outer region (in spite of the small $\delta\rho$ in that region) is certainly too low.

Let us next consider the inner region of the molecule. In this region there are two contributions to the exchange energy, there is a contribution from the bond midpoint region and a contribution from the atoms. To determine the relative importance of the atomic regions and the bond midpoint region we integrated δe_x using the cylindrical coordinate system. In figure 10.4b the functions δI_x and δI_c of equation 10.46 are plotted as a function of Z. The integration regions are thus slabs of thickness $2Z$ perpendicular to the bond axis centered around the bond midpoint. For $Z = 1.035$ bohr the slab is just touching the nuclei. We can see from figure 10.4b that at $Z = 0.5$ bohr which is about halfway the distance from the bond midpoint to the nucleus about half of the Becke correction to the exchange energy is obtained. If we denote the region $Z > 0.5$ bohr (somewhat arbitrarily) as atomic region then we can conclude that about half of the Becke correction is due to a replacement of high atomic gradients by a region of low molecular gradients (equation 10.35) around the bond midplane and half is due to changes in the density on the atoms as a result of bond formation. Note that the peaks in δe_x around the atoms in figure 10.2 are only seemingly more important than the bond midpoint peak in δe_x as they are only restricted to a small area around the nucleus whereas the bond midpoint peak also has a considerable extension in the direction perpendicular to the bond axis as demonstrated in figure 10.2c.

About the same conclusions as for Beckes exchange correction can be drawn for Perdews correlation correction. As can be seen from figure 10.2b the function δe_c has a somewhat larger extension than δe_x into the outer region of the molecule. However from figures 10.4a and 10.4b we still can conclude

that the bond midpoint region ($Z < 0.5$ bohr) contributes about half of the nonlocal correlation correction to the bonding energy.

We may wonder, as we did for the asymptotic region, whether the nonlocal corrections from the inner region are given correctly by the Becke-Perdew functional. The analysis of section 10.3 indicates that the Becke-Perdew potentials are overestimated in the bond midpoint region and therefore probably also the positive contribution to the bond energy. This fits in with the suggestion above that the contribution from the asymptotic region may be underestimated, since the total Becke-Perdew correction yields quite accurate bond energies. The conclusion that the contribution from the inner region is overestimated is supported by the fact that for solid state bulk systems (which of course exhibit no asymptotic region) the Becke correction often overshoots and leads instead of the overestimation by LDA to an underestimation of the cohesive energy [136, 137].

10.6   Conclusions

In this work we analysed the working of some gradient corrected density functionals in the case of chemical bonding. From an analysis of the exact Kohn-Sham potential based on the work of Buijse et al. [86] it follows that the exact exchange-correlation potential is peaked in the bond midpoint region. This fact can be related to left-right correlation in the dissociating molecule. It is shown that the Becke-Perdew potential also exhibits a peak in the bond midpoint region but with a wrong functional dependence on the density which makes this peak go to infinity at infinite separation of the atoms instead of going to a constant. Then we showed that one important reason that gradient corrected density functionals reduce the bonding energy in overbinded LDA-molecules is that these functionals replace a region of high gradients in the sum of atomic exchange-correlation energy densities by a region of low gradients in the molecular exchange-correlation energy density around the bond midpoint. This region of zero gradients arises purely due to the topological properties of the molecular density and is not related to density changes upon bond formation. We further argued that probably the relative importance of the bond midpoint contribution as compared to the asymptotic contribution to the total nonlocal correction is not represented completely correctly by the Becke-Perdew exchange-correlation energy densities. It appears that the positive contribution from the bond region is overemphasised particularly in weak-bonding (near dissociation) situations, in keeping with the overestimation of the bond midpoint peak in the potential in weak bonding situations. On the other hand the contribution from the asymptotic region is probably underestimated, in keeping with the too fast decay to zero of the Becke-Perdew potentials [32].

In order to improve upon existing density functionals and to obtain further insight in the relation between electron correlation and Kohn-Sham potentials we feel one should closely investigate the relation between the structure of the exchange-correlation potential (shell- and bond midpoint peaks, asymptotic Coulombic behaviour) and conditional amplitudes. The kinetic potential in atoms for instance measures the in-out correlation effect and induces some structure in the atomic correlation potential, just as it reflects the left-right correlation in the peak at the bond midpoint. It will obviously be very helpful to have exact molecular Kohn-Sham potentials available not just for the two-electron H_2 system but also for general many-electron molecules. This may be feasible with the procedure proposed in ref. [32]. From there one might endeavour to construct very accurate exchange-correlation potentials from the density and its derivatives. Work along this line is in progress.

Chapter 11

Exchange and correlation: a density functional analysis of the hydrogen molecule

In this work we analyze the exact quantities T_s, E_x and E_c from density functional theory (DFT) for the dissociating hydrogen molecule.These quantities obtained from configuration interaction calculations are compared with the approximate quantities within the local density approximation (LDA) and within LDA with nonlocal corrections (LDA+NL). We also make a comparison of the densities obtained within LDA and within the Hartree-Fock (HF) approximation. We further analyze the atomic fragments in the dissociation limit within the LDA and Local Spin Density (LSD) approximation. We conclude that LDA+NL gives a very good approximation to the exact exchange and correlation energies at equilibrium distance, but a not so good approximation for these quantities at large bond distances due to a incomplete cancelation of the Hartree potential by the exchange-correlation potential in the atomic fragments in the limit of infinite separation. This error can be corrected by using the LSD and LSD+NL approach. We also demonstrate that the present correlation functionals do not give a proper account of the correlation energy in the dissociation limit, an error which is partly compensated by an overestimation of the exchange energy providing a good total exchange-correlation energy. We will however show that conventional definition of correlation energy within DFT is not a useful one in describing the physics of molecular dissociation but one should rather consider the total exchange-correlation energy to describe the physical situation. We analyze how several types of density changes due to bond formation affect the various LDA and LDA+NL exchange and correlation energies, explaining the sign of the corrections to the bond energy.

11.1   Introduction

Density functional theory [5] has become a very successful method in the calculation of electronic properties of systems ranging from atoms to molecules [70, 122] and solids [47]. Especially in the last few years after the the introduction of the so-called Generalized Gradient Approximations (GGA) [44, 45, 43, 46, 50] or nonlocal (NL) corrections density functional theory has gained wide popularity within the field of quantum chemistry through its ability to yield very good bond energies which can compete with those obtained from other correlated quantum chemical methods. It has for instance been widely applied to the field of transition metal chemistry [71] where many systems

can not be studied by more conventional methods for practical reasons. Observing this large field of application within chemistry it would be very desirable to investigate whether the physics of for instance bond breaking is correctly described by the presently used density functionals. Such an investigation would not only contribute to a better understanding of the physics behind the phenomena in question but might also lead to the development of more accurate density functionals. For this reason we calculated the some of the quantities which are used in DFT exactly for the case of dissociating hydrogen molecule and compare them with the approximate ones obtained from LDA and LDA+NL. The hydrogen molecule is one of the first molecules used as a test for the LSD method [142, 143]. Its Kohn-Sham potential has been constructed and thoroughly analyzed [86]. An analysis of the exact exchange-correlation energies has however not been presented. Bond breaking is a specially interesting testcase for the presently used functionals as here correlation effects (especially left-right correlation effects) are large.

We further want to analyze how the various exchange and correlation functionals (LDA and GGA) contribute to the bond energy. In order to understand this we also studied the atomic fragments in the infinite separation limit. A question we like to answer in this paper is why the Becke-Perdew nonlocal exchange-correlation correction as a rule gives a positive (repulsive) contribution for the bond energy (as in all homonuclear first and second row diatomic molecules) with the exception of the hydrogen molecule which is the subject of this paper. In order to answer this question we analyze how the exchange and correlation contributions to the bond energy are changed by certain density variations resulting from bond formation.

The paper is divided as follows. In section 11.2 we give a short account of the exact quantities we want to calculate. In section 11.3 we discuss some properties and give some numerical results for the atomic fragments in the dissociation limit in order to understand the role of the atomic error in bond breaking. In section 11.4 we discuss the exact results for the DFT quantities as a function of the bond distance obtained from accurate calculations. In section 11.5 we analyze the bond energy contribution of several exchange and correlation functionals and analyze how they are influenced by density variations. In section 11.6 we finally present our conclusions.

11.2   Exchange and correlation in density functional theory

In this section we define some of the quantities we will calculate in section 11.4. The energy functional E_v within density functional theory is usually defined by the constrained search approach of Levy and Lieb [67, 68]:

$$E_v[\rho] = \int \rho(\mathbf{r})v(\mathbf{r}) + F_L[\rho] \tag{11.1}$$

where the functional F_L is defined as:

$$F_L[\rho] = \min_{\hat{D} \to \rho} Tr\{\hat{D}(\hat{T} + \hat{W})\} \tag{11.2}$$

which minimizes the trace value of the kinetic energy T and the interparticle interaction W with the set of N-particle density matrices $\hat{D}$ yielding density ρ. The above formulation is important for the consideration of degenerate ground states where $\hat{D}$ can be a linear combination of density matrices corresponding to pure states within the ground state multiplet. This is of importance for instance when a molecule dissociates into open shell atoms with degenerate ground states. Moreover E_v is also invariant under unitary transformations within the multiplet of degenerate wavefunctions. We will return to this point again in section 11.3 when we discuss the infinite separation limit of the

dissociating hydrogen molecule. The functional F_L is universal in the sense that it can in principle be calculated for each density ρ independent of the external potential v. From its definition it is also clear that F_L is invariant under rotations and translations of the density ρ, providing important constraints for approximate exchange and correlation functionals [32]. Within the Kohn-Sham approach to density functional theory the functional F_L is usually split up as follows:

$$F_L[\rho] = T_L[\rho] + \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\rho] \tag{11.3}$$

in which T_L is the Kohn-Sham kinetic energy which is the kinetic energy of a noninteracting particle system with the same density ρ as the interacting system ,defined by constrained search as:

$$T_L[\rho] = \min_{\hat{D} \to \rho} Tr\{\hat{D}\hat{T}\} \tag{11.4}$$

If the ground state Kohn-Sham wavefunction can be represented by a pure state density matrix corresponding to a single Kohn-Sham determinant we will denote T_L by T_s. We will do this for instance for the hydrogen molecule. As we now have defined all functionals except E_{xc} equation 11.3 actually defines E_{xc}. The Euler-Lagrange equations obtained by minimizing E_v with respect to the density can be written in the form of the well-known Kohn-Sham equations. For future reference we give the spinpolarized equations with spin index σ:

$$(-\frac{1}{2}\nabla^2 + v_{s,\sigma}(\mathbf{r}))\phi_{i,\sigma}(\mathbf{r}) = \epsilon_{i,\sigma}\phi_{i,\sigma}(\mathbf{r}) \tag{11.5}$$

where

$$\rho(\mathbf{r}) = \sum_\sigma \rho_\sigma(\mathbf{r}) = \sum_\sigma \sum_{i=1}^N f_{i\sigma}|\phi_{i,\sigma}(\mathbf{r})|^2 \tag{11.6}$$

where $f_{i\sigma}$ are the orbital occupation numbers (in general fractional for degerate ground states) and the $v_{s,\sigma}$ is the Kohn-Sham potential:

$$v_{s,\sigma}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1 + v_{xc,\sigma}([\rho]; \mathbf{r}) \tag{11.7}$$

where the exchange-correlation potential v_{xc} is the functional derivative of the exchange-correlation functional E_{xc} with respect to the density ρ_σ. We further define the exchange functional as:

$$E_x[\rho_\uparrow, \rho_\downarrow] = \sum_\sigma -\frac{1}{2} \int \frac{|\gamma_{s,\sigma}(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.8}$$

where $\gamma_{s,\sigma}$ is the Kohn-Sham density matrix:

$$\gamma_{s,\sigma}(\mathbf{r}_1, \mathbf{r}_2) = \sum_i^{N_\sigma} f_{i\sigma}\phi_{i,\sigma}^*(\mathbf{r}_1)\phi_{i,\sigma}(\mathbf{r}_2) \tag{11.9}$$

This is for a fixed density ensemble, i.e. for fixed occupation numbers $f_{i\sigma}$, a density functional [67, 68] as the Kohn-Sham orbitals and therefore also the Kohn-Sham one-particle density matrix are uniquely determined by the ensemble density (an explicit construction procedure has been given in [32]). The Kohn-Sham one-particle density matrices corresponding to different ground state ensemble densities are different and therefore they lead in general to different exchange energies.

The exchange potential $v_{x,\sigma}(\mathbf{r}) = \delta E_x / \delta \rho_\sigma(\mathbf{r})$ for spherical atoms has the following long range asymptotic behaviour [49, 35]:

$$v_{x,\sigma}(\mathbf{r}) \sim -\frac{f_{m,\sigma}}{r} \qquad (r \to \infty) \tag{11.10}$$

where $f_{m,\sigma}$ is the occupation of the highest occupied Kohn-Sham orbital. From this formula it is clear that it depends on the occupation $f_{m,\sigma}$ of the highest occupied Kohn-Sham orbital to which amount the self-energy is being cancelled. For nondegenerate ground states one always has $f_{m,\sigma} = 1$ and for this case the self-energy is always being cancelled. This has important implications for dissociating molecules which dissociate into open shell fragments that have degenerate ground states as will be discussed in the next section.

For closed shell systems we can simplify the above exchange functional as:

$$E_x[\rho] = -\frac{1}{4} \int \frac{|\gamma_s(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.11}$$

where $\gamma_s = \sum_\sigma \gamma_{s,\sigma}$. Finally the correlation energy functional is defined as:

$$E_c[\rho] = E_{xc}[\rho] - E_x[\rho] \tag{11.12}$$

Note that the definition of exchange-correlation energy in other quantum chemical methods ,which we will denote by W_{xc}, differs from the above definition in density functional theory, the relation being:

$$W_{xc}[\rho] = E_{xc}[\rho] - T_{xc}[\rho] \tag{11.13}$$

where T_{xc} is given by:

$$T_{xc}[\rho] = T[\rho] - T_s[\rho] \tag{11.14}$$

and where T is the kinetic energy functional which for nondegenerate ground states is defined as the expectation value of the kinetic energy operator with the wavefunction which minimises F_L within constrained search, which is just the ground state kinetic energy if we insert the ground state density. Note that $T[\rho]$ is only defined for nondegenerate ground state densities as otherwise ρ could be generated by two different degenerate ground state wavefunctions, corresponding to the same total energy but having different expectation values with the kinetic energy operator [5]. For a density corresponding to a nondegerate ground state W_{xc} is just the expectation value of the ground state wavefunction with the interparticle interaction operator $\hat{W}$ minus the Coulomb repulsion energy. It is readily shown using equations 11.2 and 11.4 that $T_{xc}[\rho] \geq 0$ so:

$$E_{xc}[\rho] \geq W_{xc}[\rho] \tag{11.15}$$

We have now defined all the quantities we need for our analysis of the dissociating hydrogen molecule. In one of the following sections we will compare the above defined functionals T_s, E_x and E_c calculated from accurate Configuration Interaction calculations with those of the Local Density Approximation (LDA) and with the Generalized Gradient Approximation (GGA) for several bond distances.

