# Electron density functional theory 

Roi Baer

Fritz Haber Center for Molecular Dynamics, Institute of Chemistry, The Hebrew University of Jerusalem

Jerusalem, June 19, 2017
Copyright © 2016 by Roi Baer
All rights reserved
Reproduction or translation of any part of this work is without the permission of the copyright owner is unlawful.

## Contents

Chapter 1. Many particle systems ..... 5
1.1. The Schrödinger equation ..... 5
1.2. Identical particles: Bosons and Fermions ..... 9
1.3. Two non-interacting particles in a trap ..... 13
1.4. Correlation: two interacting particles in a trap ..... 18
1.5. Electron density: simpler than wave function ..... 23
Chapter 2. My first density functional: Thomas-Fermi theory ..... 27
2.1. Basic concepts in Thomas-Fermi Theory ..... 32
2.2. The virial theorem for Thomas-Fermi theory ..... 39
2.3. Minimization of Thomas-Fermi energy ..... 40
2.4. Thomas-Fermi theory of atoms ..... 42
2.5. Thomas-Fermi theory cannot explain chemical bonds ..... 46
2.6. Linear response theory and Thomas-Fermi screening ..... 47
2.7. von-Weizsäcker kinetic energy ..... 49
Chapter 3. Hartree-Fock Theory ..... 51
3.1. Slater functions ..... 51
3.2. Density, reduced density matrix of Slater functions ..... 53
3.3. Expectation values in Slater functions ..... 55
3.4. Hartree-Fock theory ..... 57
3.5. The HF theory of the homogeneous electron gas ..... 62
Chapter 4. Hohenberg-Kohn theory ..... 67
4.1. Hohenberg-Kohn theorem ..... 67
4.2. Hohenberg-Kohn variational principle ..... 69
4.3. The Kohn-Sham density functional theory ..... 76
Chapter 5. Time-dependent DFT ..... 78
Appendix A. Befriending Functionals ..... 79
A.1. Functionals, linear functionals ..... 79
A.2. Functional derivatives ..... 80
A.3. Convex functionals and Legendre transforms ..... 82
Appendix B. The virial theorem ..... 88
B.1. Dilation relations ..... 88
B.2. The virial theorem - classical mechanics ..... 89
B.3. The virial theorem - quantum mechanics ..... 91
B.4. Other Dilation Facts ..... 92
Appendix C. Minimization Theory ..... 94
C.1. Necessary conditions for minimization ..... 94
C.2. Constrained minimization ..... 96
C.3. Minimization of functionals ..... 100
Appendix D. Spin states ..... 102
D.1. Spin: The Stern-Gerlach experiment ..... 102
D.2. Spin operators ..... 103
D.3. Many electron spin states ..... 105
Appendix. Bibliography ..... 109

## CHAPTER 1

## Many particle systems

### 1.1. The Schrödinger equation

Consider a system of $N$ 3D particles having with position variables $(1, \ldots N) \equiv\left(\boldsymbol{r}_{1}, \ldots, \boldsymbol{r}_{N}\right)$, in a potential

$$
\begin{equation*}
W(1, \ldots, N)=\sum_{n=1}^{N} v\left(\boldsymbol{r}_{n}\right)+\frac{1}{2} \sum_{n \neq m} u\left(\boldsymbol{r}_{m}, \boldsymbol{r}_{n}\right) \tag{1.1.1}
\end{equation*}
$$

where $v(x)$ is a potential of an external force operating on each of the particles and $u\left(\boldsymbol{r}_{m}, \boldsymbol{r}_{n}\right)$ is the potential describing the pairwise force between particles $n$ and $m$. The Schrödinger equation $\hat{H} \psi_{\alpha}(\boldsymbol{r})=E_{\alpha} \psi_{\alpha}(\boldsymbol{r})$ determines the energy levels $E_{\alpha}$ $\alpha=0,1,2, \ldots$ (in ascending order) as eigenvalues and the corresponding stationary states as eigenstates (or eigenfunctions) of the Hermitean energy operator

$$
\begin{align*}
\hat{H} & =\sum_{n=1}^{N}\left(-\frac{\hbar^{2}}{2 \mu} \nabla^{2}(n)\right)+W(1, \ldots, N)  \tag{1.1.2}\\
& \equiv-\frac{\hbar^{2}}{2 \mu} \nabla^{2}+W(1, \ldots, N)
\end{align*}
$$

where the first term is the total kinetic energy operator.
The system can be in a non-stationary state, which is a linear combination of the eigenstates:

$$
\begin{equation*}
\Psi(1, \ldots, N)=a_{\alpha} \Psi_{\alpha}(1, \ldots, N) \tag{1.1.4}
\end{equation*}
$$

where now 1 is the collection of spatial coordinates of the particle $\boldsymbol{r}_{1}$ and its spin coordinates $s_{1}$, and where $a_{\alpha}$ are the linear combination amplitudes. When a system is in this state and a measurement of the energy is performed the result is always an eigenstate of $\hat{H}$. However, one cannot predict with certainty the outcome only the probability $p_{\alpha}=\frac{\left|a_{\alpha}\right|^{2}}{\sum_{\alpha^{\prime}}\left|a_{\alpha^{\prime}}\right|^{2}}$ that the value $E_{\alpha}$ will be found. The linear amplitudes can be obtained by an integration

$$
\begin{equation*}
a_{\alpha}=\left\langle\Psi_{\alpha} \mid \Psi\right\rangle \equiv \int \Psi_{\alpha}(1, \ldots, N)^{*} \Psi(1, \ldots, N) d^{N} \boldsymbol{r} s \tag{1.1.5}
\end{equation*}
$$

where the integral is over the 3 N -dimensional spatial and the N spin "variables". This result assumes that the eigenstates are orthonormal (as can be assumed for a Hermitean operator):

$$
\begin{equation*}
\left\langle\Psi_{\alpha} \mid \Psi_{\alpha^{\prime}}\right\rangle=\delta_{\alpha \alpha^{\prime}} . \tag{1.1.6}
\end{equation*}
$$

When solving the Schrödinger equation we demand the functions belong to a space of acceptable solutions, called the Hilbert space. Basically we require that the integrals in Eqs. 1.1.6 and 1.1.5 are finite. This is achieved by selecting the Hilbert space to include all wave functions that go to zero rapidly enough when diminish to zero as $|x| \rightarrow \infty$.

One of our main theoretical tools is the variational principle of quantum mechanics. Before we state and prove it, let is cover the following exercise:

Exercise 1. How to properly define kinetic energy: A functional over a domain of functions is a mapping between each function in the domain and a number. Consider two "kinetic energy functionals" defined on the domain of normalized functions that decay as $|x| \rightarrow \infty$ fast enough for the integrals to converge. The first (we omit the spins),

$$
\begin{equation*}
T_{1}[\psi]=-\frac{\hbar^{2}}{2 m} \int \psi(\boldsymbol{r})^{*} \nabla^{2} \psi(\boldsymbol{r}) d^{3 N} r \tag{1.1.7}
\end{equation*}
$$

is the usual quantum mechanical definition of kinetic energy which we learn in QM courses, and the second,

$$
\begin{equation*}
T_{2}[\psi]=\frac{\hbar^{2}}{2 m} \int|\nabla \psi(\boldsymbol{r})|^{2} d^{3 N} r \tag{1.1.8}
\end{equation*}
$$

(1) Show that if the domain is limited to smooth $\psi$ 's these two functionals are identical: $T_{1}[\psi]=T_{2}[\psi]$.
(2) But in general they are not. For example: $\psi_{0}(x)=$ $\sqrt{\alpha} e^{-\alpha|x|}$ is a normalized wave function but has a cusp (non-smooth pointed shape) at $x=0$. Show that in this case the two definitions are in complete disagreement: $T_{1}\left[\psi_{0}\right]=-T_{2}\left[\psi_{0}\right]$ !
(3) Based on this example, which of the two functionals $T_{1}$ or $T_{2}$ is physically unacceptable as kinetic energy? Why?
(4) From this discussion, try to find a natural definition for a function which is a kinetic energy density related to the wave function, i.e. a function that:
(a) is everywhere positive.
(b) the integrates over space to the value of kinetic energy.
An important theorem in quantum mechanics shows that the ground state wave function minimizes am energy functional:

Theorem 2. (The variational principle of Quantum Mechanics) Given a Schrödinger equation $\hat{H} \Psi_{n}=E_{n} \Psi_{n}$, the solutions are orthonormal eigenfunctions $\Psi_{n}(1, \ldots, N)$ and eigenvalues indexed in increasing energy order: $E_{0} \leq E_{1} \leq \cdots \leq$ $E_{n} \leq E_{n+1} \leq \ldots$, the ground state wave function, $\Psi_{0}(1, \ldots, N)$ is the function that among all normalized functions minimizes the energy functional:

$$
E[\Psi]=T[\Psi]+W[\Psi]
$$

$$
\begin{equation*}
=\frac{\hbar^{2}}{2 m} \int|\nabla \Psi(1, \ldots, N)|^{2} d^{N} \boldsymbol{r} s+\langle\Psi| \hat{W}|\Psi\rangle \tag{1.1.9}
\end{equation*}
$$

where $\langle\Psi| \hat{W}|\Psi\rangle \equiv \int|\Psi(1, \ldots, N)|^{2} W(1, \ldots, N) d^{N} \boldsymbol{r} s$.
Proof. Let's work with one particle and one dimension (ignoring spin). The generalization to several particles and spins is trivial. Any normalized wave function is a linear combination of eigenstate: $\psi(x)=a_{\alpha} \psi_{\alpha}(x)$ where $a_{\alpha} a_{\alpha}^{*}=1$ (we assume summation over repeating indices). We note that

$$
E[\psi]=a_{\alpha^{\prime}}^{*} a_{\alpha} \int_{-\infty}^{\infty} \psi_{\alpha^{\prime}}(x)^{*}\left(-\frac{\hbar^{2}}{2 m} \psi_{\alpha}^{\prime \prime}(x)+W(x) \psi_{\alpha}(x)\right) d x
$$

$$
\begin{align*}
& =a_{\alpha^{\prime}}^{*} a_{\alpha} E_{n} \int_{-\infty}^{\infty} \psi_{\alpha^{\prime}}(x)^{*} \psi_{\alpha}(x) d x  \tag{1.1.10}\\
& =\sum_{\alpha}\left|a_{\alpha}\right|^{2} E_{\alpha} \geq \sum_{\alpha}\left|a_{\alpha}\right|^{2} E_{0}=E_{0}
\end{align*}
$$

with equality when $\psi(x)=\psi_{0}(x)$.
EXERCISE 3. Consider the quantum ground-state problem for a particle in a well $V(x)=\frac{1}{4} k x^{4}$. We consider the family of functions:

$$
\phi_{\sigma}(x)=\frac{e^{-\frac{x^{2}}{4 \sigma^{2}}}}{\sqrt{\sqrt{2 \pi \sigma}}}
$$

where $\sigma$ is a positive constant. What is the best function if this form for representing the ground state?