11.3 The infinite separation limit

For the case of a hydrogen molecule the Kohn-Sham equations (in the spin restricted case) reduce to the following differential equation involving only the electron density:

$$(-\frac{1}{2}\nabla^2 + v_s([\rho];\mathbf{r}))\sqrt{\rho(\mathbf{r})} = \epsilon\sqrt{\rho(\mathbf{r})} \tag{11.16}$$

The Kohn-Sham kinetic energy is for a two-electron system given by:

$$T_s[\rho] = \frac{1}{8}\int \frac{(\nabla\rho)^2}{\rho}d\mathbf{r} \tag{11.17}$$

The exchange functional in this case is then given by:

$$E_x[\rho] = -\frac{1}{4}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.18}$$

and the corresponding exchange potential :

$$v_x([\rho];\mathbf{r}) = \frac{\delta E_x}{\delta\rho(\mathbf{r})} = -\frac{1}{2}\int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|}d\mathbf{r}_1 \tag{11.19}$$

which is just minus onehalf times the Hartree potential. So the Kohn-Sham potential becomes:

$$v_s([\rho];\mathbf{r}) = v(\mathbf{r}) + \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|}d\mathbf{r}_1 + v_c([\rho];\mathbf{r}) \tag{11.20}$$

If we neglect the correlation potential v_c then we just obtain the Hartree-Fock equations or the optimized potential model [33, 35] equations for this two-electron system. This is of course a bad approximation for cases where correlation effects are large such as in the dissociation limit where left-right correlation effects are large. In this limit the Kohn-Sham potential can be approximated by the sum of atomic Kohn-Sham potentials and the Kohn-Sham potential becomes:

$$v_s([\rho];\mathbf{r}) \approx v_s([\rho_A];\mathbf{r} \approx R_A) + v_s([\rho_B];\mathbf{r} \approx R_B) \tag{11.21}$$

where the atomic nuclei are on positions R_A and R_B. The last step in this approximation is only valid in the regions of the atoms and not in the bond midpoint region. In this region the Kohn-Sham potential v_s is small in the dissociation limit and approximating this by the sum of atomic densities might be a bad approximation. This is reflected by the fact that a peak structure is develops in the Kohn-Sham potential [86, 90] indicating strong left-right correlation effects. In the region of each atom we have :

$$v_s([\rho];\mathbf{r} \approx R_A) = v_A(\mathbf{r}) + \frac{1}{2}\int \frac{\rho_A(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|}d\mathbf{r}_1 + v_c([\rho];\mathbf{r} \approx R_A) \tag{11.22}$$

where v_A is the potential of the atomic nucleus. For the hydrogen molecule the atoms are just the hydrogen atoms which are one electron systems for which the Kohn-Sham potential should of course just be the nuclear potential v_A. It follows that the correlation potential v_c should be equal to minus onehalf the Hartree potential and thus be equal to the exchange potential in this limit. The Hartree-Fock approximation is obtained by neglecting this correlation potential and therefore the potential is too repulsive making the electron too weakly bound and the density too diffuse. This is a well-known error of the Hartree-Fock approximation which is usually solved by doing a spin unrestricted Hartree-Fock calculation. Let us now turn this Hartree-Fock problem for a more general homonuclear diatomic molecule. Suppose we consider dissociating closed shell diatomic

molecule and suppose that in the dissociation limit (where we can neglect the overlap of the atomic orbitals) we have M_σ bonding molecular orbitals $\phi_{i\sigma} = (\phi_{i\sigma}^A + \phi_{i\sigma}^B)/\sqrt{2}$ and their corresponding antibonding combinations $\phi_{i\sigma} = (\phi_{i\sigma}^A - \phi_{i\sigma}^B)/\sqrt{2}$,where $\phi_{i\sigma}^{A,B}$ are the atomic orbitals of atoms A and B, and $N_\sigma - M_\sigma$ molecular orbitals without corresponding antibonding combinations. Then the one-particle density matrix γ_σ of the molecule in the region of atom A becomes:

$$\gamma_\sigma^A(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i=1}^{M_\sigma} \phi_{i\sigma}^A(\mathbf{r}_1)\phi_{i\sigma}^{A*}(\mathbf{r}_2) + \sum_{i=M_\sigma+1}^{N_\sigma} \frac{1}{2}\phi_{i\sigma}^A(\mathbf{r}_1)\phi_{i\sigma}^{A*}(\mathbf{r}_2) \tag{11.23}$$

and the corresponding atomic exchange energy becomes:

$$E_x^A = \sum_\sigma -\frac{1}{2} \int \frac{|\gamma_\sigma^A(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.24}$$

The atom A in this limit has $N_\sigma - M_\sigma$ degenerate orbitals with occupation number $\frac{1}{2}$ in the outermost atomic shell. The corresponding one-particle density matrix can not be represented by a single determinant but it can be written as:

$$\gamma_\sigma^A(\mathbf{r}_1, \mathbf{r}_2) = \sum_i^{N_A} w_i \gamma_{i\sigma}^A(\mathbf{r}_1, \mathbf{r}_2) \tag{11.25}$$

where γ_i^A is the one-particle density matrix correponding to the i-th degenerate configuration in the N_A-fold degenerate ground state multiplet and $\sum_i^{N_A} w_i = 1$. The Hartree-Fock solution for the atom A in the infinite separation limit can be obtained by minimizing the energy:

$$E^{FHF} = \sum_\sigma -\frac{1}{2} \int \nabla_2^2 \gamma_\sigma^A(\mathbf{r}_1, \mathbf{r}_2)|_{1=2} d\mathbf{r}_1 d\mathbf{r}_2 + \int \rho(\mathbf{r}_1)v_A(\mathbf{r}_1)d\mathbf{r}_1 +$$

$$\frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - \sum_\sigma \frac{1}{2} \int \frac{|\gamma_\sigma^A(\mathbf{r}_1, \mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.26}$$

where γ_σ^A is given by 11.23. This procedure we will call fractional Hartree-Fock (FHF). We will present numerical results for this procedure for the case of the hydrogen atom in section 11.4. Application of this FHF method leads to an incomplete cancellation of the self-energy for open shell atoms, which is the cause of the Hartree-Fock error for dissociating molecules.

Before we go further into this point we first discuss the behaviour of the Local Density Approximation in this limit. Within LDA the exchange energy is given by:

$$E_x[\rho_\uparrow, \rho_\downarrow] = \sum_\sigma -c_x \int \rho_\sigma(\mathbf{r})^{\frac{4}{3}} d\mathbf{r} \tag{11.27}$$

where the constant c_x is given by $\frac{3}{4}(\frac{6}{\pi})^{\frac{1}{3}}$. For the spin-restricted case we obtain:

$$E_x[\frac{1}{2}\rho, \frac{1}{2}\rho] = -\frac{c_x}{2^{\frac{1}{3}}} \int \rho(\mathbf{r})^{\frac{4}{3}} d\mathbf{r} \tag{11.28}$$

The LDA correlation potential is much smaller than the exchange potential so we will neglect this potential for the moment and make a comparison with the Hartree-Fock approximation. Doing a restricted LDA calculation for an open shell atom is similar to doing an FHF calculation, one

Figure 11.1: Minus the Hartree potential and the exchange-correlation potential for the hydrogen atom from LDA and LSD calculations as a function of the distance from the atomic nucleus

Figure 11.2: Minus half of the Hartree potential and the exchange-correlation energy density for the hydrogen atom from LDA and LSD calculations as a function of the distance from the atomic nucleus

H-atom	LDA	LDA (ρ_e)	LSD	LSD (ρ_e)	KS	FHF
ϵ	-0.2335		-0.2690		-0.5000	-0.2295
T_s, T^{HF}	0.4250	0.5000	0.4666	0.5000	0.5000	0.3578
E_{Ne}	-0.9210	-1.0000	-0.9656	-1.0000	-1.0000	-0.8437
E_C	0.2828	0.3125	0.2984	0.3125	0.3125	0.2565
E_{tot}	-0.4457	-0.4416	-0.4787	-0.4776	-0.5000	-0.3577
E_{tot}+NL	-0.4650	-0.4629	-0.4999	-0.4996		
E_x	-0.1931	-0.2127	-0.2564	-0.2680	-0.1563	-0.1283
E_x+NL	-0.2357	-0.2588	-0.2968	-0.3097		
E_c	-0.0394	-0.0414	-0.0217	-0.0221	-0.1563	
E_c+NL	-0.0161	-0.0166	-0.0025	-0.0024		
E_{xc}	-0.2325	-0.2541	-0.2781	-0.2901	-0.3125	
E_{xc}+NL	-0.2518	-0.2754	-0.2993	-0.3121		

Table 11.1: Energies of the atomic fragment in the infinite separation limit

just fractionally occupies each of the degenerate LDA Kohn-Sham orbitals in the outermost atomic shell. The exchange-only LDA Kohn-Sham potential for atom A is then given by:

$$v_s([\rho_A]; \mathbf{r}) = v_A(\mathbf{r}) + \int \frac{\rho_A(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1 - \frac{4}{3} \frac{c_x}{2^{\frac{1}{3}}} \rho_A(\mathbf{r})^{\frac{1}{3}} \qquad (11.29)$$

Also this exchange-potential gives an incomplete cancelation of the Hartree potential, however the cancelation is better than in the FHF case. More important, as has been pointed out by Gunnarsson et al. [144] and shown in figure 11.1, is the fact that the difference between the LDA exchange correlation potential and the Hartree potential is an almost constant function. A constant shift (of about 0.2 a.u.) in the potential affects the eigenvalue but not the density. Note that this constancy is approximately independent of the nuclear charge as both the Hartree and the exchange potential (which is the main part of the exchange-correlation potential) have a similar scaling property with respect to the nuclear charge.

If we look at the energy expressions in LDA we see a much better cancelation of the exchange-correlation energy and the electron repulsion than for the corresponding potentials. This has been displayed for in figure 11.2 where we plot the LDA (and LSD to be discussed later on) exchange-correlation density ϵ_{xc} and minus half of the Hartree potential $-v_H$. The most important difference between a LDA and an Hartree-Fock calculation is therefore the difference in energy expressions for the exchange energy leading to a better cancelation of the self-energy of the atoms in the infinite separation limit. The importance of this difference in energy expressions is apparent from table 11.1 in which we display the several energy terms of the hydrogen atomic fragments in the infinite separation limit for the self-consistent LDA ,LSD, FHF approximations and for the exact Kohn-Sham values. To show the dependency on the quality of the density we also included the LDA and LSD results calculated from the exact density ρ_e. As we can see from this table both LDA and HF make an error in the electron nuclear attraction energy and in the kinetic energy which are almost equal but of opposite sign, which leaves a small error. Now we note the importance of the different energy expressions for the exchange energy. The LDA exchange energy is roughly equal to the exact exchange-correlation energy, whereas the LDA correlation energy is much smaller than the exact correlation energy, which leads to a rather good value of the exchange-correlation energy and a cancellation of the electron repulsion energy. This then gives a good approximation of the total atomic energy. The HF exchange energy is however only half of the electron repulsion energy

and therefore gives a bad estimate of the total atomic energy.

The Hartree-Fock results for the dissociating hydrogen molecule can be improved (an in fact will give the exact limit for this special case) by performing an unrestricted Hartree-Fock (UHF) calculation in which one breaks the spin symmetry and localizes one spin up electron on one atom and a spin down electron on the other atom. A similar procedure can be carried out in density functional calculations using a spin density functionals. A better description than LDA for the dissociated atom in the infinite separation limit is obtained in the Local Spin Density (LSD) approximation. This method has been thorougly investigated by Gunnarsson and Lundqvist [127]. In spin density functional theory one uses different potentials for the spin up and spin down electrons. As in UHF this leads to a violation of the spin rotational symmetry. Using LSD one can for instance solve the Kohn-Sham equations for the hydrogen atom containing only one electron with quantum number $S_z = \pm 1$. Performing a spin rotation which leads to a different mixture of spin densities will give a different LSD energy. We will now investigate this problem more closely.

The exact ground state solution for the hydrogen atom A at infinite separation ,as obtained from singlet ground state wavefunction of H_2, is given by:

$$\psi(\mathbf{r}, s) = \sqrt{\rho_A(\mathbf{r})}(c_1\alpha(s) + c_2\beta(s)) \tag{11.30}$$

with $c_1 = 1/\sqrt{2}$ and $c_2 = i/\sqrt{2}$. The ground state energy should not change when we perform an unitary transformation on the vector (c_1, c_2) and the energy should therefore not depend on the spinpolarisation [145, 146] of the two degenerate spin states. In general ,if we have an electron density ρ which is a linear combination of densities $\rho_i = \sum_\sigma \rho_{i\sigma}$ corresponding to pure eigenstates within one M-fold degenerate ground state multiplet, i.e:

$$\rho(\mathbf{r}) = \sum_i^M w_i \rho_i(\mathbf{r}) \tag{11.31}$$

with $\sum_i^M w_i = 1$, then any density obtained from the above density by a unitary transformation within the space of degenerate states should yield the same ground state energy. This is an immediate consequence of the definition of the functional F_L. This means for the example of the hydrogen atom that the spinrestricted density which is a linear combination ($w_1 = w_2 = \frac{1}{2}$ in equation 11.31) of the degenerate pure spin up and spin down densities ($(c_1, c_2) = (1, 0)$ and $(0, 1)$ in equation 11.30) should yield the ground state energy for the hydrogen atom. As the invariance property for unitary transformations within the degenerate ground state multiplet is not fulfilled for the presently used approximate density functionals they do not decribe properly the dissociation of closed shell molecules into open shell atoms with degenerate ground states. These atomic fragments in the dissociation limit are spinrestricted atoms with fractional occupation (as in the FHF case, see equation 11.23) of the degenerate orbitals. This is also true in the exact Kohn-Sham case for which the corresponding density then corresponds to an ensemble (see equation 11.31) of pure state multiplet densities which gives the exact ground state energy of the atom when inserted in the energy functional E_v. The total energies of the dissociated atomic fragments in the LDA case however are too large, but smaller than the Hartree-Fock (FHF) atoms in this limit. As an unrestricted LSD calculation lowers the atomic energy this yields an improvement of the total atomic energy although such a correction should not be needed from a theoretical points of view. From table 11.1 we can see the LSD improvements in the density reflected in the good estimates for electron nuclear attraction energy and the kinetic energy, leading to a very good estimate of the total energy especially if the nonlocal corrections are included.

After studying the infinite separation limit we will in the next section study some exact density functional quantities as a function of the bond distance.

Figure 11.3: Electron-electron repulsion energy as a function of the bond distance

11.4   Numerical results

In order to calculate exact Kohn-Sham quantities T_s, E_x and E_c for the hydrogen molecule we performed some accurate Configuration Interaction (CI) calculations at several bond distances. We compare these accurate results with Hartree-Fock and self-consistent density functional calculations within the Local Density Approximation. The nonlocal exchange functional of Becke [44] and the nonlocal correlation functional of Perdew [45, 147] where added on top of the LDA calculation (using the VWN-correlation potential [85]) in a post-scf manner. The effect of not including the Becke and Perdew functionals in a selfconsistent manner is negligible [139] as the influence of these potentials on the density is very small.

We first look at some density dependent quantities. In figure 11.3 we plotted the Coulomb repulsion energy of the electronic cloud given by:

$$E_C[\rho] = \frac{1}{2} \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{11.32}$$

calculated from the HF, CI and LDA densities. We can see from this figure that the Coulomb repulsion energy approaches much earlier the $1/R(R \to \infty)$ behaviour, which is just the repulsion of point charges with charge one, than the HF and LDA densities. That the LDA graph is closer to the CI graph than the HF graph is in accordance with the result of the previous section that the HF-density density is the most diffuse one. This fact is supported by figure 11.4 which plots the electron nuclear attraction energy:

$$E_{Ne}[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \tag{11.33}$$

Figure 11.4: Electron nuclear attraction energy as a function of the bond distance

where v is the Coulomb potential of the atomic nuclei. We again can see that the correct asymptotic behaviour $-2 - 2/R(R \to \infty)$ is reached much earlier by the CI density, the worst density again being the HF-density.

Another measure of the quality of the electron density is the kinetic energy T and the Kohn-Sham kinetic energy T_s. In figure 11.5 we plotted the kinetic energies T^{CI} and T^{HF} and the Kohn-Sham kinetic energies T_s and T_s^{LDA}, where T_s is calculated from the CI-density. We can see from this figure that the graphs of T^{CI} and T_s are very close at all bond distances, the difference being the exchange-correlation part T_{xc} of the kinetic energy. The LDA kinetic energy T_s^{LDA} has a large deviation from T_s at large bond distances which approaches quickly the atomic error in this quantity. The HF kinetic energy at these bond distances then deviates even more from T^{CI} again reflecting the diffuseness of the density and the bad representation of the atomic fragments. Note however that at large distances the LDA and HF errors in the electron nuclear attraction energy and the kinetic energy are approximately equal and of opposite sign, giving a good approximation for the sum of those quantities.

In the vicinity of the equilibrium bond distance (1.401 bohr) the errors in the HF and LDA densities are small and in that case we have a good estimate of the energy terms described above.

In figure 11.6 we plotted the total energies for the CI, HF, LDA and LDA+NL with the nonlocal corrections of Becke and Perdew added. We also calculated the expection value of the LDA Kohn-Sham determinant, which is just the HF expression for the energy but with LDA orbitals inserted. We will denote the corresponding graph by HFLDA. We see that the curves HF, HFLDA and LDA are very close around the equilibrium distance. Also the CI and LDA+NL are very close in this bond distance region. At large bond distances all the curves deviate from the CI curve, the worst ones being HF and HFLDA (the HFLDA a little higher than HF because the LDA determinant

Figure 11.5: Kinetic energy and Kohn-Sham kinetic energy as a function of the bond distance

Figure 11.6: Total energy as a function of the bond distance

Figure 11.7: Exchange-correlation energy as a function of the bond distance

Figure 11.8: Exchange energy as a function of the bond distance

does not minimize the HF energy expression). The fact that the LDA curve is so much closer to the CI curve than the HFLDA curve is mainly due to a different expression for the exchange energy. The LDA exchange energy gives a better cancelation of the electron repulsion energy in the atomic fragments than the HF and HFLDA exchange energy which cancel only half of the electron repulsion energy. As the nonlocal corrections give an additional energy lowering the best curve at all distances is given by the LDA+NL curve.