Solution. The functions are normalized. The energy is the expectation value of the Hamiltonian which can be obtained by integration:

$$
\begin{aligned}
E(\sigma) & =\left\langle\phi_{\sigma}\right| \hat{T}\left|\phi_{\sigma}\right\rangle+\left\langle\phi_{\sigma}\right| \hat{V}\left|\phi_{\sigma}\right\rangle \\
& =\frac{\hbar^{2}}{8 m_{e} \sigma^{2}}+\frac{3}{4} a \sigma^{4} .
\end{aligned}
$$

Minimizing this is looking for $\sigma_{*}$ for which $E^{\prime}\left(\sigma_{*}\right)=0$, hence:

$$
-\frac{\hbar^{2}}{4 m_{e} \sigma_{*}^{3}}+3 a \sigma_{*}^{3}=0
$$

From which the optimal value for $\sigma$ is:

$$
\sigma_{*}=\left(\frac{\hbar^{2}}{12 a m_{e}}\right)^{1 / 6}
$$

and the minimal energy is then

$$
\begin{aligned}
E_{*} & =E\left(\sigma_{*}\right) \\
& =\frac{3}{8}\left(\frac{3}{2}\right)^{1 / 3}\left(\frac{\hbar^{4} a}{m_{e}}\right)^{1 / 3}
\end{aligned}
$$

### 1.2. Identical particles: Bosons and Fermions

The particles in the preceding section are identical. We will write $\Psi(\ldots, i, \ldots)$ as a short hand for $\Psi\left(\ldots,\left(\boldsymbol{r}_{i}, s_{i}\right), \ldots\right)$ where $\boldsymbol{r}_{i}$ is the position variable of particle $i$ and $s_{i}$ is the spin. The basic requirement for identical :

$$
|\Psi(\ldots, i, \ldots, j, \ldots)|^{2}=|\Psi(\ldots, j, \ldots, i, \ldots)|^{2}
$$

Here $x_{n}$ are the spatial and spin coordinates of the particles in the system. Bosons (Fermions) are particles which achieve this requirement by demanding that

$$
\Psi(\ldots, i, \ldots, j, \ldots)= \pm \Psi(\ldots, j, \ldots, i, \ldots)
$$

with " + " for Bosons and "-" for Fermions. This symmetry relation has deep consequences. For example, two Fermions cannot be in the same position because

$$
\begin{equation*}
\Psi(\ldots, i, \ldots, j \equiv i, \ldots)=0 \tag{1.2.1}
\end{equation*}
$$

where $j \equiv i$ means that the spatial state and spin of particles $i$ and $j$ are the same.

Ground-state wave function of Bosons. It is seen that Fermion wave function must have zeros or "nodes". By the term nodes we mean the following:

Definition 4. A node of a continuous 1D function $\psi(x)$ is a point $x_{0}$ through which the function changes sign. (Technically we say that $\psi\left(x_{0}\right)=0$ and there exists a positive number $\delta$ such that $\psi\left(x_{1}\right) \psi\left(x_{2}\right)<0$ for all $x_{1}$ and $x_{2}$ for which $x_{0}-\delta<x_{1}<x_{0}<x_{2}<x_{0}+\delta$ ).

We will now show that the opposite exists for the ground state of Bosons: wave function has no nodes. We will then return to Fermions and show it has nodes of very special structure, which in some sense are "minimal" nodes.

Theorem 5. The ground-state of a 1D Schrödinger equation is "nodeless".

Proof. By reductio ad absurdum. Assume that $\psi(x)$ is the normalized ground state and it has a node at $x=0^{1}$. We will build a new function $\tilde{\psi}(x)$ every where equal to $|\psi(x)|$ except in a tiny interval, the $\delta$-interval, enclosing the origin: $x \in[-\delta, \delta]$ (where $\delta>0$ is finite but as small as we wish) where we will make a change that such that $E[\tilde{\psi}]<E[\psi]$. This will be the absurd part since this contradicts the variational principle. Throughout the development we will neglect

[^0]any quantity which is higher than first order in $\delta$. Considering the interval $\psi(x)$ is very nearly a straight line $\psi(x) \approx a x$, where $a=\psi^{\prime}(0)$. From this, the contribution of this interval to the kinetic energy is to leading order:
\[

$$
\begin{align*}
\delta T[|\psi|] & =\frac{\hbar^{2}}{2 m} \int_{-\delta}^{\delta}\left|\psi^{\prime}(x)\right|^{2} d x  \tag{1.2.2}\\
& =\frac{\hbar^{2}}{2 m} 2|a|^{2} \times \delta . \tag{1.2.3}
\end{align*}
$$
\]

The potential has much lower contribution to leading order is:

$$
\begin{align*}
\delta V[|\psi|] & =\int_{-\delta}^{\delta} V(x)|\psi(x)|^{2} d x  \tag{1.2.4}\\
& \approx \frac{2}{3} V(0)|a|^{2} \times \delta^{3} \tag{1.2.5}
\end{align*}
$$

and likewise, the norm:

$$
\begin{equation*}
\int_{-\delta}^{\delta}|\psi(x)|^{2} d x \approx \frac{2}{3}|a|^{2} \times \delta^{3} . \tag{1.2.6}
\end{equation*}
$$

Since we are going to change the function $\psi(x)$ in the $\delta$-interval, we see that we should only concern ourselves with the first order changes in the kinetic energy as the potential and norm will not change to first order. We introduce a new function $\tilde{\psi}(x)$ which is identical to $|\psi(x)|$ outside of the $\delta$-interval and is parabolic inside it (see Fig 1.2.1):

$$
\tilde{\psi}(x)=\left\{\begin{array}{cc}
\frac{a}{\delta} x^{2} & x \in[-\delta, \delta]  \tag{1.2.7}\\
|\psi(x)| & o / w
\end{array} .\right.
$$

Note $\tilde{\psi}(x)$ is continuous (to first order in $\delta$ ). The interval contribution to the kinetic energy of $\tilde{\psi}$ is


Figure 1.2.1. Reductio ad absurdum proof: $\psi(x)$ is the "noded" groundstate with supposedly minimum energy $E[\psi]$ (blue); $|\psi(x)|$ replaces the node by a pointed cusp preserving energy, $E[\psi]=E[|\psi|] . \tilde{\psi}(x)$ is the function of Eq. (1.2.7) for which $E[\tilde{\psi}]<E[\psi]$ thereby contradicting the variational principle. Note that $\tilde{\psi}$ is discontinuous to second order in $\delta$.

$$
\begin{align*}
\delta T[\tilde{\psi}] & =\frac{\hbar^{2}}{2 m} \int_{-\delta}^{\delta}\left|\tilde{\psi}^{\prime}(x)\right|^{2} d x  \tag{1.2.8}\\
& =\frac{1}{5} \frac{\hbar^{2}}{2 m} a^{2} \times \delta^{3} \tag{1.2.9}
\end{align*}
$$

hence is negligible to first order. Thus we found a function of lower energy than $E[\psi]$ contradicting the variational principle. This proves the theorem that ground states must be nodeless.

A similar proof exists for wave function of $N$ identical bosons has no nodes.

Fermions. The Fermions have antisymmetric wave functions which therefore must be excited states of the Hamiltonian (since the ground state is symmetric, bosonic, and has no nodes). The ground-state wave function of $N>1$ Fermions is more complicated than that of $N>1$ Bosons. This is because of the Pauli principle which states $\psi\left(x_{1}, x_{2}\right)=-\psi\left(x_{2}, x_{1}\right)$, i.e. $\psi$ is bound to have nodes in general. Thus, the Fermion ground-state is actually a highly excited eigenstate of the Hamiltonian. It can still be described by a minimum principle, namely the normalized many-body antisymmetric wave-function which minimizes $E[\psi]=\langle\psi| \hat{H}|\psi\rangle$. This formulation of the ground state problem is very convenient. We do not have to look for many eigenstates of $\hat{H}$, we need only search for minimum. Finding the minimal energy and the expectation values of various physically important operators is the main challenge of electronic structure theory. Doing the same for low-lying excited states is the grand challenge of this field.

The concept of correlation is highly important in quantum chemistry. It is a general concept which arises from the interaction between particles, causing them to correlate their motion. This correlated motion results in a very complicated structure of the wave function. This issue is

### 1.3. Two non-interacting particles in a trap

We have stressed that the electronic structure problem is difficult because of the complexity of the many-body wave function when electrons interact. Equation Section (Next)In order to appreciate this complexity of the electronic wave function, let us first study a simple system, of two non-interacting electrons in a 1D "atomic" well. We consider an atomic well given by the potential v_ext (x) and we place in it an electron. The Hamiltonian is:

$$
\begin{equation*}
\hat{h}=-\frac{\hbar^{2}}{2 m_{e}} \nabla^{2}+v_{e x t}(\boldsymbol{x}) . \tag{1.3.1}
\end{equation*}
$$

The eigenstates and eigenvalues of this Hamiltonian are the $\psi_{\alpha}(x)$ and $\varepsilon_{\alpha}(\alpha=0,1, \ldots)$ respectively:

$$
\hat{h} \psi_{\alpha}(x)=\varepsilon_{\alpha} \psi_{\alpha}(x),
$$

and we follow the convention that the energies form a non decreasing series $\varepsilon_{0} \leq \varepsilon_{1} \leq \ldots$. Now consider a 2-electron problem.