To analyze the several exchange and correlation functionals more closely we plotted in figure 11.7 the exchange-correlation energies $E_{xc}, W_{xc}, E_{xc}^{LDA}$ and E_{xc}^{LDA+NL} as a function of the bond distance. From this plot we can see that indeed $E_{xc} \geq W_{xc}$ the difference being the exchange-correlation part T_{xc} of the kinetic energy. At large bond distances E_{xc} and W_{xc} approach $-E_C$ thereby canceling the electron repulsion energy. Due to incomplete cancelation of this term within LDA and LDA+NL exchange-correlation energies we expect the largest errors in these quanties at large separation. We can see from figure 11.7 that this is indeed the case. Both E_{xc}^{LDA} and especially E_{xc}^{LDA+NL} give a very good approximation to the exact Kohn-Sham exchange-correlation energy E_{xc} around the equilibrium distance. At larger distances both E_{xc}^{LDA} and E_{xc}^{LDA+NL} deviate from the exact E_{xc} the difference approaching fast the atomic error in these quantities.

We will now investigate the separate behaviour of the exchange and correlation energies. In figure 11.8 we plot the exchange energies E_x, E_x^{HF}, E_x^{LDA} and E_x^{LDA+NL} as a function of the bond distance. We can see the same trends as for the total exchange-correlation energies. We have a very good description of the exact exchange energy E_x at equilibrium distance especially by E_x^{LDA+NL} and large errors at large distance. The E_x^{LDA+NL} exchange will of course approach two times its value for the spinrestricted atoms in the large distance limit. However one should realize that the definition of exact exchange as defined in equation 11.8 is not a useful one to describe the physics of bond breaking as the corresponding exact exchange potential does not contain the complete self-energy correction of the atomic fragments. This means that the correlation potential which has to correct this should contain part of the self-energy which gives obviously a rather unphysical description of the correlation energy. As discussed this not only a feature of bond breaking but a general problem in the calculation of open shell systems within a spin restricted formalism. Let us look more closely at the correlation energies. In figure 11.9 we plot the correlation energies E_c, E_c^{LDA} and E_c^{LDA+NL}. We can see that E_c^{LDA} and E_c^{LDA+NL} have a very slow variation with respect to the bond distance. The E_c^{LDA} represents the VWN parametrization of electron gas correlation energy data giving a correlation energy which about a factor two too large in the molecule at equilibrium distance and in the atomic fragments. Addition of the Perdew gradient correction to the correlation energy gives very good values for the correlation energy for the molecule at equilibrium distance and in atoms and is also almost constant as a function of the bond distance. This is not the case for the correlation energy E_c as defined in equation 11.12 .The graph of E_c approaches just as E_x the value $-\frac{1}{2}E_C$ at large bond distances and becomes much larger than E_c^{LDA} and E_c^{LDA+NL}. This is due to the fact that E_c has to cancel half of self-energy. The approximate functionals E_c^{LDA}, which is fitted to electron gas data, and the E_c^{LDA+NL} , which is fitted to atomic data, are of course unable to imitate this behaviour.

In view of the inconvenient definition for exchange and correlation in the dissociation limit it is better to regard the total exchange-correlation energy. The total exchange-correlation energy of an atom is a quantity which automatically cancels the self-energy, as the exchange-correlation hole always will integrate to one electron, also for degenerate ground state wavefunctions. However E_{xc} is in general still different for different ground state ensemble densities but due to the correct inclusion of the self-energy these differences are smaller than for the exchange or correlation alone. These differences in E_{xc} can of course not be avoided as different degenerate ground state wavefunctions

Figure 11.9: Correlation energy as a function of the bond distance

will in general have different expectation values for any other operator than the total energy operator. From the above considerations it is clear that for the case of dissociating molecules it is much easier to find appropriate approximations for E_{xc} than for E_x and E_c as defined in equations 11.8 and 11.12.

We can therefore conclude that LDA+NL gives a very good description of the properties of the molecule (including exchange and correlation energies) at equilibrium distance. The remaining errors for LDA and LDA+NL are due to errors in the Coulomb repulsion and in the exchange-correlation energy and the molecular error is quickly approaching the atomic errors in this quantity. These is however not a problem in practical calculations as a spin unrestricted calculation with one spin density on one atom gives a good description of the total atomic energy and one can get good binding curves at all distances using spin unrestricted Kohn-Sham calculations similar to UHF calculations. This is however unsatisfactory from a theoretical point of view as ,due to unitary transformations among the degenerate ground states of the exact density functional E_v, spin restricted and spin unrestricted calculations on the hydrogen atom should give the same energy. In the next section we will investigate how and why several LDA and nonlocal (GGA) influence the bond energy as they do and how this is related to density changes.

11.5 An analysis of bond energy contributions

In this section we will investigate how the various exchange and correlation energy contributions (LDA and GGA) to molecular bonding energy are affected by certain density changes. We therefore construct the molecular density of the hydrogen molecule in two steps. First ρ_Σ, which is the sum

H_2	LDA		KS	HF
	ρ_Σ	ρ_λ		
T_s, T^{HF}	1.107	0.863	1.116	1.140	1.125
E_{Ne}	-3.602	-3.220	-3.556	-3.647	-3.631
E_C	1.297	1.128	1.257	1.321	1.316
E_{tot}	-1.137	-1.091	-1.109	-1.172	-1.133
E_{tot}+NL	-1.178	-1.126	-1.150		
E_x	-0.559	-0.488	-0.547	-0.661	-0.658
E_x+NL	-0.648	-0.566	-0.636		
E_c	-0.089	-0.089	-0.093	-0.039	
E_c+NL	-0.041	-0.046	-0.045		
E_{xc}	-0.648	-0.577	-0.640	-0.700	
E_{xc}+NL	-0.689	-0.612	-0.681		

Table 11.2: Energies of hydrogen molecule at equilibrium distance (R=1.401 bohr)

of atomic densities, is made. For the hydrogen molecule this density is given by:

$$\rho_\Sigma(\mathbf{r}) = \frac{1}{\pi}(e^{-2|\mathbf{r}|} + e^{-2|\mathbf{r}-\vec{R}|}) \tag{11.34}$$

where $\vec{0}$ and $\vec{R}$ are the positions of the atomic nuclei. The bond energy contribution of a quantity E corresponding to the density ρ_Σ we will denote by ΔE_Σ, so we have:

$$\Delta E_\Sigma = E[\rho_\Sigma] - E[\rho_A] - E[\rho_B] \tag{11.35}$$

In the second step there is a change from ρ_Σ to real molecular density which we will approximate by a density contraction:

$$\rho_\lambda(\mathbf{r}, \vec{R}) = \lambda^3 \rho_\Sigma(\lambda\mathbf{r}, \lambda\vec{R}) \tag{11.36}$$

which is just a scaling transformation. The bond energy contribution caused by this second density change we will denote by ΔE_λ:

$$\Delta E_\lambda = E[\rho_\lambda] - E[\rho_\Sigma] \tag{11.37}$$

A third contribution to the bond energy is caused by going from spin restricted to spin unrestricted atoms. This energy change is not caused by a change in the spatial electron density but by a change in the electron spin densities. This third energy change we will denote by ΔE_s. We will now discuss how the two described density changes and the spin density change affect the bond energy contributions of the LDA and GGA exchange and correlation functionals.

LDA exchange. The bond energy contribution $\Delta E_{x,\Sigma}$ for this case is given by:

$$\Delta E_{x,\Sigma} = -c_x \int ((\rho_A(\mathbf{r}) + \rho_B(\mathbf{r}))^{\frac{4}{3}} - \rho_A(\mathbf{r})^{\frac{4}{3}} - \rho_B(\mathbf{r})^{\frac{4}{3}})d\mathbf{r} < 0 \tag{11.38}$$

This follows directly from the fact that:

$$(x+y)^{\frac{4}{3}} > x^{\frac{4}{3}} + y^{\frac{4}{3}} \tag{11.39}$$

if $x, y > 0$. For our model density we have $\Delta E_{x,\Sigma} = -1.70$ eV. In order to estimate the bond energy contribution ΔE_λ we use the scaling property of the exchange functional [22].

$$E_x[\rho_\lambda(\mathbf{r}, \vec{R})] = \lambda E_x[\rho(\mathbf{r}, \lambda\vec{R})] \tag{11.40}$$

where $\vec{R}$ is the relative position vector of the atomic nuclei. An expansion of equation 11.40 in powers of $\lambda - 1$ yields:

$$\Delta E_{x,\lambda} = (\lambda - 1)(E_x[\rho_\Sigma] + \vec{R} \cdot \int v_x([\rho_\Sigma]; \mathbf{r})\nabla_{\vec{R}}\rho_\Sigma(\mathbf{r})d\mathbf{r}) + O((\lambda - 1)^2) \tag{11.41}$$

If we neglect the second term (assuming that $E_x[\rho_\lambda(\mathbf{r}, \vec{R})]$ varies much faster with respect to λ than with respect to $\vec{R}$) and we obtain:

$$\Delta E_{x,\lambda} \approx (\lambda - 1)E_x[\rho_\Sigma] < 0 \tag{11.42}$$

The optimal λ with minimizes the LDA+GGA functional is given by $\lambda = 1.140$ and therefore we obtain to a good approximation:

$$\Delta E_{x,\lambda} = 0.140E_x[\rho_\Sigma] \tag{11.43}$$

Using the above formula we find $\Delta E_{x,\lambda} = -1.86$ eV. An actual calculation of $E_x[\rho_\lambda]$ gives $\Delta E_{x,\lambda} = -1.60$ eV. Finally there is the correction by going from spin unrestricted to spin restricted atoms. As can be seen from formula 11.27 this lowers the LDA exchange energy by a factor $2^{\frac{1}{3}}$. We therefore obtain the following positive correction to the bond energy:

$$\Delta E_{s,x} = -(2^{\frac{1}{3}} - 1)(E_x[\rho_A] + E_x[\rho_B]) = 2(2^{\frac{1}{3}} - 1)(\frac{3}{4})^4(\frac{3}{\pi^2})^{\frac{1}{3}} = 3.01 \text{ eV} \tag{11.44}$$

In general we can say that both formation of ρ_Σ and contraction of the density give a negative contribution to the bond energy contribution of the LDA exchange energy and that the spin unrestricted corrections give a positive contribution.

GGA exchange. We consider in particular the Becke exchange correction. However our conclusions remain valid for a broad class of nonlocal exchange corrections of the Becke form. The Becke exchange energy correction is given by:

$$E_x[\rho_\uparrow, \rho_\downarrow] = \sum_\sigma \int \rho_\sigma(\mathbf{r})^{\frac{4}{3}}f(x_\sigma)d\mathbf{r} \tag{11.45}$$

where

$$f(x) = -\beta\frac{x^2}{1 + 6\beta x \sinh^{-1} x} \tag{11.46}$$

with $\beta = 0.0042$ and

$$x_\sigma(\mathbf{r}) = \frac{|\nabla\rho_\sigma(\mathbf{r})|}{\rho_\sigma(\mathbf{r})^{\frac{4}{3}}} \tag{11.47}$$

Spin-unrestricted we have:

$$E_x[\rho] = \frac{1}{2^{\frac{1}{3}}} \int \rho(\mathbf{r})^{\frac{4}{3}}f(2^{\frac{1}{3}}x)d\mathbf{r} \tag{11.48}$$

The bond energy correction $\Delta E_{x,\Sigma}$ due to the Becke exchange is given by:

$$\Delta E_{x,\Sigma} = \frac{1}{2^{\frac{1}{3}}} \int (\rho_\Sigma(\mathbf{r})^{\frac{4}{3}} f(2^{\frac{1}{3}} x_\Sigma) - \rho_A(\mathbf{r})^{\frac{4}{3}} f(2^{\frac{1}{3}} x_A) - \rho_B(\mathbf{r})^{\frac{4}{3}} f(2^{\frac{1}{3}} x_B)) d\mathbf{r} \tag{11.49}$$

The main contribution to the integrand of the above integral arizes in the bond midpoint region of the molecule where the molecular gradients are small (in the bond midpoint they are exactly zero). In that case it is easily seen that we have:

$$\Delta E_{x,\Sigma} > 0 \tag{11.50}$$

Within our model we find $\Delta E_{x,\Sigma} = 0.38$ eV. If we now contract the density then we have consider- able changes in the density at the atomic nuclei. As the Becke exchange functional scales according to equation 11.40 we can use the same arguments as for the LDA exchange and use equation 11.42 in order to calculate $\Delta E_{x,\lambda}$. Both an exact calculation and the approximate formula 11.42 give $\Delta E_{x,\lambda} = -0.30$ eV.
The atomic spin unrestricted correction to the Becke exchange is given by:

$$\Delta E_{x,s} = 2 \int \rho_A(\mathbf{r})^{\frac{4}{3}} (f(x_A) - \frac{1}{2^{\frac{1}{3}}} f(2^{\frac{1}{3}} x_A)) d\mathbf{r} < 0 \tag{11.51}$$

We find $\Delta E_{x,s} = -0.24$ eV. So we can say in general that both formation of the density ρ_Σ and the spin unrestricted corrections give a positive contribution to the bond energy. The density con- traction leads to a negative contribution.

LDA correlation. Due to the rather complicated nature of the Vosko-Wilk-Nusair expression for the LDA correlation energy it is difficult to predict the change in this expression due to the described density variations. We find by numerical evaluation that $\Delta E_{c,\Sigma} = -0.17$ eV, $\Delta E_{c,\lambda} = -0.11$ eV and $\Delta E_{c,s} = -1.05$ eV.

GGA correlation. We consider the Perdew correlation correction which is given by:

$$E_c[\rho_\uparrow, \rho_\downarrow] = \int \frac{1}{d(\zeta)} C(\rho) g(\rho, |\nabla\rho|) d\mathbf{r} \tag{11.52}$$

where

$$g(\rho, |\nabla\rho|) = \rho^{\frac{4}{3}} x^2 e^{-\Phi(\rho,|\nabla\rho|)} \tag{11.53}$$

with

$$x = \frac{|\nabla\rho|}{\rho^{\frac{4}{3}}} \tag{11.54}$$

where $\zeta = (\rho_\uparrow - \rho_\downarrow)/\rho$ is the spin polarization and:

$$d(\zeta) = 2^{\frac{1}{3}} ((\frac{1+\zeta}{2})^{\frac{5}{3}} + (\frac{1-\zeta}{2})^{\frac{5}{3}})^{\frac{1}{2}} \tag{11.55}$$

The function $C(\rho)$ is a local function of the density and has been obtained from the gradient expansion for slowly varying densities of the correlation energy of the electron gas. The function Φ is given by:

$$\Phi(\rho, |\nabla\rho|) = 1.745 \tilde{f} \frac{C(\infty)}{C(\rho)} \frac{|\nabla\rho|}{\rho^{\frac{7}{6}}} \tag{11.56}$$

where $\tilde{f} = 0.11$. This function is obtained from the wavevector analysis of Langreth and Mehl. For the spin unpolarized case we have (d=1) :

$$E_c[\rho] = \int C(\rho)\rho^{\frac{4}{3}}x^2 e^{-\Phi(\rho,|\nabla\rho|)} d\mathbf{r} \tag{11.57}$$

We then have:

$$\Delta E_{c,\Sigma} = \int \left(C(\rho_\Sigma)g(\rho_\Sigma,|\nabla\rho_\Sigma|) - C(\rho_A)g(\rho_A,|\nabla\rho_A|) - C(\rho_B)g(\rho_B,|\nabla\rho_B|)\right)d\mathbf{r} \tag{11.58}$$

As with the Becke exchange correction the main contribution to the above integral arises from the molecular bond midpoint region where the density gradients of the molecule are small (the function g is exactly zero in the bond midpoint region) so we find:

$$\Delta E_{c,\Sigma} < 0 \tag{11.59}$$

Within our model we find $\Delta E_{c,\Sigma} = -0.17$ eV. The Perdew correlation functional does not scale like the exchange functional. However due to the fact that the function C is almost constant as a function of the density and the importance of the term $\rho^{\frac{4}{3}}x^2$ it is to a good approximation true (as we checked for several values of λ) that (at least for atomic and molecular densities):

$$E_c[\rho_\lambda(\mathbf{r},\vec{R})] \approx \lambda E_c[\rho(\mathbf{r},\lambda\vec{R})] \tag{11.60}$$

which yields just as for the exchange case:

$$\Delta E_{c,\lambda} \approx (\lambda - 1)E_c[\rho_\Sigma] > 0 \tag{11.61}$$

A calculation using the above formula (using the optimal value $\lambda = 1.140$) yields $\Delta E_{c,\lambda} = 0.16$ eV. An actual calculation gives $\Delta E_{c,\lambda} = 0.13$ eV.

We finally consider the atomic spin corrections. For spin polarized hydrogen atom we have an spin polarisation $\zeta = 1$ and $d(\zeta) = 2^{\frac{1}{3}}$ which means that the Perdew correlation energy for the spin unpolarized hydrogen atom is a factor $2^{\frac{1}{3}}$ times this energy for the spin polarized atom (assuming that the spatial densities are the same). So we find:

$$\Delta E_{c,s} = 2(2^{\frac{1}{3}} - 1)E_c[\rho_A] > 0 \tag{11.62}$$

Within our model we find $\Delta E_{c,s} = 0.28$ eV. So for the Perdew correction the formation of ρ_Σ gives a negative contribution to the bond energy and both the density contraction and spin polarization corrections give a positive contribution to the bond energy.

As can be seen from the above analysis the actual sign of the bond energy contribution of the several exchange and correlation energies to the bond energy depends on some opposing effects. However as a rule the Becke exchange correction is positive and lowers the bond energy and the Perdew correlation correction is negative and increases the bond energy. This is for instance true for all first and second row homonuclear diatomic molecules. A notable exception being the hydrogen molecule which is the subject of this paper. However the hydrogen molecule is a rather special case due the absence of core electrons so that density changes on the nuclei , which affect ΔE_λ are rather important. Also the spin corrections for the atoms are rather large because the spin polarized hydrogen atoms are an extreme case of spin polarization with $\zeta = 1$. Those corrections are smaller for atoms with a less extreme spin polarisation. This means that in applications to other molecules the most important correction to the bond energy is the term ΔE_Σ which is positive for the Becke exchange correction and negative for the Perdew correlation correction which confirms the actual trend observed.

11.6   Conclusions

In this work we analyzed the exact exchange and correlation energy E_x and E_c as well as the Kohn-Sham kinetic energy T_s as a function of the bond distance for the case of a dissociating hydrogen molecule. These quantities are compared with the correspnding ones from the local density approximation and nonlocal corrections. The LDA+NL gives a very good account for these quantities at equilibrium separation and gives a binding curve close to the exact one. At large separation the LDA and LDA+NL give a not so good description of the binding curve due to errors in the total atomic energies. These errors are caused by the fact the presently used approximate density functionals for the case of degenerate ground states are not invariant under unitary transformations within the ground state multiplet. Correct dissociation can however be obtained through the construction of approximate density density functionals which give a good total ground state energy for open shell atoms within a spin restricted formalism. The construction of such density functionals remains a future challenge. It is further shown that LDA plus nonlocal correlation is unable to describe the left-right correlation by underestimating this correlation energy. This error is partly corrected by a corresponding overestimation of the exchange energy.

The fact that the Becke-Perdew correction in the hydrogen molecule is negative instead of the the usual positive correction is shown to be related to special property of the hydrogen molecule of having no core electrons which make density contraction effects important. Another important fact in this respect is the fact the spin corrections for the atomic fragments in this molecule have a large effect due to the extreme polarization of the polarized atoms.

Chapter 12

Energy expressions in density functional theory using line integrals

In this paper we will adress the question of how to obtain energies from functionals when only the functional derivative is given. It is shown that one can obtain explicit expressions for the exchange-correlation energy from approximate exchange-correlation potentials using line integrals along paths within the space of densities. The path dependence of the results is discussed and criteria for path independence are given. Derivations are given of upper and lower bounds to the exchange-correlation energy in terms of the exchange-correlation potential at the beginning and the end point of a certain path. We further express the kinetic part T_{xc} of the exchange-correlation energy in terms of a line integral and derive a constraint on approximate correlation potentials. We show how to use the line-integral formalism to derive the requirements which exchange-correlation potentials must fulfil in order to make the exchange-correlation functional satisfy some symmetry property such as rotational and translational invariance and scaling properties. Finally we will discuss how to obtain bonding energies of molecules from approximate potentials.

12.1   Introduction

Density functional theory [5] has in the last few years been succesfully applied with increasing accuracy to systems ranging from atoms and molecules [71, 70, 122] to surfaces and solids [47]. Especially the introduction of the so-called generalized gradient expansion approximations (GGA) [44, 50, 46, 43, 45] for the exchange-correlation energy has constituted a great improvement over the local density approximation (LDA) in the calculation of for instance molecular bond energies. In view of the accuracy of the GGA *energy* expressions for the exchange-correlation functional and their improvement of the LDA energy expressions it is surprising that the GGA exchange-correlation *potential* gives little improvement over the LDA exchange-correlation potential [49, 32] and has a small influence on the density. For this reason calculating the GGA corrections from an LDA density yields almost the same results as an inclusion of the GGA potentials in a self-consistent calculation [139]. This fact is also reflected in the LDA+GGA eigenvalues which are very close to the LDA eigenvalues which are much too small for finite systems such as atoms and molecules and also for surfaces. As the eigenvalue of the highest occupied Kohn-Sham orbital has been proven to be equal to the ionisation energy of the system [31, 72] the LDA gives too weakly bound electrons. This latter fact is due to the LDA exchange-correlation potential having exponential decay instead of the correct Coulombic decay. One way to improve the one-electron energies and the density (and

related quantities such as dipole moments) is to directly approximate the exchange or exchange-correlation potential instead of the exchange or exchange-correlation functional. Over the years several potentials (not only within density functional theory) have been proposed. Well-known is the average Hartree-Fock exchange potential proposed by Slater [101, 148]. A so-called optimized effective potential (OEP) and an approximation to this was suggested by Sharp and Horton [34] and by Talman and Shadwick [33]. This potential received considerable attention within the context of exchange-only density functional theory or the optimized effective potential model (OPM) and was shown to give one-electron energies close to the Hartree-Fock ones. An approximation to this potential was recently proposed by Krieger et al. [35] and Gritsenko et al. [121]. Also within the weighted density approximation (WDA) method approximate exchange potentials with the correct asymptotic behaviour have been used [149, 150, 124]. A more crude way to improve the asymptotic behaviour of the exchange potential was given by Latter [151] within the context of the Thomas-Fermi model which due to its simplicity has found widespread use in atomic structure calculations. Much less work has been done regarding the full exchange-correlation potential. An approximate exchange-correlation potential has been proposed by Harbola and Sahni [131] . This exchange-correlation potential $v_{xc}(\mathbf{r})$ has been defined as the work done by bringing an electron from infinity to point $\mathbf{r}$ against the force of the electric field of the exchange-correlation hole. However to insure path-independence for non-spherical systems like molecules one must only consider the irrotational part of this electric field [152]. A gradient approximation to v_{xc} has recently been proposed in reference [32] which was shown to give very good ionisation energies as calculated from the highest occupied Kohn-Sham orbital.

In view of the fact that correct one-electron energies, the asymptotic Coulombic behaviour and the atomic shell structure can be reasonably obtained by approximate potentials it is an interesting question of theoretical and practical importance whether we can calculate good exchange and correlation energies from these approximate potentials. This is also of importance since it is possible to test the quality of approximate exchange-correlation potentials by comparing them with accurate ones constructed from accurate densities. It is for instance an interesting question how certain features in the exchange-correlation potential (such as the bond midpoint peaks in molecules [86, 90] related to the left-right correlation effect) contribute to the exchange-correlation energy. This requires an understanding of the relation between potential and energy expression. A further understanding of the features displayed in the Kohn-Sham potential might then also lead to more improved expressions for the exchange-correlation energy. If we want to assign some energy expression for instance to the model potentials considered above we immediately run into some theoretical difficulties because none of the potentials considered, with the exception of the OPM potential, is a functional derivative of some energy density functional. However for approximate exchange potentials which satisfy the exchange scaling property it is still possible to assign an exchange energy using the Levy-Perdew relation [121, 35]. However such a relation is not available for the exchange-correlation potential due to the unknown scaling properties of the correlation functional. In section 12.2 of this article we will show how to use line integrals to calculate the exchange-correlation energy from an exchange-correlation potential. We also discuss some criteria of path independence.In section 12.3 we discuss some bounds on the exchange-correlation energy in terms of the exchange-correlation potential and we give an line integral expression of the kinetic part T_{xc} of the exchange-correlation energy. In section 12.4 we derive some constraints on the correlation potential. In section 12.5 we derive some properties with respect to rotation, translation and scaling of the exchange-correlation energy from given properties with respect to these symmetries of the exchange-correlation potential. In section 12.6 we will adress the question how to calculate molecular bond energies from given potentials. Finally in section 12.7 we will present

our conclusions.

12.2 Line integrals and path dependence

In this section we will discuss the problem of obtaining energies from approximate potentials. The exchange-correlation energy for a system with Coulombic two-particle interactions is within density functional theory usually defined as:

$$E_{xc}[\rho] = F[\rho] - T_s[\rho] - \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{12.1}$$

where the universal functional F is as usual defined by constrained search [18, 128] over all anti-symmetric wavefunctions yielding density ρ :

$$F[\rho] = \min_{\psi \to \rho} \langle \psi | \hat{T} + \hat{W} | \psi \rangle \tag{12.2}$$

where $\hat{T}$ is the kinetic energy operator and $\hat{W}$ is the interparticle interaction operator. The Kohn-Sham kinetic energy $T_s[\rho]$ is the kinetic energy of a noninteracting particle system with the same density as the interacting system defined by:

$$T_s[\rho] = \min_{\psi_s \to \rho} \langle \psi_s | \hat{T} | \psi_s \rangle = \langle \phi_s[\rho] | \hat{T} | \phi_s[\rho] \rangle \tag{12.3}$$

where we search over all Slater-determinants yielding density ρ and $\phi_s[\rho]$ is the Kohn-Sham determinant. The exchange energy functional is usually defined as [22]:

$$E_x[\rho] = \langle \phi_s[\rho] | \hat{W} | \phi_s[\rho] \rangle - \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{12.4}$$

and the correlation functional by:

$$E_c[\rho] = E_{xc}[\rho] - E_x[\rho] \tag{12.5}$$

Suppose we have a parametrisation $\gamma(t)$ of a set of electron densities starting at $\gamma(0) = \rho_1$ and ending in $\gamma(1) = \rho_2$. Suppose further that we have an exchange-correlation energy functional $E_{xc}[\rho]$. Then we have:

$$E_{xc}[\rho_2] - E_{xc}[\rho_1] = \int_0^1 dt \frac{dE_{xc}}{dt} =$$

$$= \int_0^1 dt \int d\mathbf{r} \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})}[\gamma(t)] \frac{d\gamma(t)}{dt} = \int_0^1 dt \int d\mathbf{r} v_{xc}([\gamma(t)]; \mathbf{r}) \frac{d\gamma(t)}{dt} \tag{12.6}$$

in which v_{xc} is the functional derivative of E_{xc} :

$$v_{xc}([\rho]; \mathbf{r}) = \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})} \tag{12.7}$$

If we take for instance the straight path from zero to ρ defined by $\gamma(t) = t\rho$ and we have $E_{xc}[0] = 0$ (which must be the case for any physically acceptable exchange-correlation functional) then $\frac{d\gamma}{dt} = \rho$ and it follows that:

$$E_{xc}[\rho] = \int d\mathbf{r} \rho(\mathbf{r}) \varepsilon_{xc}([\rho]; \mathbf{r}) \tag{12.8}$$

in which ε_{xc} is defined as:

$$\varepsilon_{xc}([\rho]; \mathbf{r}) \equiv \int_0^1 dt v_{xc}([t\rho]; \mathbf{r}) \tag{12.9}$$

This thus yields an explicit expression of the exchange-correlation energy in terms of the exchange-correlation potential. It should be remembered however that in this expression v_{xc} is a functional derivative. Suppose however that we only have an expression of v_{xc} for instance in terms of the density and its gradients. Then we can define the line integral of v_{xc} along a path γ in the space of densities as:

$$\int_\gamma v_{xc} \equiv \int_0^1 dt \int d\mathbf{r} v_{xc}([\gamma(t)]; \mathbf{r}) \frac{d\gamma(t)}{dt} \tag{12.10}$$

As we can see from the above formulas the outcome of this line integral does not depend on the chosen path (except for the endpoints of course) whenever v_{xc} is the functional derivative of some functional E_{xc}. If one uses approximate potentials to obtain energy expressions this way one has to make sure that the result does not depend on the path γ. One therefore needs some criteria to determine whether v_{xc} yields path independent line integrals. This is for instance the case when:

$$\frac{\delta v_{xc}(\mathbf{r})}{\delta\rho(\mathbf{r}')} - \frac{\delta v_{xc}(\mathbf{r}')}{\delta\rho(\mathbf{r})} = 0 \tag{12.11}$$

which is a vanishing curl condition which is equivalent to:

$$\frac{\delta^2 E_{xc}}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} = \frac{\delta^2 E_{xc}}{\delta\rho(\mathbf{r}')\delta\rho(\mathbf{r})} \tag{12.12}$$

To prove this we only have to show that a v_{xc} satisfying the above relation is the functional derivative of some functional E_{xc}. Suppose we *define* the functional E_{xc} by integrating v_{xc} along the straight path $\gamma(t) = t\rho$:

$$E_{xc}[\rho] \equiv \int d\mathbf{r}\rho(\mathbf{r})\varepsilon_{xc}([\rho]; \mathbf{r}) \tag{12.13}$$

in which ε_{xc} is defined as in equation 12.9. If v_{xc} satisfies condition 12.11 then v_{xc} is the functional derivative of the above defined functional E_{xc} and hence path independent. This is readily shown by differentiation of the above equation 12.13. This yields:

$$\frac{\delta E_{xc}}{\delta\rho(\mathbf{r})} = \varepsilon_{xc}([\rho]; \mathbf{r}) + \int d\mathbf{r}'\rho(\mathbf{r}')\frac{\delta\varepsilon_{xc}([\rho]; \mathbf{r}')}{\delta\rho(\mathbf{r})} =$$

$$= \varepsilon_{xc}([\rho]; \mathbf{r}) + \int_0^1 dt \int d\mathbf{r}'\rho(\mathbf{r}')\frac{\delta v_{xc}([t\rho]; \mathbf{r}')}{\delta\rho(\mathbf{r})} \tag{12.14}$$

Now using condition 12.11 yields:

$$\frac{\delta E_{xc}}{\delta\rho(\mathbf{r})} = \varepsilon_{xc}([\rho]; \mathbf{r}) + \int_0^1 dt \int d\mathbf{r}'t\rho(\mathbf{r}')\frac{\delta v_{xc}([t\rho]; \mathbf{r})}{\delta(t\rho(\mathbf{r}'))} =$$

$$= \varepsilon_{xc}([\rho]; \mathbf{r}) + \int_0^1 dt\, t\frac{dv_{xc}([t\rho]; \mathbf{r})}{dt} = \varepsilon_{xc}([\rho]; \mathbf{r}) +$$

$$\int_0^1 dt\frac{d}{dt}(tv_{xc}([t\rho]; \mathbf{r})) - \int_0^1 dt v_{xc}([t\rho]; \mathbf{r}) = v_{xc}([\rho]; \mathbf{r}) \tag{12.15}$$

Hence v_{xc} is the functional derivative of E_{xc} and therefore path independent.

We will now give another criterium for path independence related to equation 12.11. In order to do this we recall Stokes theorem in vector calculus which states for a vectorfield $\vec{v}$ in n-dimensional space:

$$\sum_i^n \oint_\gamma v_i dx_i = \sum_{i,k}^n \int_O (\frac{\partial v_i}{\partial x_k} - \frac{\partial v_k}{\partial x_i}) dx_i dx_k \tag{12.16}$$

in which γ is a closed curve which is the boundary of surface O. A generalisation of this theorem to function spaces would be:

$$\oint_\gamma v\delta\rho(\mathbf{r})d\mathbf{r} = \int_O (\frac{\delta v(\mathbf{r})}{\delta\rho(\mathbf{r}')} - \frac{\delta v(\mathbf{r}')}{\delta\rho(\mathbf{r})})\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{12.17}$$

in which in the first integral the variations are restricted to be on the path γ and in the second integral on the surface O. From this equation we can see that if relation 12.11 is satisfied then the integral of v along any closed path is zero and hence the line integral of v is path independent. On the other hand if the line integral is zero for any closed path then we obtain the integrability condition 12.11.