Now assume we have two electrons, Fermions, non-interacting in this potential well $v_{\text {ext }}(x)$. The Hamiltonian of such a system is

$$
\hat{H}=\hat{h}(1)+\hat{h}(2) .
$$

The notation $\hat{h}(i)$ means the Hamiltonian of Eq. (1.3.1) applied to particle $i=1,2$. What are the eigenstates in this case? First, since each electron can have a spin, we must decide on the spin of the state. For now, let us assume the state is that of "total spin 1", both electrons are in spin-up orbitals $\alpha$. Hence we try the following form as a wave function:

$$
\Psi(1,2)=\frac{1}{\sqrt{2}}\left[\psi_{0}(1) \psi_{1}(2)-\psi_{0}(2) \psi_{1}(1)\right] \alpha(1) \alpha(2) .
$$

EXERCISE 6. Show that the factor $\frac{1}{\sqrt{2}}$ is necessary for normalization of $\Psi$ (assume that $\psi_{\alpha}$ are a orthonormal set) and $\langle\alpha(1) \alpha(2) \mid \alpha(1) \alpha(2)\rangle=\langle\alpha \mid \alpha\rangle\langle\alpha \mid \alpha\rangle=1$.

Solution. Note the following

$$
\begin{aligned}
2\langle\Psi \mid \Psi\rangle & =\left\langle\psi_{0}(1) \psi_{1}(2) \mid \psi_{0}(1) \psi_{1}(2)\right\rangle \\
& -\left\langle\psi_{0}(1) \psi_{1}(2) \mid \psi_{0}(2) \psi_{1}(1)\right\rangle \\
& -\left\langle\psi_{0}(2) \psi_{1}(1) \mid \psi_{0}(1) \psi_{1}(2)\right\rangle \\
& +\left\langle\psi_{0}(2) \psi_{1}(1) \mid \psi_{0}(2) \psi_{1}(1)\right\rangle
\end{aligned}
$$

This is then:

$$
\begin{aligned}
2\langle\Psi \mid \Psi\rangle & =\left\langle\psi_{0} \mid \psi_{0}\right\rangle\left\langle\psi_{1} \mid \psi_{1}\right\rangle \\
& -\left\langle\psi_{0} \mid \psi_{1}\right\rangle\left\langle\psi_{1} \mid \psi_{0}\right\rangle \\
& -\left\langle\psi_{0} \mid \psi_{1}\right\rangle\left\langle\psi_{1} \mid \psi_{0}\right\rangle \\
& +\left\langle\psi_{0} \mid \psi_{0}\right\rangle\left\langle\psi_{1} \mid \psi_{1}\right\rangle
\end{aligned}
$$

Since $\left\langle\psi_{1} \mid \psi_{0}\right\rangle=0$ and $\left\langle\psi_{0} \mid \psi_{0}\right\rangle=\left\langle\psi_{1} \mid \psi_{1}\right\rangle$ we see that

$$
2\langle\Psi \mid \Psi\rangle=1 \times 1-0-0+1 \times 1=2
$$

hence $\langle\Psi \mid \Psi\rangle=1$.
EXERCISE 7. Show that the $\Psi(1,2)$ is an eigenstate of $\hat{H}$ with eigenvalue equal to $\varepsilon_{0}+\varepsilon_{1}$ :
(1.3.2) $\hat{H} \Psi(1,2)=\left(\varepsilon_{0}+\varepsilon_{1}\right) \Psi(1,2)$

Solution. It is easy to see that $\psi_{0}(1) \psi_{1}(2)$ is an eigenstate of $\hat{H}$ :

$$
\begin{aligned}
(\hat{h}(1)+\hat{h}(2)) \psi_{0}(1) \psi_{1}(2) & =\hat{h}(1) \psi_{0}(1) \psi_{1}(2) \\
& +\psi_{0}(1) \hat{h}(2) \psi_{1}(2) \\
& =\left(\varepsilon_{0}+\varepsilon_{1}\right) \psi_{0}(1) \psi_{1}(2)
\end{aligned}
$$

From this, Eq. (1.3.2) follows immediately .
We now build a simple electronic structure model that will allow us to study in some detail the most basic concepts. For this, we suppose that the electrons are in a harmonic atom, that is the potential well:

$$
\begin{equation*}
v_{e x t}(x)=\frac{1}{2} m_{e} \omega^{2} x^{2} \tag{1.3.3}
\end{equation*}
$$

The two lowest eigenstates of the Harmonic oscillator are:

$$
\begin{align*}
& \psi_{0}(x)=N_{0} e^{-\frac{1}{2} \frac{m_{e} \omega}{\hbar} x^{2}}  \tag{1.3.4}\\
& \psi_{1}(x)=N_{1} e^{-\frac{1}{2} \frac{m_{e} \omega}{\hbar} x^{2}} \sqrt{\frac{m_{e} \omega}{\hbar} x} \tag{1.3.5}
\end{align*}
$$

where the eigenvalues are

$$
\begin{equation*}
\varepsilon_{\alpha}=\left(\frac{1}{2}+\alpha\right) \hbar \omega, \alpha=0,1, \ldots \tag{1.3.6}
\end{equation*}
$$

The normalization constants in Eqs, (1.3.4)-(1.3.5) are:

$$
\begin{aligned}
& N_{0}=\left(\frac{m_{e} \omega}{\pi \hbar}\right)^{1 / 4} \\
& N_{1}=\left(\frac{4 m_{e} \omega}{\pi \hbar}\right)^{1 / 4}
\end{aligned}
$$

As discussed above singlet and triplet two-electron ground state wave functions composed of two orbitals must be spacesymmetric or antisymmetric, respectively. We consider below 3 wave functions. The first $\Psi_{S, 00}$ is the ground state singlet where both electrons are in $\psi_{0}$ :

$$
\begin{align*}
\Psi_{S, 00}(1,2) & =\psi_{0}(1) \psi_{0}(2) \\
& =N_{00} e^{-\frac{1}{2} \frac{m_{e \omega}}{\hbar}\left(x_{1}^{2}+x_{2}^{2}\right)} \tag{1.3.7}
\end{align*}
$$

The second and third are a singlet made from one electron in $\psi_{0}$ and the other in $\psi_{1}$ :

$$
\begin{align*}
\Psi_{S, 01}(1,2) & =\frac{1}{\sqrt{2}}\left[\psi_{0}(1) \psi_{1}(2)+\psi_{0}(2) \psi_{1}(1)\right] \\
& =N_{01} e^{-\frac{1}{2} \frac{m_{e \omega}}{\hbar}\left(x_{1}^{2}+x_{2}^{2}\right)}\left(x_{1}+x_{2}\right) \tag{1.3.8}
\end{align*}
$$

and finally, a triplet from the same orbitals

$$
\begin{align*}
\Psi_{T, 01}(1,2) & =\frac{1}{\sqrt{2}}\left[\psi_{0}(1) \psi_{1}(2)-\psi_{0}(2) \psi_{1}(1)\right] \\
& =N_{01} e^{-\frac{1}{2} \frac{m_{e \omega}}{\hbar}\left(x_{1}^{2}+x_{2}^{2}\right)}\left(x_{1}-x_{2}\right) \tag{1.3.9}
\end{align*}
$$

The normalization constants are:

$$
\begin{aligned}
& N_{00}=N_{0}^{2}=\sqrt{\frac{m_{e} \omega}{\pi \hbar}} \\
& N_{01}=\frac{1}{\sqrt{2}} \sqrt{\frac{m_{e} \omega}{\hbar}} N_{0} N_{1} \equiv \frac{m_{e} \omega}{\pi \hbar}
\end{aligned}
$$

Eqs.(1.3.7)-(1.3.9) describe the distribution of positions of both electrons in their corresponding states. How much are the electrons in this state aware of each other? Do they correlate their motion in some way? One way to measure correlation is to consider the position correlation constant defined as

$$
\begin{equation*}
C_{\Psi}=\langle\Psi| x_{1} x_{2}|\Psi\rangle-\langle\Psi| x_{1}|\Psi\rangle\langle\Psi| x_{2}|\Psi\rangle . \tag{1.3.10}
\end{equation*}
$$

If electrons are completely unaware of each other this quantity is zero because then the average of the product of their position must decompose to the product of the average. Any deviance from zero indicates some degree of correlation.

EXERCISE 8. Prove the following, for the wave function in Eqs.(1.3.7)-(1.3.9) :

$$
\begin{gathered}
\left\langle x_{1}\right\rangle_{\Psi}=\langle\Psi| x_{1}|\Psi\rangle=0 \\
\left\langle x_{1} x_{2}\right\rangle_{S, 00}=\left\langle\Psi_{S, 00}\right| x_{1} x_{2}\left|\Psi_{S, 00}\right\rangle=0 \\
\left\langle x_{1} x_{2}\right\rangle_{S, 01}=\left\langle\Psi_{S, 01}\right| x_{1} x_{2}\left|\Psi_{S, 01}\right\rangle=\frac{1}{2} \frac{\hbar}{m_{e} \omega} \\
\left\langle x_{1} x_{2}\right\rangle_{T, 01}=\left\langle\Psi_{T, 01}\right| x_{1} x_{2}\left|\Psi_{T, 01}\right\rangle=-\frac{1}{2} \frac{\hbar}{m_{e} \omega}
\end{gathered}
$$

We find from Exercise 8 the following position correlation constants:

$$
\begin{gathered}
C_{S, 00}=0 . \\
C_{S, 01}=\frac{1}{2} \frac{\hbar}{m_{e} \omega} .
\end{gathered}
$$

$$
C_{T, 01}=-\frac{1}{2} \frac{\hbar}{m_{e} \omega} .
$$

Since there is no interaction between the electrons, the correlation in these wave functions arises only from the Pauli principle, i.e. because we impose the fact that electrons are Fermions. This is called Fermi correlation. Our lesson is this:
(1) Wave functions that are mere products of singe-particle orbitals have no correlation.
(2) If the products are symmetrized (antisymmetrized) like in the case of the excited singlet (triplet) the position correlation constant $C$ is positive (negative) indicating that the particles "like to (not) be together" i.e. a bias for (not) being both on the right or on the left of the origin.