12.3 Exchange-correlation energy and the kinetic part: bounds from potentials

The present formalism can be used to obtain energy expressions from approximate potentials from line integrals. In order to obtain sensible results from such a calculation one has to make sure that any approximate v_{xc} is either a functional derivative (for instance by requiring the integrability condition 12.11) or by requiring that the approximate potential is a good approximation to the exact potential for every density along the integration path. This favors in practical applications some paths over other paths. For instance the path defined by $\gamma(t) = t\rho$ has the disadvantage of not conserving the particle number which can therefore be fractional which gives theoretical problems if one wants to assign a potential to the corresponding density. Another more appealing choice of path is the following:

$$\gamma(t) = t^3\rho(t\mathbf{r}) \tag{12.18}$$

If we let the path parameter t run from 0 to 1 then we are integrating from 0 to ρ. The most important feature of this path is that it is particle number conserving, thus:

$$\int \gamma(t)d\mathbf{r} = \int \rho(\mathbf{r})d\mathbf{r} = N \tag{12.19}$$

along the path. For hydrogen-like atoms for instance we have:

$$\gamma(t) = \frac{(tZ)^3}{\pi}e^{-tZr} \tag{12.20}$$

in which Z is the nuclear charge. So if we let t approach zero then the density $\gamma(t)$ becomes increasingly diffuser and approaches zero in every point of space but keeping its normalization. Our main task is now to construct for the case of many electron systems exchange-correlation

potentials which approximate the exact exchange-correlation potential along this path. Application of equation 12.6 for this path leads to the following expression for the exchange-correlation energy:

$$E_{xc}[\rho] = \int_0^1 dt \int d\mathbf{r} v_{xc}([\gamma(t)]; \mathbf{r})(3t^2 \rho(t\mathbf{r}) + t^3 \mathbf{r} \cdot \nabla_{t\mathbf{r}} \rho(t\mathbf{r})) =$$

$$= \int_0^1 dt \int d\mathbf{r} \frac{1}{t} v_{xc}([\gamma(t)]; \frac{1}{t}\mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \tag{12.21}$$

in which we performed the substitution $t\mathbf{r} \to \mathbf{r}$. Before we go on to discuss the exchange-correlation potential let us first discuss the simpler case of the exchange-potential alone. Suppose we have an approximate exchange potential $\tilde{v}_x$ for instance of the form:

$$\tilde{v}_x(\mathbf{r}) = \rho^{\frac{1}{3}}(\mathbf{r}) f(x, y) \tag{12.22}$$

where

$$x(\mathbf{r}) = \frac{|\nabla \rho(\mathbf{r})|}{\rho^{\frac{4}{3}}(\mathbf{r})} \tag{12.23}$$

and

$$y(\mathbf{r}) = \frac{\nabla^2 \rho(\mathbf{r})}{\rho^{\frac{5}{3}}(\mathbf{r})} \tag{12.24}$$

are dimensionless functions. This approximate potential satisfies the following well-known scaling property [23] of the exchange potential:

$$v_x([\rho_t]; \mathbf{r}) = t v_x([\rho]; t\mathbf{r}) \tag{12.25}$$

in which $\rho_t = \gamma(t)$ is just the path of equation 12.18 (in the following we will for notational convenience use ρ_t instead of $\gamma(t)$ for this particular path). Let v_x denote the exact exchange potential being the functional derivative of the exact exchange functional. The difference between the exact exchange energy $E_x[\rho]$ and the approximate exchange energy $\tilde{E}_x[\rho]$ using the line integral of equation 12.21 and scaling property 12.25 is then given by:

$$E_x[\rho] - \tilde{E}_x[\rho] = \int_\gamma v_x - \int_\gamma \tilde{v}_x = \int d\mathbf{r} (v_x([\rho]; \mathbf{r}) - \tilde{v}_x([\rho]; \mathbf{r}))(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \tag{12.26}$$

For a correctly scaling approximate exchange potential the difference in the two line integrals thus only depends on the difference between the two potentials at the endpoints of the path and can be made arbitrarily small by better fits of $\tilde{v}_x$ to v_x at this endpoint. This can in practice be done for instance by fitting to the so-called Optimised Effective Potential (OPM) [35, 121].

To show that if $\tilde{v}_x$ is not a functional derivative we can make the difference in the line integral of v_x and $\tilde{v}_x$ arbitrarily large we consider the following path;

$$\tilde{\gamma}(t) = t^3 \rho(t\mathbf{r} + (1-t)\vec{R}) \tag{12.27}$$

In which $\vec{R}$ is an arbitrary vector in three-dimensional space. We have:

$$\frac{d\tilde{\gamma}(t)}{dt} = 3t^2 \rho(t\mathbf{r}+(1-t)\vec{R}) + t^3 \mathbf{r} \cdot \nabla_{t\mathbf{r}+(1-t)\vec{R}} \rho(t\mathbf{r}+(1-t)\vec{R}) - t^3 \vec{R} \cdot \nabla_{t\mathbf{r}+(1-t)\vec{R}} \rho(t\mathbf{r}+(1-t)\vec{R}) \tag{12.28}$$

Integrating along this path (which can easily be checked to be particle number conserving) yields using the fact that both the exact v_x and the approximate potential $\tilde{v}_x$ satisfy some translational invariance property (see equation 12.67 of section 12.5) we find:

$$\int_{\tilde{\gamma}} v_x - \int_{\tilde{\gamma}} \tilde{v}_x = \int d\mathbf{r}(v_x(\mathbf{r}) - \tilde{v}_x(\mathbf{r}))(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) - \vec{R} \cdot \int d\mathbf{r}(v_x(\mathbf{r}) - \tilde{v}_x(\mathbf{r}))\nabla\rho(\mathbf{r}) \quad (12.29)$$

We see that this expression can be made arbitrarily large by choosing an arbitrarily large $\vec{R}$ (which amounts to choosing a different path) if the difference between v_x and $\tilde{v}_x$ is not zero.

Things are more complicated if we want to calculate the total exchange-correlation energy due to the fact that we do not know the scaling properties of the correlation potential. This means that we cannot calculate the exchange-correlation energy directly from knowledge of $v_{xc}([\rho]; \mathbf{r})$ alone, we must know v_{xc} along some path $\gamma(t)$ in density space. However some useful inequalities can be derived from the knowledge of $v_{xc}([\rho]; \mathbf{r})$. Averill and Painter [30] and Levy and Perdew [22] have derived the following useful relation:

$$E_{xc}[\rho] = \int d\mathbf{r}v_{xc}([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) - T_{xc}[\rho] \tag{12.30}$$

where

$$T_{xc}[\rho] = T[\rho] - T_s[\rho] \tag{12.31}$$

is the kinetic part of the exchange-correlation energy. The kinetic energy $T[\rho]$ is as usual defined by:

$$T[\rho] = \langle \psi[\rho]|\hat{T}|\psi[\rho]\rangle \tag{12.32}$$

where $\hat{T}$ is the kinetic energy operator and $\psi[\rho]$ is the anti-symmetric wavefunction yielding density ρ and minimising the universal functional $F[\rho]$. As can be inferred directly from equation 12.6 using the scaling path 12.18 it follows from equation 12.30 that equivalently:

$$T_{xc}[\rho] = \frac{dE_{xc}[\gamma(t)]}{dt}|_{t=1} - E_{xc}[\rho] = \frac{dE_c[\gamma(t)]}{dt}|_{t=1} - E_c[\rho] \tag{12.33}$$

which is a useful relation that can be used to obtain the kinetic energy part of the exchange-correlation energy from approximate exchange-correlation or correlation functionals such as used in LDA and GGA. Relation 12.30 is also useful to obtain an upper bound to the exchange-correlation energy from knowledge of the exchange-correlation potential. From the definitions for $T[\rho]$ and $T_s[\rho]$ we can deduce the well-known result [5]:

$$T_s[\rho] = \langle \phi_s[\rho]|\hat{T}|\phi_s[\rho]\rangle \leq \langle \psi[\rho]|\hat{T}|\psi[\rho]\rangle = T[\rho] \tag{12.34}$$

and thus $T_{xc}[\rho] \geq 0$ and we find the following inequality:

$$E_{xc}[\rho] \leq \int d\mathbf{r}v_{xc}([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{12.35}$$

which provides an upper bound to E_{xc}. A similar inequality can be obtained for the correlation functional E_c (with corresponding correlation potential v_c being the functional derivative of E_c) by substracting the Levy-Perdew relation 12.26 for the exchange from equation 12.30 with yields:

$$E_c[\rho] = \int d\mathbf{r}v_c([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) - T_{xc}[\rho] \leq \int d\mathbf{r}v_c([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{12.36}$$

This gives an upper bound to the correlation energy when the correlation potential is known.

To provide a lower bound to E_c in terms of potentials we use the following relation valid for systems with Coulombic interparticle interactions proved by Levy [27]:

$$\lim_{t \to 0} \frac{1}{t} E_c[\rho_t] = -b[\rho] \tag{12.37}$$

where $b[\rho]$ is a positive functional satisfying the following inequality:

$$T_{xc}[\rho] - E_c[\rho] \le b[\rho] \le \langle \phi_s[\rho] | \hat{W} | \phi_s[\rho] \rangle \tag{12.38}$$

where $\phi_s[\rho]$ is the Kohn-Sham determinant. Relation 12.37 immediately leads to:

$$-b[\rho] = \lim_{t \to 0} \frac{1}{t} E_c[\rho_t] = \frac{dE_c[\rho_t]}{dt}\Big|_{t=0} = \int d\mathbf{r}\, \hat{v}_c([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \tag{12.39}$$

where

$$\hat{v}_c([\rho]; \mathbf{r}) = \lim_{\lambda \to 0} \frac{1}{\lambda} v_c([\rho_\lambda]; \frac{\mathbf{r}}{\lambda}) \tag{12.40}$$

Combining relations 12.38 and 12.33 then gives:

$$E_c[\rho] \ge \frac{1}{2}(\frac{dE_c[\rho_t]}{dt}\Big|_{t=0} + \frac{dE_c[\rho_t]}{dt}\Big|_{t=1}) = \int d\mathbf{r} \frac{1}{2}(\hat{v}_c([\rho]; \mathbf{r}) + v_c([\rho]; \mathbf{r}))(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \tag{12.41}$$

So we have now derived an upper and a lower bound to the correlation energy functional in terms of the correlation potential. Adding the Levy-Perdew relation for the exchange potential on both sides of the inequalities gives corresponding bounds for the exchange-correlation energy:

$$\int d\mathbf{r} \frac{1}{2}(\hat{v}_{xc}([\rho]; \mathbf{r}) + v_{xc}([\rho]; \mathbf{r}))(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \le E_{xc}[\rho]$$

$$\le \int d\mathbf{r}\, v_{xc}([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \tag{12.42}$$

where $\hat{v}_{xc} = v_x + \hat{v}_c$ (using $\hat{v}_x = v_x$ which is due to exchange scaling property of v_x). So upper and lower bounds to the exchange-correlation energy can be given from the knowledge of the exchange-correlation potential at the beginning ($t = 0$) and the end ($t = 1$) of the scaling path.

We first will now give an expression for the kinetic part of the exchange-correlation energy. From relation 12.30 and equation 12.6 we find:

$$T_{xc}[\rho] = \int d\mathbf{r}(v_{xc}([\rho]; \mathbf{r}) - \bar{v}_{xc}([\rho]; \mathbf{r}))(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla \rho(\mathbf{r})) \tag{12.43}$$

where

$$\bar{v}_{xc}([\rho]; \mathbf{r}) = \int_0^1 dt \frac{1}{t} v_{xc}([\gamma(t)]; \frac{\mathbf{r}}{t}) \tag{12.44}$$

This gives an explicit expression for T_{xc} in terms of the exchange-correlation potential. It can therefore directly be calculated from approximate expressions of v_{xc} or from the knowledge of v_{xc} along the scaling path. As a result of the scaling property of the exchange potential the above equation can be further split up as follows.

$$\bar{v}_{xc}([\rho]; \mathbf{r}) = \int_0^1 dt \frac{1}{t} v_x([\gamma(t)]; \frac{\mathbf{r}}{t}) + \int_0^1 dt \frac{1}{t} v_c([\gamma(t)]; \frac{\mathbf{r}}{t})$$

$$= v_x([\rho]; \mathbf{r}) + \int_0^1 dt \frac{1}{t} v_c([\gamma(t)]; \frac{\mathbf{r}}{t}) = v_x([\rho]; \mathbf{r}) + \bar{v}_c([\rho]; \mathbf{r}) \tag{12.45}$$

where

$$\bar{v}_c([\rho]; \mathbf{r}) = \int_0^1 dt \frac{1}{t} v_c([\gamma(t)]; \frac{\mathbf{r}}{t}) \tag{12.46}$$

This leads to another expression for T_{xc} in terms of the correlation potential:

$$T_{xc}[\rho] = \int d\mathbf{r} (v_c([\rho]; \mathbf{r}) - \bar{v}_c([\rho]; \mathbf{r}))(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{12.47}$$

From $T_{xc}[\rho] \geq 0$ it follows further that:

$$\int d\mathbf{r} \bar{v}_c([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \leq \int d\mathbf{r} v_c([\rho]; \mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{12.48}$$

which provides a further constraint on approximate correlation potentials. Further constraints are derived in the next section.

12.4 Constraints on v_c

We now will derive some constraints on the correlation potential. Levy has proven the following useful relation [27] which is valid for systems with repulsive Coulombic interparticle interactions:

$$\lim_{t \to \infty} E_c[\rho_t] > -\infty \tag{12.49}$$

Using relation 12.49 it then follows directly from:

$$E_{xc}[\rho_t] = tE_x[\rho] + E_c[\rho_t] \tag{12.50}$$

that

$$E_x[\rho] = \lim_{t \to \infty} \frac{1}{t} E_{xc}[\rho_t] \tag{12.51}$$

Equation 12.49 immediately leads to a constraint on the correlation potential. If we use the scaling path we find the following line integral:

$$E_c[\rho_\lambda] = \int_0^\lambda dt \int d\mathbf{r} \frac{1}{t} v_c([\rho_t]; \frac{\mathbf{r}}{t})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{12.52}$$

Application of equation 12.49 then immediately gives:

$$\int_0^\infty dt \int d\mathbf{r} \frac{1}{t} v_c([\rho_t]; \frac{\mathbf{r}}{t})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) > -\infty \tag{12.53}$$

which puts constraints on approximate correlation potentials. The above constraint is for example not satisfied by the LDA correlation potential due to the logarithmic divergence of the LDA correlation energy as function of the scaling parameter t. We can also write equation 12.51 in line integral form giving:

$$E_x[\rho] = \lim_{\lambda \to \infty} \frac{1}{\lambda} \int_0^\lambda dt \int d\mathbf{r} \frac{1}{t} v_{xc}([\rho_t]; \frac{\mathbf{r}}{t})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) \tag{12.54}$$

Inserting $v_{xc} = v_x + v_c$ in the above equation and using the scaling property of v_x we find the following constraint on v_c:

$$\lim_{\lambda \to \infty} \frac{1}{\lambda} \int_0^\lambda dt \int d\mathbf{r} \frac{1}{t} v_c([\rho_t]; \frac{\mathbf{r}}{t})(3\rho(\mathbf{r}) + \mathbf{r} \cdot \nabla\rho(\mathbf{r})) = 0 \tag{12.55}$$

The above constraint which is weaker than constraint 12.53 follows also directly from constraint 12.53. The correlation potential of the local density approximation for instance satisfies the above equation 12.55 but not constraint 12.53.

12.5 Invariance properties of potentials

The above formalism of line integrals provides an easy way to obtain conditions on the energy functional in the case that the exchange-correlation potential has some invariance or symmetry property. Suppose for instance that we vary the densities along our path γ by varying our path parameter t but that the potential v_{xc} has some symmetry property under such changes. Using equation 12.6 we then can deduce some properties of E_{xc}. In the following we will apply this idea to translation, rotation and scaling properties of E_{xc}.
We define a path $\gamma(t)$ by:

$$\gamma(t) = \rho(R(t)\mathbf{r}) \tag{12.56}$$

where $R(t)$ is a rotation in threedimensional space which rotates the vector $\mathbf{r}$ around a vector $\vec{\omega}$ by an angle t. If the functional E_{xc} is invariant under rotations we have that $E_{xc}[\gamma(t)] = E_{xc}[\rho]$ is constant as a function of t and we find:

$$0 = \frac{dE_{xc}[\gamma(t)]}{dt}|_{t=\theta} = \int d\mathbf{r}\frac{\delta E_{xc}}{\delta\rho(\mathbf{r})}[\gamma(\theta)]\frac{d\gamma(t)}{dt}|_{t=\theta} =$$

$$= \int d\mathbf{r}v_{xc}([\gamma(\theta)];\mathbf{r})(\vec{\omega}\times R(\theta)\mathbf{r})\cdot\nabla_{R(\theta)\mathbf{r}}\rho(R(\theta)\mathbf{r}) \tag{12.57}$$

For $\theta = 0$ in particular we find:

$$0 = \int d\mathbf{r}v_{xc}([\rho];\mathbf{r})(\vec{\omega}\times\mathbf{r})\cdot\nabla\rho(\mathbf{r}) =$$

$$= \int d\mathbf{r}v_{xc}([\rho];R(\theta)\mathbf{r})(\vec{\omega}\times R(\theta)\mathbf{r})\cdot\nabla_{R(\theta)\mathbf{r}}\rho(R(\theta)\mathbf{r}) \tag{12.58}$$

The above equations 12.57 and 12.58 yield the same result for all densities ρ and all rotations $R(\theta)$. We thus must have:

$$v_{xc}([\rho(R(\theta)\mathbf{r})];\mathbf{r}) = v_{xc}([\rho(\mathbf{r})];R(\theta)\mathbf{r}) \tag{12.59}$$

So if we insert in v_{xc} the rotated density then we obtain the same value in point $\mathbf{r}$ as v_{xc} with the original density in the rotated point $R(\theta)\mathbf{r}$. This result and some other results in this section have been presented elsewhere [23, 32] and are only presented here for clarity and to demonstrate the usefulness of the density path formalism in deriving properties of the potential and functional. The line integral of equation 12.6 offers the possibililty to make statements about the energy functional when we know properties of the potential. We can now prove the following statement. If a potential v_{xc} satisfies the rotation symmetry property of equation 12.59 and if v_{xc} is a functional derivative of some functional E_{xc} then E_{xc} is rotationally invariant. To prove this we insert the path of equation 12.56 in equation 12.6 which gives:

$$E_{xc}[\rho(R(\theta)\mathbf{r})] - E_{xc}[\rho(\mathbf{r})] = \int_0^\theta dt\int d\mathbf{r}v_{xc}([\gamma(t)];\mathbf{r})(\vec{\omega}\times R(t)\mathbf{r})\cdot\nabla_{R(t)\mathbf{r}}\rho(R(t)\mathbf{r}) =$$

$$= \int_0^\theta dt\int d\mathbf{r}v_{xc}([\rho];\mathbf{r})(\vec{\omega}\times\mathbf{r})\cdot\nabla\rho(\mathbf{r}) = \theta\vec{\omega}\cdot\int d\mathbf{r}v_{xc}([\rho];\mathbf{r})\mathbf{r}\times\nabla\rho(\mathbf{r}) \tag{12.60}$$

in which we in the second step used the rotation symmetry property of equation 12.59 and carried out a substitution $R(t)\mathbf{r} \to \mathbf{r}$. If we insert $\theta = 2\pi$ in the above formula 12.60 then we are integrating along a closed path and we obtain:

$$\oint_\gamma v_{xc} = 2\pi\vec{\omega} \cdot \int d\mathbf{r} \, v_{xc}([\rho]; \mathbf{r})\mathbf{r} \times \nabla\rho(\mathbf{r}) \tag{12.61}$$

As v_{xc} is a functional derivative the loop integral should be zero. This is also obvious from the left hand side of equation 12.60 as the 2π-rotated density is equal to the starting density and we obtain:

$$0 = \int d\mathbf{r} \, v_{xc}([\rho]; \mathbf{r})\mathbf{r} \times \nabla\rho(\mathbf{r}) \tag{12.62}$$

for any density ρ. If we insert this equation back into equation 12.60 then we obtain:

$$E_{xc}[\rho(R(\theta)\mathbf{r})] = E_{xc}[\rho(\mathbf{r})] \tag{12.63}$$

which proves our statement.