### 1.4. Correlation: two interacting particles in a trap

Up to now, we assumed no e-e interaction. So now let's include it and add to the Hamiltonian an interaction term:

$$
\begin{equation*}
\hat{H}=\hat{h}(1)+\hat{h}(2)+W(1,2), \tag{1.4.1}
\end{equation*}
$$

where, as before, $\hat{h}$ is given in Eq. (1.3.1) and $v_{e x t}(x)$ in Eq. (1.3.3) and we take a coupling which is simple enough to yield to analytical analysis:

$$
\begin{equation*}
W(1,2)=m_{e} \gamma^{2} x_{1} x_{2}, \tag{1.4.2}
\end{equation*}
$$

with $\gamma^{2}<\omega^{2}$. This interaction seems strange at first site because it does not depend on the distance between the particles, as we are used to from electrostatics, yet, it does describe a repulsion. We can see this in two ways. Once, since if 1 and 2 are both large and of the same sign this is energy-costly; if they are both large and of opposite sign that lowers energy. Second,
the Hamiltonian can be written as:

$$
\begin{equation*}
\hat{H}=\hat{h}^{\prime}(1)+\hat{h}^{\prime}(2)+W^{\prime}(1,2) \tag{1.4.3}
\end{equation*}
$$

where the same $\hat{H}$ in Eq. (1.4.1) is now written as describing two identical particles, each with $\hat{h}^{\prime}=-\frac{\hbar^{2}}{2 m_{e}} \frac{d^{2}}{d x^{2}}+v_{e x t}^{\prime}(x)$ with single particle Hamiltonian in a new potential well $v_{e x t}^{\prime}(x)=$ $\frac{1}{2} m_{e}\left(\omega^{2}+\gamma^{2}\right) x^{2}$ and interacting via a new repulsive spring force depending on their distance:

$$
\begin{equation*}
W^{\prime}(1,2)=-\frac{1}{2} \gamma^{2} m_{e}\left(x_{1}-x_{2}\right)^{2} \tag{1.4.4}
\end{equation*}
$$

Exercise 9. Show that if $\gamma^{2}>\omega^{2}$ the total potential $V\left(x_{1}, x_{2}\right)=$ $\frac{1}{2} m_{e}\left[\left(\omega^{2}+\gamma^{2}\right) x_{1}^{2}+\left(\omega^{2}+\gamma^{2}\right) x_{2}^{2}-\gamma^{2}\left(x_{1}-x_{2}\right)^{2}\right]$ cannot support bound states.

Solution. We'll show that the energy drops as particles receded from the origin. . Take $x_{1}=-x_{2}$ the potential is $V\left(x_{1},-x_{1}\right)=\frac{1}{2} m_{e}\left[2\left(\omega^{2}-\gamma^{2}\right) x_{1}^{2}\right]$ hence, if $\gamma^{2}>\omega^{2}$ the energy can be lowered in definitely when from below as $x_{1}=$ $-x_{2} \rightarrow \infty$.

Exercise 10. Find the eigenvalues and eigenfunctions of this Hamiltonian.

Solution. Define new coordinates $X$ and $x$ by: $(X+x) / \sqrt{2}=$ $x_{1}$ and $(X-x) / \sqrt{2}=x_{2}$. The conjugate momenta are: $p_{1}=$ $(P+p) / \sqrt{2}$ and $p_{2}=(P-p) / \sqrt{2}$ (show that the new momenta are indeed conjugate to the new positions by testing that the commutation relations $[P, X]=[p, x]=-i \hbar$ and $[p, X]=[P, x]=0)$. Then, a straightforward calculation shows the new Hamiltonian separates into two non-interacting harmonic oscillators of mass $m_{e}$ :

$$
\begin{equation*}
\hat{H}=\left[\frac{P^{2}}{2 m_{e}}+\frac{1}{2} m_{e} \Omega_{1}^{2} X^{2}\right]+\left[\frac{p^{2}}{2 m_{e}}+\frac{1}{2} m_{e} \Omega_{2}^{2} x^{2}\right], \tag{1.4.5}
\end{equation*}
$$



Figure 1.4.1. The frequencies $\Omega_{1}$ and $\Omega_{2}$ of the decoupled oscillators in the problem of two particles in a Harmonic well of frequency $\omega$ coupled by a repulsive harmonic interaction of frequency $\gamma=\omega \sqrt{\cos \theta}$ and where $0 \leq \theta \leq \pi / 2$.
where $\Omega_{1}=\sqrt{\omega^{2}+\gamma^{2}}$ and $\Omega_{2}=\sqrt{\omega^{2}-\gamma^{2}}$. Let us define $\cos \theta=(\gamma / \omega)^{2}$ for $\theta \in[0, \pi / 2]$ so that when $\theta=\pi / 2$ the two particles are non-interacting $(\gamma=0)$, and when $\theta=0$ the interaction is maximal $(\gamma=\omega)$. With this, $\Omega_{1}=\sqrt{2} \omega \cos \frac{\theta}{2}$ and $\Omega_{2}=\sqrt{2} \omega \sin \frac{\theta}{2}$. This representation shows that $\left(\Omega_{1}, \Omega_{2}\right)$ is a always point on the arc of radius $\sqrt{2} \omega$ making an angle $\theta / 2$ with the x axis of the x - y plain.

We find the eigenstates are simple products of eigenstates of the left times the right Hamiltonian in Eq. (1.4.5):

$$
\begin{align*}
\Psi_{\alpha \beta}(X, x) & =\psi_{\alpha}\left(X ; \Omega_{1}\right) \psi_{\beta}\left(x ; \Omega_{2}\right)  \tag{1.4.6}\\
E_{\alpha \beta} & =\left(\alpha+\frac{1}{2}\right) \hbar \Omega_{1}+\left(\beta+\frac{1}{2}\right) \hbar \Omega_{2} \\
\alpha, \beta & =0,1,2, \ldots
\end{align*}
$$

where:

$$
\begin{aligned}
\psi_{\alpha}(z ; \Omega) & =N_{\alpha} e^{-\frac{q^{2}}{2}} H_{\alpha-1}(q) \\
q & =\sqrt{\frac{m_{e} \Omega}{\hbar}} z
\end{aligned}
$$

and $H_{\alpha}(q)$ are the Hermite polynomials of order $\alpha$ and $N_{\alpha}$ are orthogonalization constants.

Exercise 11. Write down the ground-state energy and wave function for the triplet state of the system in the previous exercise. Determine the effect of interaction on the energy by calculating $r=E_{g s, \gamma} / E_{g s, \gamma=0}$ and on the correlation coefficient $C$.

Solution. We need to find the lowest energy solution which is antisymmetric in 1 and 2 . The two variables $X=\left(x_{1}+x_{2}\right) / \sqrt{2}$ and $x=\left(x_{1}-x_{2}\right) / \sqrt{2}$ are respectively symmetric and antisymmetric combinations of the positions of the electrons. Since $x$ is the antisymmetric combination we require the Hermite polynomial in $x$ to be odd. The lowest antisymmetric state is the combination $\alpha=0$ for $X$ and $\beta=1$ for $x$, i.e. the wave function $\Psi_{01}(X, x)$ of Eq. (1.4.6) which can be written in terms of $x_{1}$ and $x_{2}$ :
(1.4.7) $\Psi_{01}(1,2)=N_{01} e^{-\frac{m_{e} \Omega_{1}}{4 \hbar}\left(x_{1}+x_{2}\right)^{2}-\frac{m_{e} \Omega_{2}}{4 \hbar}\left(x_{1}-x_{2}\right)^{2}} \frac{x_{1}-x_{2}}{\sqrt{2}}$.
where

$$
\begin{equation*}
N_{01}=\sqrt{\frac{2 m^{2}}{\pi} \sqrt{\Omega_{1} \Omega_{2}^{3}}} \tag{1.4.8}
\end{equation*}
$$



Figure 1.4.2. The triplet wave functions for two Fermions of mass $m_{e}=1$ at positions $x_{1}$ and $x_{2}$ in an harmonic well with frequency $\omega$ interacting via a repulsive interaction (Eq. (1.4.2)) of various interaction strengths $\gamma$ determined by the angle $\theta$ where $\cos \theta=$ $(\gamma / \omega)^{2}$.

In Fig. 1.4.2 we see how the interaction "distorts" the wave function forcing the particles to become more distant from each other. This triplet groundstate energy can be written as

$$
E_{01}=\frac{1}{2} \hbar\left(\Omega_{1}+3 \Omega_{2}\right)=\frac{1}{2} \hbar \omega \sqrt{2}\left(\cos \frac{\theta}{2}+3 \sin \frac{\theta}{2}\right) .
$$

The ratio $r(\theta)=E_{g s, \gamma} / E_{g s, \gamma=0}=\left(\cos \frac{\theta}{2}+3 \sin \frac{\theta}{2}\right) / 2 \sqrt{2}$ shows the dependence of the energy on the interaction strength plotted in Fig. 1.4.3(top). It is seen that the repulsive interaction
lowers the total energy because the. This is explained by observing that on the average $X=0$ irrespective of $\gamma$ and so in the Hamiltonian of Eq. (1.4.5) the first term does not increase much as $\gamma$ grows is small due to because the triplet wave function structure promotes electrons being away from each other. Thus one is pushed towards the $+x$ direction and the other towards that of $-x$ and thus they acquire a large negative value of $x_{1} x_{2}$. To see this note that the expectation values of $x$ and $X$ are both zero and therefore $\left\langle x_{1}\right\rangle$ and $\left\langle x_{2}\right\rangle$ are zero as well. Furthermore, note that $X^{2}-x^{2}=2 x_{1} x_{2}$ and the hence the position correlation constant is: $C(\theta)=\frac{1}{2}\left(\left\langle X^{2}\right\rangle-\left\langle x^{2}\right\rangle\right)=$ $\frac{1}{4 \sqrt{2}} \frac{\hbar}{m_{e} \omega}\left(\frac{1}{\cos \frac{\theta}{2}}-\frac{3}{\sin \frac{\theta}{2}}\right)$. In Fig. 1.4.3(bottom) we plot $C(\theta) / C(\pi / 2)$, where $C(\pi / 2)=-\frac{1}{2} \frac{\hbar}{m_{e} \omega}$ is negative as discussed in Exercise 8 for non-interacting particles.