We will now carry out a similar derivation for the translation properties of the potential. If we define the path $\gamma(t)$ by:

$$\gamma(t) = \rho(\mathbf{r} + t\vec{R}) \tag{12.64}$$

in which $\vec{R}$ is an arbitrary translation vector. Now suppose that E_{xc} is translationally invariant then we have that $E_{xc}[\gamma(t)] = E_{xc}[\rho]$ is constant as a function of t and :

$$0 = \frac{dE_{xc}[\gamma(t)]}{dt}\Big|_{t=1} = \int d\mathbf{r} \frac{\delta E_{xc}}{\delta\rho(\mathbf{r})}[\gamma(1)]\frac{d\gamma(t)}{dt}\Big|_{t=1} =$$

$$= \int d\mathbf{r} \, v_{xc}([\gamma(1)]; \mathbf{r})\vec{R} \cdot \nabla\rho(\mathbf{r} + \vec{R}) \tag{12.65}$$

Similarly:

$$0 = \frac{dE_{xc}[\gamma(t)]}{dt}\Big|_{t=0} = \int d\mathbf{r} \, v_{xc}([\rho]; \mathbf{r})\vec{R} \cdot \nabla\rho(\mathbf{r})$$

$$= \int d\mathbf{r} \, v_{xc}([\rho]; \mathbf{r} + \vec{R})\vec{R} \cdot \nabla\rho(\mathbf{r} + \vec{R}) \tag{12.66}$$

As the above equations 12.65 and 12.66 yield the same result for all densities ρ and translation vectors $\vec{R}$ it follows that:

$$v_{xc}([\rho(\mathbf{r} + \vec{R})]; \mathbf{r}) = v_{xc}([\rho(\mathbf{r})]; \mathbf{r} + \vec{R}) \tag{12.67}$$

Thus v_{xc} with the translated density inserted yields the same value in point $\mathbf{r}$ as v_{xc} with the original density inserted in point $\mathbf{r} + \vec{R}$. This result has been derived elsewhere [32] and is only presented here for the sake of clarity. Using the line integral of equation 12.6 we now however also prove the opposite statement under some restrictions. If a potential satisfies relation 12.67 and if v_{xc} is a functional derivative of some functional E_{xc} which is bounded from above or below (The exact exchange-correlation functional is bounded from above as $E_{xc} \leq 0$) then this functional is translational invariant. If we use the path of equation 12.64 and insert it in equation 12.6 we obtain:

$$E_{xc}[\rho(\mathbf{r} + \vec{R})] - E_{xc}[\rho(\mathbf{r})] = \int_0^1 dt \int d\mathbf{r} \, v_{xc}([\gamma(t)]; \mathbf{r})\vec{R} \cdot \nabla\rho(\mathbf{r} + t\vec{R}) \tag{12.68}$$

If we use the translation symmetry property 12.67 we have:

$$E_{xc}[\rho(\mathbf{r}+\vec{R})] - E_{xc}[\rho(\mathbf{r})] = \int_0^1 dt \int d\mathbf{r} v_{xc}([\rho];\mathbf{r}+t\vec{R})\vec{R}\cdot\nabla\rho(\mathbf{r}+t\vec{R}) =$$

$$= \int_0^1 dt \int d\mathbf{r} v_{xc}([\rho];\mathbf{r})\vec{R}\cdot\nabla\rho(\mathbf{r}) = \vec{R}\cdot\int d\mathbf{r} v_{xc}([\rho];\mathbf{r})\nabla\rho(\mathbf{r}) \tag{12.69}$$

in which we carried out a substitution $\mathbf{r}+t\vec{R}\to\mathbf{r}$ which makes the t-integration trivial. The right hand side of this equation can be made arbitrarily large (both positive and negative) by making $\vec{R}$ arbitrarily large. As E_{xc} is bounded from above or below this righthand side of the equation must therefore be zero which yields:

$$\int d\mathbf{r} v_{xc}([\rho];\mathbf{r})\nabla\rho(\mathbf{r}) = 0 \tag{12.70}$$

and therefore one has:

$$E_{xc}[\rho(\mathbf{r}+\vec{R})] = E_{xc}[\rho(\mathbf{r})] \tag{12.71}$$

We can carry out a similar analysis for scaling properties. We then define a path:

$$\gamma(t) = t^3\rho(t\mathbf{r}) \tag{12.72}$$

Suppose we have an energy functional E with functional derivative v which satisfies the following scaling relation:

$$E[\gamma(t)] = t^n E[\rho] \tag{12.73}$$

with n some positive integer. If we differentiate the above equation in $t=1$ we obtain:

$$nE[\rho] = \frac{dE[\gamma(t)]}{dt}\Big|_{t=1} = \int d\mathbf{r}\frac{\delta E}{\delta\rho(\mathbf{r})}[\gamma(t)]\frac{d\gamma(t)}{dt}\Big|_{t=1} =$$

$$= \int d\mathbf{r} v([\rho];\mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r}\cdot\nabla\rho(\mathbf{r})) \tag{12.74}$$

Differentiating the same expression in $t=\lambda$ gives:

$$n\lambda^{n-1}E[\rho] = \frac{dE[\gamma(t)]}{dt}\Big|_{t=\lambda} = \int d\mathbf{r} v([\gamma(\lambda)];\mathbf{r})(3\lambda^2\rho(\lambda\mathbf{r}) + \lambda^3\mathbf{r}\cdot\nabla_{\lambda\mathbf{r}}\rho(\lambda\mathbf{r})) \tag{12.75}$$

This yields:

$$E[\rho] = \frac{1}{n}\int d\mathbf{r}\frac{1}{\lambda^n}v([\gamma(\lambda)];\mathbf{r})(3\lambda^3\rho(\lambda\mathbf{r}) + \lambda^4\mathbf{r}\cdot\nabla_{\lambda\mathbf{r}}\rho(\lambda\mathbf{r})) \tag{12.76}$$

On the other hand from equation 12.74 it follows:

$$E[\rho] = \frac{1}{n}\int d\mathbf{r} v([\rho];\mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r}\cdot\nabla\rho(\mathbf{r})) =$$

$$= \frac{1}{n}\int d\mathbf{r} v([\rho];\lambda\mathbf{r})(3\lambda^3\rho(\lambda\mathbf{r}) + \lambda^4\mathbf{r}\cdot\nabla_{\lambda\mathbf{r}}\rho(\lambda\mathbf{r})) \tag{12.77}$$

As the outcome of the integrals in equation 12.76 and 12.77 is equal for any density ρ and scaling parameter λ it follows:

$$v([\lambda^3\rho(\lambda\mathbf{r})];\mathbf{r}) = \lambda^n v([\rho];\lambda\mathbf{r}) \tag{12.78}$$

Thus result has been derived before [23] but is presented here to motivate our following statement. If a potential v satisfies the scaling relation 12.78 and is the functional derivative of some functional E with $E[0] = 0$ then this functional satisfies the scaling property 12.73. If we use the line integral of equation 12.6 and insert the path of equation 12.72 we obtain:

$$E[\gamma(\lambda)] - E[0] = \int_0^\lambda dt \int d\mathbf{r} v([\gamma(t)];\mathbf{r})(3t^2\rho(t\mathbf{r}) + t^3\mathbf{r}\cdot\nabla_{t\mathbf{r}}\rho(t\mathbf{r})) \tag{12.79}$$

If we now use the scaling property 12.78 we have after a substitution:

$$E[\gamma(\lambda)] - E[0] = \int_0^\lambda dt \int d\mathbf{r} t^{n-1} v([\rho];\mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r}\cdot\nabla\rho(\mathbf{r})) =$$

$$= \frac{\lambda^n}{n} \int d\mathbf{r} v([\rho];\mathbf{r})(3\rho(\mathbf{r}) + \mathbf{r}\cdot\nabla\rho(\mathbf{r})) \tag{12.80}$$

Hence we have:

$$E[\gamma(\lambda)] - E[0] = \lambda^n (E[\gamma(1)] - E[0]) \tag{12.81}$$

In this equation $E[0]$ is just an integration constant. For instance if we add to the functional 12.73 an arbitrary constant then it will satisfy the above equation 12.81 and its potential will still satsify 12.78. So if we set $E[0] = 0$ then we obtain:

$$E[\lambda^3\rho(\lambda\mathbf{r})] = \lambda^n E[\rho(\mathbf{r})] \tag{12.82}$$

which proves our statement.

12.6 Calculating molecular binding energies from potentials

One of the most succesful applications of density functional theory has been the calculation of molecular binding energies. This is largely due to the development of gradient corrected functionals. However the potentials corresponding to these functionals are not much improved [32] . An improvement of the existing gradient corrections might be sought in the development of better potentials as they can be compared to exact potentials obtainable from accurate densities. In this section we will discuss how to obtain binding energies from approximate potentials.

Suppose we have a diatomic molecule consisting of atom A and atom B. The total density of the molecule is given by $\rho_M = \rho_\Sigma + \Delta\rho$ where $\rho_\Sigma = \rho_A + \rho_B$ is the sum of atomic densities and $\Delta\rho$ is the deformation density of the molecule. The binding energy contribution of the exchange-correlation energy is then given by:

$$\Delta E_{xc} = E_{xc}[\rho_M] - E_{xc}[\rho_A] - E_{xc}[\rho_B] = (E_{xc}[\rho_M] - E_{xc}[\rho_\Sigma]) + (E_{xc}[\rho_\Sigma] - E_{xc}[\rho_A] - E_{xc}[\rho_B]) \tag{12.83}$$

The first term between brackets we will call ΔE_1 and the second term between brackets we will call ΔE_2. This division is useful in connection with the application of the transition-state method for bond energies developed by Ziegler and Rauk [153] to ΔE_1 and is helpful to obtain numerical

stability for Coulombic terms analogous to ΔE_2 in schemes that rely on numerical integration [154]. The first term ΔE_1 can in practice accurately be calculated using the linear path [154, 155]:

$$\gamma(t) = \rho_\Sigma + t(\rho_M - \rho_\Sigma) = \rho_\Sigma + t\Delta\rho \tag{12.84}$$

which yields:

$$\Delta E_1 = \int_0^1 dt \int d\mathbf{r} v_{xc}([\gamma(t)]; \mathbf{r})\Delta\rho(\mathbf{r}) = \int d\mathbf{r}\Delta\rho(\mathbf{r}) \int_0^1 dt v_{xc}([\gamma(t)]; \mathbf{r}) \tag{12.85}$$

This procedure is equivalent to the transition state method [153] as is evident when using the Simpson rule for the t-integration:

$$\Delta E_1 = \int d\mathbf{r}\Delta\rho(\mathbf{r})(\frac{1}{6}v_{xc}([\rho_\Sigma]; \mathbf{r}) + \frac{2}{3}v_{xc}([\rho_T]; \mathbf{r}) + \frac{1}{6}v_{xc}([\rho_M]; \mathbf{r})) \tag{12.86}$$

where $\rho_T = \gamma(\frac{1}{2}) = \frac{1}{2}(\rho_\Sigma + \rho_M)$ is the so-called transition state density. For practical applications of the above method with approximate exchange-correlation potentials it is necessary that the approximate potential gives a good description of the real exchange-correlation potential in the region where the deformation density $\Delta\rho$ is the largest. The other part of the binding energy ΔE_2 is easily obtained by numerical integration if explicit exchange-correlation energy expressions are known. However if no energy expressions are known we have to calculate also this term from the potential.One way to do this is is just to calculate the total energies of the atom and the molecule using equations 12.8 and 12.9 and to calculate the difference. This procedure has however some disadvantages. First of all along the path $\gamma(t) = t\rho$ the number of particles is not conserved which presents some difficulties from a theoretical point of view. Secondly if one makes an approximate expression for the potential v for some N-electron system, for instance by some fitting procedure to some known exact potential, one makes unknown errors for systems with a different number of electrons. It is therefore of some advantage to use particle number conserving paths. An other way of calculating the energy difference ΔE_2 is to choose the following path:

$$\gamma_\Sigma(t) = \rho_A(\mathbf{r}) + \rho_B(\mathbf{r} + t\vec{R}) \tag{12.87}$$

in which we let t run from 1 to ∞. This path is particle number conserving. If we have $\rho_\Sigma = \gamma(1)$ then:

$$\Delta E_2 = \int_{\gamma_\Sigma} v_{xc} = \int d\mathbf{r} \int_\infty^1 v_{xc}([\gamma_\Sigma(t)]; \mathbf{r})\vec{R} \cdot \nabla\rho_B(\mathbf{r} + t\vec{R}) \tag{12.88}$$

Application of the above formula puts some severe restrictions on approximate exchange-correlation potentials. This approximate potential should give a good description of the exact v_{xc} at all bond distances. For instance the bond midpoint peak in v_{xc} in dissociating molecules [86, 90] must also be described by this approximate v_{xc}. This might be feasible for instance in a gradient expansion using Laplacians of the density in any approximate v_{xc}. However care should be taken for potentials which are not functional derivatives as in that case the value of ΔE_2 will depend on the dissociation path taken.

12.7 Conclusions

In this work we adressed the question how to obtain exchange-correlation energies from approximate exchange-correlation potentials. This is of some theoretical and practical importance as many approximate exchange- and correlation potentials have been proposed. It is shown how one can

use line integrals to express the exchange-correlation energy in terms of the exchange-correlation potentials. It was further shown how to derive symmetry properties of the exchange-correlation energy functional from the corresponding properties of the exchange-correlation potential. We further obtained some upper and lower bounds to the exchange-correlation energy which can be calculated if the exchange-correlation potential is known in two limiting cases, at the beginning and the end of the scaling path. We showed how to express the kinetic part T_{xc} of the exchange-correlation energy in terms of line integrals over the exchange-correlation v_{xc} or correlation potential v_c and derived some constraints on the correlation potential. We finally adressed the problem of calculating the exchange-correlation part of molecular binding energies from approximate potentials. The constraints and inequalities derived in this article might prove useful to the development of more accurate exchange-correlation potentials improving the LDA and GGA potentials. Work along this line is in progress.

Samenvatting

In dit proefschrift wordt de dichtheidsfunktionaaltheorie (DFT) onderzocht en worden nieuwe benaderingen voorgesteld voor de Kohn-Sham potentiaal.