### 1.5. Electron density: simpler than wave function

The complexity of the wave function $\Psi\left(1,2, \ldots, N_{e}\right)$ is overwhelming. One piece of information is the single electron probability density $p(\boldsymbol{r})$. This function gives the probability density of finding an electron at point $r$ in space. It is the expectation value of an operator, $p\left(\boldsymbol{r}_{1}\right)=\langle\Psi| \frac{1}{N_{e}} \sum_{n=1}^{N_{e}} \delta\left(\hat{\boldsymbol{r}}_{n}-\boldsymbol{r}_{1}\right)|\Psi\rangle$. It can also be represented as a partial integral over the wave function (assuming all spins are integrated over):

$$
\begin{equation*}
p\left(\boldsymbol{r}_{1}\right)=\int\left|\Psi\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2} \ldots, \boldsymbol{r}_{N_{e}}\right)\right|^{2} d^{3} r_{2} \cdots d^{3} r_{N} \tag{1.5.1}
\end{equation*}
$$

The density of electrons in real space is closely related to the probability density

$$
\begin{equation*}
n(\boldsymbol{r})=N_{e} p\left(\boldsymbol{r}_{1}\right), \tag{1.5.2}
\end{equation*}
$$

clearly, $n(\boldsymbol{r}) d^{3} r$ is the number of electrons in an infinitesimal cube of volume $d^{3} r$ around the point $r$. The integral of the density $n(\boldsymbol{r})$ over entire space must equal the total number of electrons: $\int n(\boldsymbol{r}) d^{3} r=N_{e}$.


Figure 1.4.3. Upper panel: The energy ratio $r(\theta)$ vs interaction strength parameter $\theta$ between interacting and non-interacting particles in the Harmonic oscillator; Lower panel: the ratio $C(\theta) / C(\pi / 2)$ (where $\theta=\frac{\pi}{2}$ is the noninteraction particle regime) of position correlation constants.

Another layer of information deals with pairs of electrons. The pair probability $P\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right)$, is the probability density of finding an electron at $\boldsymbol{r}_{1}$ and another electron at $\boldsymbol{r}_{2}$. This quantity too is an expectation value,

$$
P\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right)=\langle\Psi| \frac{1}{N_{e}\left(N_{2}-1\right)} \sum_{n \neq m} \delta\left(\hat{\boldsymbol{r}}_{n}-\boldsymbol{r}_{1}\right) \delta\left(\hat{\boldsymbol{r}}_{m}-\boldsymbol{r}_{2}\right)|\Psi\rangle
$$

and can be represented as an integral (assuming all spins are integrated over):

$$
\begin{equation*}
P\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right)=\int\left|\Psi\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2} \ldots, \boldsymbol{r}_{N_{e}}\right)\right|^{2} d^{3} r_{3} \cdots d^{3} r_{N} \tag{1.5.3}
\end{equation*}
$$

Another information is the reduced density matrix (RDM), which is given as a partial integral:
(1.5.4)
$\gamma\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)=N \int \Psi\left(\boldsymbol{r}, \boldsymbol{r}_{2} \ldots, \boldsymbol{r}_{N_{e}}\right) \Psi\left(\boldsymbol{r}^{\prime}, \boldsymbol{r}_{2} \ldots, \boldsymbol{r}_{N_{e}}\right)^{*} d^{3} r_{2} \cdots d^{3} r_{N}$
Clearly, the density can be obtained as a diagonal element of the RDM: $n(\boldsymbol{r})=\gamma(\boldsymbol{r}, \boldsymbol{r})$.

One can go on and obtain information concerning triples of particles etc. However, for most purposes $n\left(\boldsymbol{r}_{1}\right)$ and $P\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right)$ and the $\operatorname{RDM} \gamma\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)$ are all we need. For example the potential energy of electron electron repulsion is given by (1.5.5)

$$
E_{e e}=N_{e}\left(N_{e}-1\right) \int P\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right) u_{C}\left(\left|\boldsymbol{r}_{1}-\boldsymbol{r}_{2}\right|\right) d^{3} r_{1} d^{3} r_{2}
$$

or the kinetic energy is

$$
\begin{equation*}
T=\iint\left[\nabla_{\boldsymbol{r}} \cdot \nabla_{\boldsymbol{r}^{\prime}} \gamma\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)\right] \delta\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) d^{3} r d^{3} r^{\prime} \tag{1.5.6}
\end{equation*}
$$

and the energy of interaction with a single particle potential:

$$
\begin{equation*}
V_{e x t}=\int v_{e x t}(\boldsymbol{r}) n(\boldsymbol{r}) d^{3} r \tag{1.5.7}
\end{equation*}
$$

Looking at Eqs. (1.5.1) and (1.5.2), we see that $n(\boldsymbol{r})$ involves integrating out a huge amount of wave-function details. Only the data concerning the density distribution of a single electron remains! This is multiplied by $N_{e}$ in Eq. (1.5.2) so $\mathrm{n}(\mathrm{r})$ accounts for the combined density of all electrons.

Exercise 12. Calculate the 1D electron density of the triplet ground state from

Solution. Use the wave function of Eq. (1.4.7) and to compute the density:
(1.5.8)

$$
\begin{aligned}
n(x) & =2 \int\left[N_{01} e^{-\frac{m_{e} \Omega_{1}}{4 \hbar}\left(x+x_{2}\right)^{2}-\frac{m_{e} \Omega_{2}}{4 \hbar}\left(x-x_{2}\right)^{2}} \frac{x-x_{2}}{\sqrt{2}}\right]^{2} d x_{2} \\
(1.5 .9) & =\frac{\sqrt{\frac{2}{\pi}} \sqrt{\frac{m}{4 \hbar} \Omega\left(\frac{\Omega_{2}}{\hbar}\right)}}{\left(\Omega_{2}+\Omega_{2}\right)^{2}} e^{-\frac{m \Omega}{2 \hbar} x^{2}}\left(4 m \Omega_{1}^{2} x^{2}+\hbar\left(\Omega_{1}+\Omega_{2}\right)\right)
\end{aligned}
$$

where $\Omega=\frac{4 \Omega_{1} \Omega_{2}}{\Omega_{1}+\Omega_{2}}$.

## CHAPTER 2

## My first density functional: Thomas-Fermi theory

The solution of the Schrodinger equation of materials or molecules allows in-principle understanding and prediction of their properties. However, such a solution is not easily available especially not with the required accuracy for modernday applications. The problem is largely in the basic concept of quantum mechanics, namely the correlated wave function $\Psi(1,2, \ldots, N)$ of $N$ interacting particles. We have seen that even for two particles, the wave function is a complicated nontrivial object. A naive approach for its representation for large number $N$ of particles quickly ends up with exponentially large resources. For example, suppose that the amplitude for a spatial degrees of freedom has $F$ possibilities. Then for $3 N$ degrees of freedom the total amount of information the wave function must hold is $F^{3 N}$. Even for a modest case of $N=30$ electrons and $F=10$ this number far exceeds the estimated number of atoms in the visible universe. The wave functions of non-interacting particles on the other hand are much simpler as each particle is described by it's own wave 3D function, and hence the amount of data is linear with system size, i.e. $N \times F^{3}$.

In view of this, scientists are are continuously seeking for new ways to bypass the need for a description of the manybody wave function. One such way is to develop a method which does not require knowledge of the correlated wave function at all. In fact, even if we had at our disposal the entire
wave function, we would probably only use a small amount of the vast data it offers. We almost always need only average properties of an electron or a pair of electrons. The average properties of a pair require knowledge of the single electron density matrix which we discuss in later chapters. Because the electron-electron Coulomb repulsion is a pairwise interaction the electronic energy depends only on the relative distance of pairs of particles, i.e. on the "pair probability", namely the probability to find a pair of electrons, one at point $r$ of space and the other at point $\boldsymbol{r}^{\prime}$. We will discuss the pair distribution in a later chapter as well.

The simplest non-trivial quantity that can be extracted from the electronic wave function is the electron density $n(\boldsymbol{r})$. This 3D-function tells us the expectation value of the density of electrons. That is, $n(\boldsymbol{r}) d^{3} r$ is the expectation value of the number of electrons in a small volume $d^{3} r$ around point $r$ :

$$
\begin{equation*}
n(\boldsymbol{r})=\langle\Psi| \hat{n}(\boldsymbol{r})|\Psi\rangle \tag{2.0.1}
\end{equation*}
$$

where

$$
\begin{equation*}
\hat{n}(\boldsymbol{r})=\sum_{n=1}^{N} \delta\left(\boldsymbol{r}-\hat{\boldsymbol{r}}_{n}\right) \tag{2.0.2}
\end{equation*}
$$

and $\hat{\boldsymbol{r}}_{n}$ is the operator corresponding to the position vector of electron $n, n=1, \ldots, N$. The 3D delta-function symbol $\delta(\boldsymbol{a})$, where $\boldsymbol{a}$ is a 3D vector, has the property that for any function $f: R^{3} \rightarrow R$ in 3D space and any point $\boldsymbol{a}$ :

$$
\begin{equation*}
\int \delta\left(\boldsymbol{a}^{\prime}-\boldsymbol{a}\right) f\left(\boldsymbol{a}^{\prime}\right) d^{3} a^{\prime}=f(\boldsymbol{a}) \tag{2.0.3}
\end{equation*}
$$

Exercise 13. Using the fact that the electrons are identical particles obeying the Pauli principle prove that

$$
\begin{equation*}
n\left(\boldsymbol{r}_{1}\right)=N \int|\Psi(1,2, \ldots 3)|^{2} d(2 \ldots N) \tag{2.0.4}
\end{equation*}
$$

Looking at Eq. (2.0.4), we see that $n(\boldsymbol{r})$ involves integrating out a huge amount of wave-function details. Only the data
concerning the average density distribution of a single electron remains! This is multiplied by $N$ so $n(\boldsymbol{r})$ accounts for the combined density of all electrons. Indeed, integrating over the entire space, one obtains:

$$
\begin{equation*}
\int n\left(\boldsymbol{r}_{1}\right) d^{3} r_{1}=N \tag{2.0.5}
\end{equation*}
$$

The density is such a "washed out" version of the wave function that we would think it is not capable of holding all the information the wave function holds. However, in a certain sense this is not true and it does. This was discovered by Hohenberg and Kohn and forms the basis of density functional theory (DFT).

Before we describe the basics of DFT, we give here an early theory developed independently by Llewellyn H Thomas [7] and a year later by Enrico Fermi [1] which is an early form of an approximation to the in-principle exact density functional theory which came some 40 years later. A detailed account of Thomas-Fermi theory including application to the study of stability of matter in the universe can be found in works of Elliott Lieb . [4, 5]

Before giving a simplified account of the Thomas Fermi theory we first examine a very crude model for the atom, which allows for a correct asymptotic scaling of the atomic energy with the atomic number $Z \rightarrow \infty$.

## EXERCISE 14. A crude model for the energy of heavy

 atomsDevelop a qualitative theory for the energy $E(Z)$ of an atom with atomic number $Z$. The model presents the energy as a function of a fictitious "atomic radius" $R$ :

$$
E(R, Z)=T+E_{e n}+E_{e e}
$$

and once the function is built $R$ is determined by minimizing the energy. Here are the assumptions for the various terms of the energy
(1) The e-e repulsive interaction energy $E_{e e}$ is estimated as the classical classically. The $Z$ electrons are modeled by a homogeneous spherical cloud of total charge $-Z e$ in a spherical volume $V$ of atomic radius $R$.
(2) The kinetic energy $T$ of the electrons is estimated, respecting the Pauli principle, under the assumption that each pair is in a separate non-overlapping cubic box of volume $a^{3}=V / 2 Z$. The $Z / 2$ boxes are distributed equally within the spherical volume $V$. If the sphere is large the inconsistency of fitting boxes in the sphere will be small.
(3) The nucleus is a positive point charge $+Z e$ located within the center of the sphere.
Find the scaling of the energy with the only variable imposed parameter $Z$. Show the model leads to a unphysical assumption that the radius $Z$ decreases as $Z$ grows.

Solution. The electron-nuclear energy is the energy to bring a nucleus from infinity into the center of the electronic charge distribution. The electric field due to the electrons is

$$
E_{r}(r)=\frac{Z e}{4 \pi \epsilon_{0}} \frac{Q(r)}{r^{2}}
$$

Where $Q(r)$ is the amount of electron charge inside the sphere of radius $r$. Assuming this positive charge distribution is homogeneous we have:

$$
Q(r)=-Z e\left[\left(\frac{r}{R}\right)^{3} \theta(R-r)+\theta(r-R)\right]
$$

where $\theta(x)=1$ for $x>0$ and 0 otherwise. The work done for bringing the positive nucleus from infinity to the center of the
sphere is then the line integral over the electric field is:

$$
\begin{aligned}
E_{e n} & =-(Z e) \int_{\infty}^{0} E_{r}(r) d r \\
& =-\frac{e^{2} Z^{2}}{4 \pi \epsilon_{0}}\left[\int_{0}^{R} \frac{1}{r^{2}}\left(\frac{r}{R}\right)^{3} d r+\int_{R}^{\infty} \frac{1}{r^{2}} d r\right] \\
& =-\frac{B_{1}}{R}
\end{aligned}
$$

where in atomic units $e^{2} / 4 \pi \epsilon_{0}=E_{h} a_{0}$ :

$$
\begin{aligned}
B_{1} & =\frac{e^{2}}{4 \pi \epsilon_{0}} \frac{3}{2} Z^{2} \\
& =1.50 E_{h} a_{0} Z^{2}
\end{aligned}
$$

The kinetic energy $T$ at the groundstate is quantum and take into account the Pauli principle dictating a different quantum state for electron of different spins. Hence, assume each pair of electrons occupies a unique small cube within the atomic sphere. The cube volume is $V_{1}=\frac{V}{Z / 2}=\frac{4 \pi}{3} \frac{R^{3}}{Z / 2}$. Hence, the cube side length is $a=\left(\frac{8 \pi}{3} \frac{1}{Z}\right)^{1 / 3} R$. The quantum energy of a free particle in a cube is $T_{1}=3 \times \frac{\hbar^{2} \pi^{2}}{2 m a^{2}}$, multiplying by $Z / 2$ pairs of electrons we find in atomic units:

$$
\begin{aligned}
T & =\frac{A}{R^{2}} \\
A & =1.8 Z^{5 / 3} E_{h} a_{0}^{2}
\end{aligned}
$$

Finally, $E_{e e}$ is the energy of Coulomb repulsion between the electrons. Here we can take a model of a uniform sphere of charge $n=\frac{Z}{V}$ where $V=\frac{4 \pi}{3} R^{3}$ is the atomic volume. We do this by repeatedly bringing a shell of radius $r$ and charge $n \times 4 \pi r^{2} d r$ from infinity to its location in the sphere, one at a
time, we have:

$$
\begin{aligned}
E_{e e} & =\frac{e^{2}}{4 \pi \epsilon_{0}} \int_{0}^{R} n \frac{4 \pi}{3} r^{3} \times \frac{1}{r} \times n 4 \pi r^{2} d r \\
& =\frac{e^{2}}{4 \pi \epsilon_{0}} \frac{(4 \pi)^{2}}{3} n^{2} \frac{R^{5}}{5} \\
& =\frac{B_{2}}{R}
\end{aligned}
$$

where

$$
\begin{aligned}
B_{2} & =\frac{3}{5} \frac{e^{2}}{4 \pi \epsilon_{0}} Z^{2} \\
& =0.6 Z^{2} E_{h} a_{0}
\end{aligned}
$$

The total energy becomes a function of the atom radius $R$ and $Z$ :

$$
E(R)=\frac{A}{R^{2}}-\frac{B}{R}
$$

Where $B=B_{1}-B_{2}=0.9 Z^{2} E_{h} a_{0}$. We now select the radius $R$ by assuming that it minimizes the energy. Solving $E^{\prime}\left(R_{*}\right)=0$ gives $R_{*}=\frac{2 A}{B}$ and the atom minimal energy is $E=-\frac{B^{2}}{4 A}$. Evaluating this in atomic units gives the following relation involving the atomic number:

$$
\begin{aligned}
E & =-0.11 \times Z^{7 / 3} E_{h} \\
R_{*} & =4.0 \times Z^{-1 / 3} a_{0}
\end{aligned}
$$

Since the H atom has radius $R=1 a_{0}$ and energy $E=-0.5 E_{h}$ this result is not very good for this small atom limit. However, the scaling of the energy with $Z$ turns out to be exact when $Z$ as it increases but the exact prefactor is much higher, close to -0.7687 instead of -0.11 .

### 2.1. Basic concepts in Thomas-Fermi Theory

In the early days of quantum mechanics there was no practical way of using the Schrödinger equation to determine the
electronic structure of many-electron systems such as heavy atoms. A simple, albeit approximate method was in need and supplied separately by Thomas. Their theory can be thought of as an application of "density functional theory", although the latter appeared only 40 years later.

The TF theory focuses on the 1-particle density function $n(\boldsymbol{r})$, giving the number of electrons per volume at point $\boldsymbol{r}$, defined in Eq. (2.0.4). They assumed that the density minimizes an "energy density functional" built from 3 terms:

$$
\begin{align*}
E_{T F}[n] & =T_{T F}[n]+\int v_{e x t}(\boldsymbol{r}) n(\boldsymbol{r}) d^{3} r  \tag{2.1.1}\\
& +\frac{1}{2} \iint n\left(\boldsymbol{r}_{1}\right) n\left(\boldsymbol{r}_{2}\right) u_{C}\left(r_{12}\right) d^{3} r_{1} d^{3} r_{2} .
\end{align*}
$$

where $r_{12}=\left|\boldsymbol{r}_{1}-\boldsymbol{r}_{2}\right|$ and

$$
\begin{equation*}
u_{C}\left(r_{12}\right)=\frac{e^{2}}{4 \pi \epsilon_{0}} \frac{1}{r_{12}} \tag{2.1.2}
\end{equation*}
$$

is the Coulomb force potential. The first term in Eq. (2.1.1) is the density-dependent kinetic energy functional which will be discussed below. The second term is the potential energy of the external potential operating on the electrons while the last term is the classical repulsive Coulomb energy.

The density $n(\boldsymbol{r})$ that minimizes the TF energy functional must be performed under the constraint that the density include the specified number $N$ of electrons of the atom or molecule:

$$
\int n(\boldsymbol{r}) d^{3} r=N
$$

The method for minimizing the constrained energy is discussed in Appendix C. Meanwhile we discuss what should we take as the "kinetic energy (KE) functional". We need to assign a kinetic energy to any density $n(\boldsymbol{r})$ we are given. We must take into account the Pauli principle, namely the Fermion and quantum nature of the electrons. The Thomas-Fermi approach is to
assume a simple local form for the KE functional:

$$
\begin{equation*}
T_{T F}[n]=\int t_{S}(n(\boldsymbol{r})) n(\boldsymbol{r}) d^{3} r . \tag{2.1.3}
\end{equation*}
$$

Where the meaning of the function $t_{S}(n)$ is elucidated by searching for a system where the form is valid. This is the system of $N_{e}$ electrons in a box of volume $V$ with periodic boundaries, which due to symmetry must have a homogeneous density characterized by a number: $n(\boldsymbol{r})=n=\frac{N_{e}}{V}$. Since we do not want the size of the system $(V)$ to play a role, we look at the "thermodynamic limit", namely the limit of $V \rightarrow \infty$ and $N_{e} \rightarrow \infty$ (but $n=N_{e} / V$ stays finite). The system is often called the non-interacting Homogeneous Electron Gas (HEG) or sometimes Uniform Electron Gas. In this case the integrand of Eq. (2.1.3) is independent of $r$ and the integral can be readily performed:

$$
T_{T F}=t_{S}(n) N_{e} .
$$

From this relation we see what is the meaning of the function $t_{S}(n)=T_{T F} / N_{e}$ : it is the average kinetic energy per electron in a HEG. Note that here too, in the thermodynamic limit both $T_{T F}$ and $N_{e}$ go to infinity but the ration $t_{S}$ remains. We will now show that this indeed is the case by calculating the kinetic energy of a non-interacting HEG.