De dichtheidsfunktionaaltheorie, die in de jaren zestig is ontwikkeld door Hohenberg en Kohn en Kohn en Sham, geeft een methode om de grondtoestandseigenschappen van veeldeeltjessystemen zoals atomen, molekulen en vaste stoffen te berekenen. In de dichtheidsfunktionaaltheorie zijn de grondtoestandsverwachtingswaarden van de kwantummechanische operatoren die de fysische observabelen voorstellen dichtheidsfunktionalen, d.w.z. het zijn funkties die gedefinieerd zijn op de oneindigdimensionale ruimte van elektronendichtheden. Een belangrijke dichtheidsfunktionaal is de grondtoestandsenergiefunktionaal. Deze funktionaal wordt geminimaliseerd voor de grondtoestandselektronendichtheid van het systeem. Minimalisering van de energiefunktionaal leidt dan tot een Euler-Lagrange-vergelijking voor de elektronendichtheid. In de Kohn-Sham aanpak van dichtheidsfunktionaaltheorie krijgt deze variationele vergelijking de vorm van een onafhankelijk deeltjesmodel dat bekend staat als het Kohn-Sham systeem. Het wisselwerkende systeem van elektronen wordt dus vervangen door een niet-wisselwerkend systeem van Kohn-Sham elektronen die bewegen in het veld van een effektieve potentiaal. Deze Kohn-Sham potentiaal is een funktionaal van de elektronendichtheid. Tengevolge van de ingewikkelde wiskundige definitie van de Kohn-Sham potentiaal zijn hiervoor geen exakte analytische uitdrukkingen in termen van de dichtheid bekend. De onbekende term in de Kohn-Sham potentiaal is de exchange-korrelatie potentiaal, die de funktionele afgeleide is naar de elektronendichtheid van de exchange-korrelatie funktionaal. Voor praktische toepassingen worden er voor deze funktionaal benaderingen gebruikt. De historisch gezien meest gebruikte benadering is de lokale dichtheidsbenadering. Deze benadering, die gebaseerd is op berekeningen aan het elektronengas, blijkt een simpele en tegelijkertijd verrassend nauwkeurige benadering. De nauwkeurigheid van de lokale dichtheidsbenadering blijkt echter in belangrijke mate te worden vergroot door de zogenaamde gegeneraliseerde gradiëntbenaderingen, die naast de lokale elektronendichtheid ook afhangen van de gradiënt van de elektronendichtheid. Om deze laatste reden worden de gradiëntfunktionalen ook wel nietlokale funktionalen genoemd. De gradiëntfunktionalen blijken echter niet op alle punten de lokale dichtheidsbenadering te verbeteren. De verbetering geldt vooral voor energietermen, zoals de bindingsenergie van molekulen. De verbetering voor lokale grootheden zoals de exchange-korrelatie potentiaal of de elektronendichtheid is veel minder. Dit betekent bijvoorbeeld ook dat grootheden die direkt van de elektronendichtheid afhangen, zoals polariseerbaarheden, weinig verbetert worden. In dit proefschrift stellen we een aantal benaderingen voor voor de exchange- en exchange-korrelatie potentiaal die de potentialen van de gegenariseerde gradiëntbenaderingen aanzienlijk verbeteren, terwijl energietermen met gelijke nauwkeurigheid uitgerekend worden.

In hoofdstuk 5 van dit proefschrift bekijken we voor een aantal atomen de gradiëntbenaderingen van Becke en Perdew voor de exchange-korrelatie potentiaal en vergelijken deze met bijna exakte exchange-korrelatie potentialen berekend uit nauwkeurige elektronendichtheden. Om dit te

doen gebruiken we een nieuw ontwikkelde methode om de Kohn-Sham potentiaal te berekenen uit
een gegeven elektronendichtheid. Vervolgens wordt een nieuwe gradiëntbenadering voorgesteld die
in tegenstelling tot de bekende gradiëntbenaderingen het goede langeafstandsgedrag van de xc-
potentiaal weergeeft. De Kohn-Sham eigenwaarden en de ionisatie-energie (die gegeven wordt door
de energie van de hoogst bezette baan van het Kohn-Sham systeem) worden hierdoor aanzienlijk
verbeterd.

In hoofdstuk 6 worden verschillende nieuwe gradiëntbenaderingen voorgesteld voor de Kohn-Sham
exchange potentiaal. Dit zijn Padé-approximanten in termen van variabelen die dimensieloze
samenstellingen zijn van de dichtheid en de gradiënt en de laplaciaan van de dichtheid. Op deze
wijze wordt automatisch aan de goede schalingseigenschappen voldaan. De vorm van de Padé-
approximant wordt verder zo gekozen dat het langeafstandsgedrag van de potentiaal goed wordt
beschreven. De voorgestelde Padé-approximanten blijken de eigenwaarden en de vorm van de po-
tentiaal te verbeteren. De exchange energiën die worden berekend met behulp van de Levy-Perdew
relatie zijn echter minder goed. Deze energiën worden vooral bepaald door de vorm van de potenti-
aal in het kerngbied van het atoom. Een betere fit van de potentiaal in dit gebied verslechtert echter
het langeafstandsgedrag van de potentiaal hetgeen de te beperkte vorm van de Padé-approximant
aantoont.

In hoofdstuk 7 analyseren we verschillende delen van de Kohn-Sham potentiaal. De vraag hier-
bij is hoe de effekten van het Pauli-principe en de effekten van de Coulomb-afstoting tussen de
elektronen de vorm van de Kohn-Sham potentiaal bepalen. In de uitgevoerde analyse wordt de
exchange-korrelatie potentiaal opgesplitst in twee termen, een aantrekkende langedrachtsterm en
een repulsieve kortdrachtsterm, die beide relatief eenvoudig te benaderen zijn. Een belangrijk ken-
merk van de exchange-korrelatie potentiaal is zijn atoomschillenstruktuur. Deze komt tot uiting
in de kortedrachtspotentiaal die een stapvorm heeft. De potentiaal is konstant binnen de atoom-
schillen en verandert snel op de schilovergangen. De atoomschillenstruktuur komt ook duidelijk
tot uiting in de eigenschappen van het exchange- en exchange-korrelatie gat en de konditionele
waarschijnlijkheidsamplitude. Beide grootheden beschrijven de verdeling van de andere elektronen
wanneer bekend is dat een gegeven elektron, die wij het referentie-elektron noemen, zich op een
bepaalde plaats in de ruimte bevindt. Het exchange gat en de konditionele waarschijnlijkheidsam-
plitude ondergaan grote veranderingen wanneer het referentie-elektron een schilovergang passeert.
Omdat verschillende delen van de Kohn-Sham potentiaal direkt in de konditionele waarschijnli-
jkheidsamplitudes kunnen worden uitgedrukt kan belangrijke informatie worden verkregen over de
Kohn-Sham potentiaal door de bestudering van deze waarschijnlijkheidsamplitudes. Bijvoorbeeld
de piekenstruktuur op de atoomschilovergangen in de zogenaamde kinetische potentiaal, die een
deel is van de funktionele afgeleide van de kinetische energie funktionaal, is direkt gerelateerd
aan de veranderingen van de waarschijnlijkheidsamplitude op de schilovergangen. Zoals besproken
wordt in hoofdstuk 10 beschrijft dezelfde kinetische potentiaal het links-rechts korrelatie-effekt in
dissociërende molekulen.

In hoofdstuk 8 leiden we de stappenstruktuur van de exchange-korrelatie potentiaal die in hoofd-
stuk 7 was gekonstateerd af. We maken hierbij een benadering voor de funktionele afgeleide naar
de elektronendichtheid van de paarkorrelatiefunktie. Een analytische uitdrukking voor deze funk-
tionele afgeleide van de paarkorrelatiefunktie, die verantwoordelijk is voor de stappenstruktuur in de
exchange-korrelatie potentiaal, was nog niet eerder afgeleid. De bijbehorende exchange potentiaal
is echter op een andere wijze eerder afgeleid door Krieger, Li en Iafrate (KLI). Deze potentiaal blijkt
een heel goede benadering voor de exakte Kohn-Sham exchange potentiaal. Een ander belangrijk
resultaat is een benaderde uitdrukking voor de inverse dichtheidsresponsfunktie die in verschillende
gebieden binnen de dichtheidsfunktionaaltheorie een belangrijke rol speelt (zoals in de konstruktie

van potentialen uit dichtheden en in de berekening van energiën van aangeslagen toestanden in tijdsafhankelijke dichtheidsfunktionaaltheorie).

In hoofdstuk 9 stellen we benaderingen voor voor het stapgedeelte en het langedrachtsgedeelte van de exchange-korrelatie potentiaal. Het langedrachtsgedeelte wordt benaderd door een gradiënt-uitdrukking met het goede langeafstandsgedrag afgeleid uit Beckes gradiëntfunktionaal voor de exchange energie. De repulsieve stappotentiaal wordt geschreven als een som van orbitaaldichthe-den met voorfaktoren die afhangen van de orbitaalenergiën. Het analytisch verband tussen deze faktoren en de orbitaalenergiën wordt afgeleid door het opleggen van schalingsrelaties, ijkinvari-antie en reduktie tot het goede elektronengasresultaat voor homogene systemen. De resulterende potentiaal blijkt een praktische en nauwkeurige weergave te zijn van de exakte exchange potenti-aal en leidt bovendien tot goede (van dezelfde kwaliteit als Beckes gradiëntbenadering) exchange energiën.

In hoofdstuk 10 en 11 bekijken we molekulaire systemen. In hoofdstuk 10 onderzoeken we de eigenschappen van de Becke en Perdew gradiëntfunktionalen voor het geval van molekulaire disso-ciatie. De Becke-Perdew gradiëntfunktionaal wordt nu binnen de kwantumchemie op grote schaal toepast in molekulaire berekeningen omdat het goede geometriën en (met een nauwkeurheid van ongeveer 0.2 elektronvolt of minder) bindingsenergiën geeft. We proberen meer inzicht te verkrij-gen in de werking van deze gradiëntfunktionalen in chemische binding. We bekijken hiervoor eerst de eigenschappen van de exakte Kohn-Sham potentiaal in het bindingsmiddelpuntgebied. Zoals eerder opgemerkt vertoont de Kohn-Sham potentiaal in het bindingsmiddelpunt een piekstruktuur die gerelateerd is aan het links-rechts korrelatie-effekt. Deze piekstruktuur vindt zijn oorsprong in het kinetische gedeelte van de exchange-korrelatie funktionaal en kan worden afgeleid uit het gedrag van de konditionele waarschijnlijkheidsamplitude. We laten zien dat er ook een piekstruk-tuur aanwezig is in de exchange-korrelatie potentialen van de Becke-Perdew gradiëntfunktionaal, echter met een verkeerde funktionele afhankelijkheid van de dichtheid. We laten verder zien waarom de Becke gradiëntfunktionaal voor de exchange in het algemeen een repulsieve bijdrage levert aan de bindingsenergie en waarom de Perdew korrelatiefunktionaal gewoonlijk een aantrekkende bij-drage levert. Dit heeft te maken met het feit dat dichtheidsgradiënten in de zadelpunten van de elektronendichtheid (het bindingsmiddelpunt) nul zijn. Numerieke resultaten worden geanalyseerd aan de hand van dichtheidheidsfunktionaalberekeningen aan het stikstofmolekuul.

In hoofdstuk 11 bekijken we exakte en benaderde dichtheidsfunktionaalresultaten voor het dis-sociërende waterstofmolekuul. Het waterstofmolekuul is gekozen voor deze analyse omdat hier-voor de exakte dichtheidsfunktionaalgetallen kunnen worden berekend uit nauwkeurige elektronen-dichtheden verkregen met grote konfiguratie interaktie berekeningen. Een nadeel is echter dat het waterstofmolekuul niet een 'standaard' molekuul is in de zin dat er door de afwezigheid van gevulde kernbanen geen Pauli-repulsie optreedt bij het vormen van de binding. De dichtheidsfunktionaal-getallen worden verder vergeleken met resultaten verkregen in de Hartree-Fock benadering en met de bijna exakte konfiguratie interaktie resultaten. Het blijkt dat DFT met gradiëntfunktionalen een zeer nauwkeurige weergave geeft van de bindingskurve van het waterstofmolekuul. De kurve wijkt echter af van de exakte kurve in de dissociatielimiet, hetgeen echter gekorrigeerd kan wor-den door het uitvoeren van een onbeperkte Kohn-Sham berekening (Engels:unrestricted Kohn-Sham). De DFT dissociatiefout is echter veel kleiner dan die in de Hartree-Fock benadering. Dit komt omdat het exchange-korrelatie gat in de lokale dichtheidsbenadering en de gegeneraliseerde gradiëntbenadering gelokaliseerd is rond het referentie-elektron en hierdoor veel meer lijkt op het exakte exchange-korrelatie gat dan het gedelokaliseerde Hartree-Fock exchange gat. Dat er nog steeds een dissociatiefout in DFT bestaat is het gevolg van het feit dat de benaderde dichtheids-funktionalen niet invariant zijn onder spinrotaties hetgeen to problemen leidt voor gedegenereerde

grondtoestanden zoals de open-schil fragmenten van een dissociërend molekuul.

In hoofdstuk 12 onderzoeken we het probleem hoe de exchange-korrelatie energie te berekenen indien alleen de exchange-correlatie potentiaal bekend is. Uit de exchange potentiaal kan de exchange energie berekend worden met behulp van de Levy-Perdew relatie die kan worden afgeleid uit de schalingseigenschappen van de exchange funktionaal. Voor korrelatiefunktionalen zijn dergelijke schalingseigenschappen niet bekend. Zoals we aantonen kan de korrelatie-energie echter wel worden afgeleid uit de korrelatiepotentiaal door middel van lijnintegralen in de ruimte van elektronendichtheden. Wegafhankelijkheid van deze lijnintegralen wordt besproken en integreerbaarheidsvoorwaarden worden gegeven. We bespreken verder hoe verschillende invariantie-eigenschappen van de exchange-korrelatie potentiaal kunnen worden afgeleid door het kiezen van paden die overeenkomen met bepaalde symmetrieën (rotatie,translatie,schaling) en leiden ongelijkheden af voor de korrelatiepotentiaal. We laten vervolgens zien hoe het lijnintegraalformalisme de overgangstoestandmethode van Ziegler voor de berekening van molekulaire bindingsenergiën generaliseert.

List of publications

R.van Leeuwen and E.J.Baerends
Phys.Rev. A49:2421,1994; see chapter 5
An exchange-correlation potential with correct asymptotic behaviour

R.van Leeuwen and E.J.Baerends
Int.J.Quant.Chem. accepted;see chapter 10
An analysis of nonlocal density functionals in chemical bonding

O.V.Gritsenko,R.van Leeuwen and E.J.Baerends
Int.J.Quant.Chem. submitted;see chapter 6
Structure of the optimized effective Kohn-Sham exchange potential and its gradient approximations

R.van Leeuwen, E.van Lenthe, E.J.Baerends and J.G.Snijders
J.Chem.Phys. 101:1272,1994
Exact solutions of regular approximate relativistic wave equations for hydrogen-like atoms

R.van Leeuwen and E.J.Baerends
Phys.Rev.A submitted;see chapter 12
Energy expressions in density functional theory using line integrals

O.V.Gritsenko,R.van Leeuwen and E.J.Baerends
J.Chem.Phys. accepted;see chapter 7
Analysis of electron interaction and atomic shell structure in terms of local potentials

E.van Lenthe,R.van Leeuwen,E.J.Baerends and J.G.Snijders
in *New Challenges in Quantum Chemistry* (eds).R.Broer,P.J.C.Aerts and P.J.Bagus,Groningen University Press 1994
Regular relativistic two-component Hamiltonians

E.van Lenthe,R.van Leeuwen,E.J.Baerends and J.G.Snijders
Int.J.Quant.Chem. submitted
Regular relativistic two-component Hamiltonians

R.van Leeuwen,O.V.Gritsenko and E.J.Baerends
Z.Phys.D submitted;see chapter 8
Step structure in the atomic Kohn-Sham potential

O.V.Gritsenko,R.van Leeuwen,E.van Lenthe and E.J.Baerends
Phys.Rev.A submitted;see chapter 9
A self-consistent approximation to the Kohn-Sham exchange potential

R.van Leeuwen and E.J.Baerends
to be submitted;see chapter 10
Exchange and correlation: a density functional analysis of the hydrogen molecule

Bibliography

[1] P.Fulde. *Electron Correlations in Molecules and Solids*. Springer-Verlag, Berlin, 1991.

[2] L.Hedin and S.Lundqvist. *Solid State Physics*, 23:1, 1969.

[3] E.K.U.Gross and E.Runge. *Vielteilchentheorie*. B.G.Teubner, Stuttgart, 1986.

[4] E.K.U.Gross, E.Runge, and O.Heinonen. *Many-Particle Theory*. Hilger, Bristol, 1991.

[5] R.M.Dreizler and E.K.U.Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer-Verlag, Berlin, 1990.

[6] R.G.Parr and W.Yang. *Density Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989.

[7] R.M.Dreizler and E.K.U.Gross (eds.). *Density Functional Theory*. NATO ASI Series. Plenum Press,New York, 1994.