ExERCISE 15. The eigenstates of a particle on a ring of radius $a$ are indexed by the integers $m=0, \pm 1, \pm 2, \ldots$ and given as $\psi_{m}(\phi)=e^{i m \phi}$, the energy eigenvalues are $E_{m}=\frac{\hbar^{2} m^{2}}{2 I}$ where $I=m_{e} a^{2}$. Show $\psi_{m}(\phi)$ is also an eigenstate of the angular momentum $L_{z}=\frac{\hbar}{i} \frac{d}{d \phi}$ and that all eigenvalues are integer multiples of the smallest momentum $\hbar$. Map the ring onto a coordinate $x=L \phi$. Show that the wave functions are now $\psi_{k_{n}}(x)=e^{i k_{n} x}$ where $k_{n}=n \frac{2 \pi}{L}, n=0, \pm 1, \pm 2, \ldots$ and the kinetic energy eigenvalues are $E_{n}=\frac{\hbar^{2} k_{n}^{2}}{2 m_{e}}$.

Solution. We compute the overlap integral for $k \neq k^{\prime}$ as follows:

$$
\begin{aligned}
\left\langle\psi_{k} \mid \psi_{k^{\prime}}\right\rangle & =\int_{0}^{L} \psi_{k}(x)^{*} \psi_{k^{\prime}}(x) d x \\
& =\frac{1}{L} \int_{0}^{L} e^{-i k x} e^{i k^{\prime} x} d x \\
& =\frac{e^{i\left(k^{\prime}-k\right) L}-1}{i L\left(k^{\prime}-k\right)} \\
& =0
\end{aligned}
$$

where the last equality is due to the quantization $k=\frac{2 \pi}{L} n$, where $n$ is an integer, and $e^{i 2 \pi n}-1=0$. When $k=k^{\prime}$ we have $\left\langle\psi_{k} \mid \psi_{k^{\prime}}\right\rangle=1$. Finally

$$
\begin{gathered}
\hat{p}_{x}\left|\psi_{k}\right\rangle=\frac{\hbar}{i} \frac{d}{d x}\left|\psi_{k}\right\rangle=\left|\psi_{k}\right\rangle \hbar k \\
\hat{T}\left|\psi_{k}\right\rangle=\frac{\hat{p}^{2}}{2 m_{e}}\left|\psi_{k}\right\rangle=\left|\psi_{k}\right\rangle \frac{\hbar^{2} k^{2}}{2 m_{e}}
\end{gathered}
$$

where $m_{e}$ is the electron mass.
Consider a homogeneous gas of $N_{e}$ uncharged electrons, non-interacting, put in a cubic cell of length $L$ with electron density uniform: $n=N_{e} / V=N_{e} / L^{3}$. We will be interested in the "thermodynamic limit", meaning in the case where both $N_{e}$ and $V$ are taken to infinity while $n$ which is their ratio stays constant. In this limit many intricate details get washed out as will be seen below. The fact that the electrons are noninteracting means that we can compute the levels of each electron separately and then just let the $N_{e}$ electrons occupy these levels, two by two, starting from the lowest energy and increasing. This imposes the Pauli principle. The periodic boundary
conditions mean that the wave functions of each electron is periodic in the box. Hence by Fourier's theorem any wave function of such a system is expressible as:

$$
\psi(\boldsymbol{r})=\sum_{\boldsymbol{k}} a_{\boldsymbol{k}} \psi_{\boldsymbol{k}}(\boldsymbol{r}),
$$

namely, a linear combination of plane-waves $\psi_{k}(\boldsymbol{r})=\frac{e^{i \boldsymbol{k} \cdot \boldsymbol{r}}}{\sqrt{V}}$ where:

$$
\begin{equation*}
\boldsymbol{k}=\left(k_{x}, k_{y}, k_{z}\right)=\frac{2 \pi}{L}\left(n_{x}, n_{y}, n_{z}\right) \tag{2.1.4}
\end{equation*}
$$

are quantized set of wave-vectors and $n_{x / y / z}$ are integers.
Exercise 16. Show that in the thermodynamic limit the sum of a function $f(k)$ over the discrete $k_{n}=\frac{2 \pi}{L} n$ values $\frac{1}{L} \sum_{k} f(k)$ is equal to integral over $k \frac{1}{2 \pi} \int f(k) d k$.

Solution. We start from the discrete approximation to the integral: $\int_{-\infty}^{\infty} f(k) d k \approx \Delta k \sum_{n=-\infty}^{\infty} f\left(k_{n}\right)$ where $k_{n}=\Delta k n$ and $\Delta k$ is any small integral. The approximation improves as $\Delta k$ decreases. The box of length $L$ in $x$-space defines a mesh in $k$-space with spacing $\Delta k=\frac{2 \pi}{L}$ and hence $\frac{1}{L} \sum_{n} f\left(k_{n}\right) \approx$ $\frac{1}{2 \pi} \int_{-\infty}^{\infty} f(k) d k$. In 3D this equation is straightforwardly generalized to $\frac{1}{V} \sum_{k} f(\boldsymbol{k}) \approx \frac{1}{(2 \pi)^{3}} \iiint f(\boldsymbol{k}) d^{3} k$. This approximation becomes exact as $V$ grows to infinity. In all cases we consider below the integrand function $f(\boldsymbol{k})$ is spherically symmetric, i.e. $f(\boldsymbol{k})=f(k)$, so the 3D integral becomes a spherical integral $\int_{0}^{\infty} 4 \pi k^{2} d k$, giving the following useful formula

$$
\begin{equation*}
\frac{1}{V} \sum_{\boldsymbol{k}} f(\boldsymbol{k}) \approx \frac{1}{2 \pi^{2}} \int_{0}^{\infty} f(k) k^{2} d k \tag{2.1.5}
\end{equation*}
$$

For the HEG we want to occupy the lowest energy states by two electrons each and once we exhaust all electrons the other states are vacant. Suppose the energy of the highest occupied state is $\mu$, this is called the Fermi energy. Then the occupation
function is $p_{k}=\theta\left(\mu-\frac{\hbar^{2} \boldsymbol{k}^{2}}{2 m_{e}}\right)$ assigns to each momentum state $k$ either 0 (empty) or 1 (full), according to whether the kinetic energy of that state $\frac{\hbar^{2} k^{2}}{2 m_{e}}$ is less than and energy $\mu$ or not. Since there are 2 electrons in each occupied state, $2 \sum_{k} p_{k}=N_{e}$. Let us set the "Fermi" wavenumber $k_{F}: \mu=\frac{\hbar^{2} k_{F}^{2}}{2 m_{e}}$, then $p_{\boldsymbol{k}}=$ $\theta\left(k_{F}-|\boldsymbol{k}|\right)$ and we must have $2 \sum_{\boldsymbol{k}} \theta\left(k_{F}-|\boldsymbol{k}|\right)=N_{e}$. Sums on $\boldsymbol{k}$ must be divided by $V$ to converge in the thermodynamic limit (Eq. 2.1.5) we find

$$
\begin{equation*}
2 \frac{1}{V} \sum_{\boldsymbol{k}} \theta\left(k_{F}-|\boldsymbol{k}|\right)=\frac{N_{e}}{V}=n \tag{2.1.6}
\end{equation*}
$$

From Eq. (2.1.5) the sum $2 \frac{1}{V} \sum_{k} \theta\left(k_{F}-|\boldsymbol{k}|\right)$ is converted to the integral $\frac{1}{\pi^{2}} \int_{0}^{k_{F}} k^{2} d k=\frac{1}{3 \pi^{2}} k_{F}^{3}$, giving the highest occupied wavenumber (momentum) in terms of the density in the noninteracting HEG:

$$
\begin{equation*}
n=\frac{k_{F}^{3}}{3 \pi^{2}} \Longleftrightarrow k_{F}=\left(3 \pi^{2} n\right)^{1 / 3} \tag{2.1.7}
\end{equation*}
$$

The kinetic energy per electron is the sum of the kinetic energies of the occupied states times 2: $T_{T F}=2 \sum_{k} \theta\left(k_{F}-|k|\right) \frac{\hbar^{2} k^{2}}{2 m_{e}}$. Hence the kinetic energy per electron is

$$
t_{S}=\frac{T_{T F}}{N_{e}}=\frac{2}{n V} \sum_{k} \theta\left(k_{F}-|k|\right) \frac{\hbar^{2} k^{2}}{2 m_{e}},
$$

where we used the relation $N_{e}=n V$. Replacing the sum by an integral (Eq. 2.1.5) and the density $n$ by $\frac{k_{F}^{3}}{3 \pi^{2}}$ (Eq. (2.1.7)) we find:

$$
\begin{equation*}
t_{S}=\frac{3}{5} \frac{\hbar^{2} k_{F}^{2}}{2 m_{e}}=\frac{3}{5} \mu . \tag{2.1.8}
\end{equation*}
$$

The average kinetic energy is $3 / 5$ of the maximal kinetic energy - the Fermi energy. In terms of the density, we use Eq. 2.1.7
and obtain:

$$
\begin{equation*}
t_{S}(n)=\frac{3}{5} C n^{2 / 3} \tag{2.1.9}
\end{equation*}
$$

where

$$
C=\frac{\hbar^{2}\left(3 \pi^{2}\right)^{2 / 3}}{2 m_{e}}=4.785 \frac{\hbar^{2}}{m_{e}}
$$