[8] R.M.Dreizler and J.da Providencia (eds.). *Density Functional Methods in Physics*. NATO ASI Series B123. Plenum Press,New York, 1985.

[9] S.Lundqvist and N.H.March (eds.). *Theory of the Inhomogeneous Electron Gas*. Plenum Press,New York, 1983.

[10] P.Hohenberg and W.Kohn. *Phys. Rev.*, 136B:864, 1964.

[11] W.Kohn and L.J.Sham. *Phys. Rev.*, 140A:1133, 1965.

[12] R.McWeeny. *Rev. Modern Phys.*, 32:335, 1960.

[13] R.Erdahl and V.H.Smith (eds.). *Density Matrices and Density Functionals*. Reidel, Dordrecht, 1987.

[14] A.J.Coleman. *Rev. Modern Phys.*, 35:668, 1963.

[15] E.H.Lieb. *Phys. Rev. Lett.*, 46:457, 1981.

[16] A.M.K.Müller. *Phys. Lett.*, 105A:446, 1984.

[17] M.A.Buijse. *Electron correlation*. PhD thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1991.

[18] M.Levy. *Proc. Natl. Acad. Sci. USA*, 76:6062, 1979.

[19] E.H.Lieb. *Int. J. Quantum Chem.*, 24:243, 1982.

[20] J.C.Slater. *Quantum Theory of Molecules and Solids*, volume 4. McGraw-Hill, 1963.

[21] R.McWeeny. *Methods of Molecular Quantum Mechanics*. Academic Press, 1992.

[22] M.Levy and J.P.Perdew. *Phys. Rev.*, A32:2010, 1985.

[23] Hui Ou-Yang and M.Levy. *Phys. Rev. Lett.*, 65:1036, 1990.

[24] O.Gunnarson and B.I.Lundqvist. *Phys. Rev.*, B13:4274, 1976.

[25] D.C.Langreth and J.P.Perdew. *Solid State Commun.*, 17:1425, 1975.

[26] D.C.Langreth. *Phys. Rev. Lett.*, 52:2317, 1984.

[27] M.Levy. *Phys. Rev.*, A43:4637, 1991.

[28] A.Görling and M.Levy. *Phys. Rev.*, B47:13105, 1993.

[29] A.Görling and M.Levy. *Phys. Rev.*, A50:196, 1994.

[30] F.W.Averill and G.S.Painter. *Phys. Rev.*, B24:6795, 1981.

[31] C.O.Almbladh and U.von Barth. *Phys. Rev.*, B31:3231, 1985.

[32] R.van Leeuwen and E.J.Baerends. *Phys. Rev.*, A49:2421, 1994.

[33] J.D.Talman and W.F.Shadwick. *Phys. Rev.*, A14:36, 1976.

[34] R.T.Sharp and G.K.Horton. *Phys. Rev.*, 90:317, 1953.

[35] J.B.Krieger, Yan Li, and G.J.Iafrate. *Phys. Rev.*, A45:101, 1992.

[36] E.Engel and S.H.Vosko. *Phys. Rev.*, A47:2800, 1993.

[37] J.D.Talman. *Computer Phys. Commun.*, 54:85, 1989.

[38] Yan Li, J.B.Krieger, M.R.Norman, and G.J.Iafrate. *Phys. Rev.*, B44:10437, 1991.

[39] J.B.Krieger, Yan Li, and G.J.Iafrate. *Phys. Rev.*, A46:5453, 1992.

[40] Yan Li, J.B.Krieger, and G.J.Iafrate. *Phys. Rev.*, A47:165, 1993.

[41] Yan Li, J.B.Krieger, and G.J.Iafrate. *Chem. Phys. Lett.*, 191:38, 1992.

[42] D.J.W.Geldart and M.Rasolt. *Phys. Rev.*, B13:1477, 1976.

[43] J.P.Perdew. *Physica*, B172:1, 1991.

[44] A.D.Becke. *Phys. Rev.*, A38:3098, 1988.

[45] J.P.Perdew. *Phys. Rev.*, B33:8822, 1986.

[46] J.P.Perdew and Wang Yue. *Phys. Rev.*, B33:8800, 1986.

[47] J.P.Perdew, J.A.Chevary, S.H.Vosko, K.A.Jackson, M.R.Pederson, D.J.Singh, and C.Fiolhais. *Phys. Rev.*, B46:6671, 1992.

[48] J.P.Perdew in. *Electronic Structure of Solids 1991*. (eds.)P.Ziesche and H.Eschrig, Akademie Verlag, Berlin, 1991.

[49] E.Engel, J.A.Chevary, L.D.Macdonald, and S.H.Vosko. *Z. Physik*, D23:7, 1992.

[50] D.C.Langreth and.J.Mehl. *Phys. Rev.*, B28:1809, 1983.

[51] D.C.Langreth and J.P.Perdew. *Phys. Rev.*, B21:5469, 1980.

[52] A.D.Becke. *J. Chem. Phys.*, 97:9173, 1993.

[53] J.P.Perdew. *Int. J. Quantum Chem. Symp.*, 27:93, 1993.

[54] T.Zhu, C.Lee, and W.Yang. *J. Chem. Phys.*, 98:4814, 1993.

[55] T.Zhu, C.Lee, and W.Yang. *J. Chem. Phys.*, 99:4239, 1993.

[56] L.J.Sham and M.Schlüter. *Phys. Rev. Lett.*, 51:1888, 1983.

[57] L.J.Sham. *Phys. Rev.*, B32:3876, 1985.

[58] K.Schönhammer and O.Gunnarsson. *J. Phys.*, C20:3675, 1987.

[59] W.Hanke, N.Meskini, and H.Weiler in. *Electronic Structure, Dynamics, and Quantum Structural Properties of Condensed Matter*, volume B121 of *NATO ASI Series*. (eds.)J.T.DeVreese and P.van Camp, Plenum Press,New York, 1985.

[60] L.Hedin. *Phys. Rev.*, 136:A796, 1965.

[61] L.Hedin, A.Johansson, B.I.Lundqvist, S.Lundqvist, and V.Samathiyakanit. *Arkiv för Fysik*, 39:97, 1968.

[62] E.L.Shirley and R.M.Martin. *Phys. Rev.*, B47:15404, 1993.

[63] R.W.Godby, M.Schlüter, and L.J.Sham. *Phys. Rev. Lett.*, 56:2415, 1986.

[64] H.Jeggle. *Nichtlineare Funktionalanalysis*. Studienbücher Mathematik. B.G.Teubner, Stuttgart, 1979.

[65] Y.Choquet-Bruhat, C.DeWitt-Morette, and M.Dillard-Bleick. *Analysis, Manifolds and Physics*. North-Holland Publishing, Amsterdam, 1978.

[66] I.Ekeland and R.Temam. *Convex Analysis and Variational Problems*. North-Holland Publishing, Amsterdam, 1976.

[67] H.Englisch and R.Englisch. *Phys. Stat. Sol.*, 123:711, 1983.

[68] H.Englisch and R.Englisch. *Phys. Stat. Sol.*, 124:373, 1983.

[69] E.H.Lieb and S.Oxford. *Int. J. Quantum Chem.*, 19:427, 1981.

[70] A.D.Becke. *J. Chem. Phys.*, 96:2155, 1992.

[71] L.Fan and T.Ziegler. *J. Chem. Phys.*, 95:7401, 1991.

[72] M.Levy, J.P.Perdew, and V.Sahni. *Phys. Rev.*, A30:2745, 1984.

[73] A.D.Becke. *J. Chem. Phys.*, 88:2, 1988.

[74] C.O.Almbladh and A.C.Pedroza. *Phys. Rev.*, A29:2322, 1984.

[75] E.K.U.Gross and R.M.Dreizler. *Z. Physik*, A302:102, 1981.

[76] F.Herman, J.P.van Dyke, and I.B.Ortenburger. *Phys. Rev. Lett.*, 22:807, 1969.

[77] C.Lee and Z.Zhou. *Phys. Rev.*, A44:1536, 1991.

[78] A.G. Görling. *Phys. Rev.*, A46:3753, 1992.

[79] F.Aryasetiawan and M.J.Stott. *Phys. Rev.*, B38:2974, 1988.

[80] Yue Wang and R.G.Parr. *Phys. Rev.*, A47:R1591, 1993.

[81] M.Levy. *Phys. Rev.*, A26:1200, 1982.

[82] R.O.Esquivel and A.V.Bunge. *Int. J. Quantum Chem.*, 32:295, 1987.

[83] A.V.Bunge and R.O.Esquivel. *Phys. Rev.*, A33:853, 1986.

[84] F.Herman and F.Skilman. *Atomic Structure Calculations*. Prentice-Hall Inc., 1963.

[85] S.H.Vosko, L.Wilk, and M.Nusair. *Can. J. Phys.*, 58:1200, 1980.

[86] M.A.Buijse, E.J.Baerends, and J.G.Snijders. *Phys. Rev.*, A40:4190, 1989.

[87] W.L.Luken and D.N.Beratan. *Theoret. Chim. Acta*, 61:265, 1982.

[88] V.W.Maslen. *Proc. Phys. Soc. (London)*, A69:734, 1956.

[89] G.Sperber. *Int. J. Quantum Chem.*, 6:881, 1972.

[90] R.van Leeuwen and E.J.Baerends. *Int. J. Quantum Chem.*, 1994. Accepted.

[91] K.Aashamar, T.M.Luke, and J.D.Talman. *At.Data.Nucl.Data.Tabl.*, 22:443, 1978.

[92] G.Ortiz and P.Ballone. *Phys. Rev.*, B43:6376, 1991.

[93] Y.Wang, J.P.Perdew, J.A.Chevary, L.D.Macdonald, and S.H.Vosko. *Phys. Rev.*, A41:78, 1990.

[94] E.Engel and S.H.Vosko. *Phys. Rev.*, B47:13164, 1993.

[95] A.D.Becke. *J. Chem. Phys.*, 84:4524, 1986.

[96] C.Lee, W.Yang, and R.G.Parr. *Phys. Rev.*, B37:785, 1988.

[97] L.Hedin and B.I.Lundqvist. *J. Phys.*, C4:2064, 1971.

[98] U.von Barth and L.Hedin. *J. Phys.*, C5:1629, 1972.

[99] D.M.Ceperley and B.J.Alder. *Phys. Rev. Lett.*, 45:566, 1980.

[100] F.Herman, I.B.Ortenburger, and J.P.Van Dyke. *Int. J. Quantum Chem. Symp.*, 3:827, 1970.

[101] J.C.Slater. *Phys. Rev.*, 81:385, 1951.

[102] R.F.W.Bader. *Atoms in Molecules: An Quantum Theory.* Clarendon Press, Oxford, 1990.

[103] R.Daudel, R.F.W.Bader, M.E.Stephens, and D.S.Borrett. *Can. J. Chem.*, 52:1310, 1974.

[104] N.H.March. *Phys. Lett.*, A113:476, 1986.

[105] N.H.March and R.F.Nalewajski. *Phys. Rev.*, A35:525, 1987.

[106] L.J.Bartolotti and P.K.Acharya. *J. Chem. Phys.*, 77:4576, 1982.

[107] Q.Zhao and R.G.Parr. *Phys. Rev.*, A46:2337, 1992.

[108] O.V.Gritsenko, R.van Leeuwen, and E.J.Baerends. *J. Chem. Phys.* Accepted.

[109] A.Isihara and E.W.Montroll. *Proc. Natl. Acad. Sci. USA*, 68:3111, 1971.

[110] S.H.Vosko and L.D.Macdonald in. *Condensed Matter Theories.* (eds.)P.Vashishta,R.K.Kalia and R.F.Bishop, Plenum Press,New York, 1987.

[111] G.A.Baker. *Essentials of Padé Approximants.* Academic Press, 1975.

[112] J.W.Cooley. *Math. Comp.*, 15:363, 1961.

[113] G.Hunter. *Int. J. Quantum Chem.*, 29:197, 1986.

[114] B.M.Deb and S.K.Ghosh. *Int. J. Quantum Chem.*, 23:1, 1983.

[115] M.Levy and H.Ou-Yang. *Phys. Rev.*, A38:625, 1988.

[116] A.Holas and N.H.March. *Phys. Rev.*, A44:5521, 1991.

[117] R.F.Nalewajski and P.M.Kozlowski. *Acta Physica Polonica*, A74:287, 1988.

[118] R.Daudel, R.F.W.Bader, M.E.Stephens, and D.S.Borett. *Can. J. Chem.*, 52:1310, 1974.

[119] M.A.Buijse and E.J.Baerends in. *Electronic density functional theory of molecules,clusters and solids.* (ed.)D.E.Ellis, Kluwer, Dordrecht, 1993.

[120] W.A.Bingel. *Z.Naturforschg.*, 18A:1249, 1963.

[121] O.V.Gritsenko, R.van Leeuwen, and E.J.Baerends. *Int. J. Quantum Chem.* Submitted.

[122] B.G.Johnson, P.M.W.Gill, and J.A.Pople. *J. Chem. Phys.*, 98:5612, 1992.

[123] O.Gunnarsson, M.Jonson, and B.I.Lundqvist. *Phys. Rev.*, B20:3136, 1979.

[124] M.D.Glossman, L.C.Balbás, A.Rubio, and J.A.Alonso. *Int. J. Quantum Chem.*, 49:171, 1994.

[125] O.V.Gritsenko, N.A.Cordero, A.Rubio, L.C.Balbás, and J.A.Alonso. *Phys. Rev.*, A48:4197, 1993.

[126] O.Gunnarsson, M.Jonson, and B.I.Lundqvist. *Solid State Commun.*, 24:765, 1977.

[127] O.Gunnarsson and B.I.Lundqvist. *Phys. Rev.*, B13:4274, 1976.

[128] E.Lieb. *Int. J. Quantum Chem.*, 24:243, 1982.

[129] U.von Barth. *Phys. Rev.*, A20:1693, 1979.

[130] O.V.Gritsenko, R.van Leeuwen, E.van Lenthe, and E.J.Baerends. *Phys. Rev.*, A. Submitted.

[131] M.K.Harbola and V.Sahni. *Phys. Rev. Lett.*, 62:489, 1989.

[132] R.van Leeuwen, O.V.Gritsenko, and E.J.Baerends. *Z. Physik*, D. Submitted.

[133] J.A.Alonso and L.A.Girifalco. *Phys. Rev.*, B17:3735, 1978.

[134] J.P.Perdew, R.G.Parr, M.Levy, and J.L.Balduz. *Phys. Rev. Lett.*, 49:1691, 1982.

[135] T.Koopmans. *Physica*, 1:104, 1933.

[136] A.Garciá, C.Elsässer, J.Zhu, S.G.Louie, and M.Cohen. *Phys. Rev.*, B46:9829, 1992.

[137] G.Ortiz. *Phys. Rev.*, B45:11328, 1992.

[138] V.Tschinke and T.Ziegler. *Theoret. Chim. Acta*, 81:65, 1991.

[139] L.Fan and T.Ziegler. *J. Chem. Phys.*, 94:6057, 1991.

[140] D.W.Smith, S.Jagannathan, and G.S.Handler. *Int. J. Quantum Chem. Symp.*, 13:103, 1979.

[141] D.R.Salahub and M.C.Zerner. *The Challenge of d and f Electrons: Theory and Computation.* ACS Symposium Series. American Chemical Society,Washington DC, 1989.

[142] O.Gunnarsson, P.Johansson, S.Lundqvist, and B.I.Lundqvist. *Int. J. Quantum Chem. Symp.*, 9:83, 1975.

[143] O.Gunnarsson and P.Johansson. *Int. J. Quantum Chem.*, 10:307, 1976.

[144] O.Gunnarsson, B.I.Lundqvist, and J.W.Wilkins. *Phys. Rev.*, B4:1319, 1974.

[145] O.Gunnarsson, J.Harris, and R.O.Jones. *J. Chem. Phys.*, 67:3970, 1977.

[146] B.I.Dunlap. *Chem. Phys.*, 125:89, 1988.

[147] J.P.Perdew. *Phys. Rev.*, B34:7406, 1986.

[148] O.L.Malkina, O.V.Gritsenko, and G.M.Zhidomirov. *Chem. Phys. Lett.*, 171:191, 1990.

[149] H.Przybylski and G.Borstel. *Solid State Commun.*, 49:317, 1984.

[150] H.Przybylski and G.Borstel. *Solid State Commun.*, 52:713, 1984.

[151] R.Latter. *Phys. Rev.*, 99:510, 1955.

[152] M.K.Harbola, M.Slamet, and V.Sahni. *Phys. Lett.*, A157:60, 1991.

[153] T.Ziegler, A.Rauk, and E.J.Baerends. *Theoret. Chim. Acta*, 46:1, 1977.

[154] E.J.Baerends. Bond energy evaluation. Technical report, Vrije Universiteit, Amsterdam, The Netherlands, 1988.

[155] P.M.Boerrigter. PhD thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1987.