Going back to the TF kinetic energy functional, Eq. (2.1.3) we find:

$$
\begin{equation*}
T_{T F}[n]=\frac{3}{5} C \int n(\boldsymbol{r})^{5 / 3} d^{3} r \tag{2.1.10}
\end{equation*}
$$

EXercise 17. The Thomas-Fermi functional for the hydrogenlike atom.
(1) Derive the TF kinetic energy functional for a "spinpolarized HEG". That is, do not assume that there are 2 electrons in each k-state (the "spin-unpolarized" case) but instead, that all spins are up and so there is only one electron per k-state.
(2) Since the electron in a hydrogen-like atom is "spinpolarized", use the Thomas-Fermi KE functional derived in the previous question and compare its estimate of the kinetic energy of the electron in a hydrogenlike atom to the exact value. Using the exact kinetic energy in the hydrogen atom (which, using the virial theorem, is just the negative of the ground state energy), assess the quality of the result as a function of the nucleus charge $Z$.
The TF energy functional is thus
(2.1.11) $\quad E_{T F}[n]=\frac{3}{5} C \int n(\boldsymbol{r})^{5 / 3} d^{3} r+\int v_{e x t}(\boldsymbol{r}) n(\boldsymbol{r}) d^{3} r$

$$
+\frac{1}{2} \iint n\left(\boldsymbol{r}_{1}\right) n\left(\boldsymbol{r}_{2}\right) u_{C}\left(r_{12}\right) d^{3} r_{1} d^{3} r_{2} .
$$

For atoms and molecules, the external potential is a result of a positive charge distribution $n_{+}(\boldsymbol{r})$ :

$$
v_{e x t}\left(\boldsymbol{r}_{1}\right)=-\int n_{+}\left(\boldsymbol{r}_{2}\right) u_{C}\left(r_{12}\right) d^{3} r_{2} .
$$

In this case one can write the TF functional as:
$E_{T F}[n]=\frac{3}{5} C \int n(\boldsymbol{r})^{5 / 3} d^{3} r+\frac{1}{2} \iint \rho\left(\boldsymbol{r}_{1}\right) u_{C}\left(r_{12}\right) \rho\left(\boldsymbol{r}_{2}\right) d^{3} r_{1} d^{3} r_{2}$.
where

$$
\begin{equation*}
\rho(\boldsymbol{r})=n_{+}(\boldsymbol{r})-n(\boldsymbol{r}) \tag{2.1.13}
\end{equation*}
$$

is the total charge density.

### 2.2. The virial theorem for Thomas-Fermi theory

The Thomas-Fermi theory enjoys some interesting scaling laws. Some of them, like the one we study here turn out to be valid in the exact Schrödinger equation. Others are unique to the theory and are correct only for infinitely heavy atoms. The virial theorem in quantum mechanics is studied in detail in chapter XXX. Here we give only the details pertinent to TF theory. We consider the TF functional for an atom, evaluated at the density $n(\boldsymbol{r})$, which we assume is its minimum:

$$
\begin{equation*}
E_{T F}[n]=T_{T F}[n]+U[n] \tag{2.2.1}
\end{equation*}
$$

where

$$
U[n]=-Z \kappa \int n(\boldsymbol{r}) \frac{1}{r} d^{3} r+\frac{1}{2} \iint n\left(\boldsymbol{r}_{1}\right) n\left(\boldsymbol{r}_{2}\right) u_{C}\left(r_{12}\right) d^{3} r_{1} d^{3} r_{2} .
$$

Let us assume that the electron density $n_{*}(\boldsymbol{r})$ is minimizes the above functional, subject to the constraint $\int n(r) d^{3} r=$ $N_{e}$. Let us now scale this electronic density in the following way, using the scaling parameter $\lambda>0$ :

$$
n_{\lambda}(\boldsymbol{r})=\lambda^{3} n_{*}(\lambda \boldsymbol{r}) .
$$

Clearly, $\int n_{*}(r) d^{3} r=\int n_{\lambda}(r) d^{3} r$, so both charge distributions ascribe to the same number of electrons. Similarly it is straightforward to check that:

$$
\begin{gathered}
T_{T F}\left[n_{\lambda}\right]=\lambda^{2} T_{T F}\left[n_{*}\right] \\
U\left[n_{\lambda}\right]=\lambda U\left[n_{*}\right]
\end{gathered}
$$

The TF energy is then

$$
E_{T F}\left[n_{\lambda}\right]=\lambda^{2} T_{T F}\left[n_{*}\right]+\lambda U\left[n_{*}\right]
$$

The scaling parameter $\lambda_{*}$ that minimizes the energy is obtained from the condition $\frac{d}{d \lambda}\left(E_{T F}\left[n_{\lambda}\right]\right)_{\lambda=\lambda_{*}}=0$. We know that the result should be $\lambda_{*}=1$. Thus:

$$
2 T_{T F}\left[n_{*}\right]+U\left[n_{*}\right]=0
$$

Using Eq. (2.2.1) we find the virial theorem for TF theory:

$$
\begin{equation*}
E_{T F}\left[n_{*}\right]=-T_{T F}\left[n_{*}\right]=\frac{1}{2} U\left[n_{*}\right] . \tag{2.2.2}
\end{equation*}
$$

Interestingly, despite the fact that the TF theory for an atom is a gross approximation it obeys the virial relation which is identical in form to the exact quantum mechanical virial theorem.

### 2.3. Minimization of Thomas-Fermi energy

The TF approach assumes that the ground-state electron density $n(\boldsymbol{r})$ should be determined by minimizing $E_{T F}[n]$, among all densities having the required number of electrons: $\int n(\boldsymbol{r}) d^{3} r=$ $N_{e}$. For imposing this constraint we introduce a Lagrangian (see XX):

$$
\begin{equation*}
L[n]=E_{T F}[n]-\mu\left(\int n(\boldsymbol{r}) d^{3} r-N_{e}\right) \tag{2.3.1}
\end{equation*}
$$

The derivative with respect to $n(\boldsymbol{r})$, at the solution density $n_{*}(\boldsymbol{r})$ gives the Euler-Lagrange equation for TF theory:

$$
\begin{equation*}
0=\frac{\delta L}{\delta n(\boldsymbol{r})}=\left.\frac{\delta E_{T F}}{\delta n(\boldsymbol{r})}\right|_{n_{*}(\boldsymbol{r})}-\mu \tag{2.3.2}
\end{equation*}
$$

We denote the minimal value of the TF functional by:

$$
\begin{equation*}
E_{T F *} \equiv E_{T F}\left[n_{*}\right] \tag{2.3.3}
\end{equation*}
$$

From Lagrange's theory XXX, the Lagrange constant $\mu$ is the derivative of $E_{T F *}$ with respect to the number of electrons:

$$
\begin{equation*}
\frac{\partial E_{T F *}}{\partial N_{e}}=\mu . \tag{2.3.4}
\end{equation*}
$$

Hence $\mu$ is the chemical potential. Applying Eq. 2.3.2to the TF energy functional of Eq. 2.1.12 we find:
(2.3.5)

$$
C n_{*}(\boldsymbol{r})^{2 / 3}=\mu+\int\left(n_{+}\left(\boldsymbol{r}^{\prime}\right)-n_{*}\left(\boldsymbol{r}^{\prime}\right)\right) u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|\right) d^{3} r^{\prime}
$$

This is the integral Thomas Fermi equation for $n(\boldsymbol{r})$. We solve it and adjust $\mu$ so that $\int n_{*}(\boldsymbol{r}) d^{3} r=N_{e}$.

Another way to define the electric potential $\phi(\boldsymbol{r})$ (shifted by $\mu$ ) as:

$$
\phi(\boldsymbol{r})=\mu+\int\left(n_{+}\left(\boldsymbol{r}^{\prime}\right)-n_{*}\left(\boldsymbol{r}^{\prime}\right)\right) u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|\right) d^{3} r^{\prime} .
$$

And then :

$$
\begin{equation*}
\phi(\boldsymbol{r})=C n_{*}(\boldsymbol{r})^{2 / 3} . \tag{2.3.6}
\end{equation*}
$$

We can obtain a differential equation for $\phi(\boldsymbol{r})$ as follows. The Laplacian of $\phi$ is equal to $4 \pi \kappa\left(n_{*}\left(\boldsymbol{r}^{\prime}\right)-n_{+}\left(\boldsymbol{r}^{\prime}\right)\right)$, hence, using Eq. (2.3.6):

$$
\begin{equation*}
\nabla^{2} \phi(\boldsymbol{r})=4 \pi \kappa\left(\left(\frac{\phi(\boldsymbol{r})}{C}\right)^{3 / 2}-n_{+}(\boldsymbol{r})\right) \tag{2.3.7}
\end{equation*}
$$

This is the differential TF equation for $\phi(\boldsymbol{r})$ from which $n_{*}(\boldsymbol{r})$ is obtained through Eq. 2.3.6. The chemical potential enters this equation as a boundary condition: $\phi(\boldsymbol{r}) \rightarrow \mu$, i.e. $\mu$ is simply an unimportant additive constant to $\phi$ that is designed to ensure that $\int n_{*}(\boldsymbol{r}) d^{3} r=N_{e}$.

### 2.4. Thomas-Fermi theory of atoms

Application of TF theory for atoms explains the main features of stability of atoms. The theory becomes ever more accurate as the atomic number $Z$ increases. We consider first a H atom in the origin, the TF equation becomes, in spherical coordinates,

$$
\frac{1}{r}(r \phi(r))^{\prime \prime}=4 \pi \kappa\left(\left(\frac{\phi(r)}{C}\right)^{3 / 2}-\frac{Z \delta(r)}{4 \pi r^{2}}\right)
$$

Which upon multiplication by $r^{2}$ becomes:

$$
r(r \phi(r))^{\prime \prime}=4 \pi \kappa r^{2}\left(\frac{\phi(r)}{C}\right)^{3 / 2}-\kappa Z \delta(r)
$$

This equation is not convenient because of the delta-function source term in the origin. We note first that for $r>0$ the equation is

$$
\begin{equation*}
r(r \phi(r))^{\prime \prime}=\frac{4 \pi \kappa}{C^{3 / 2}} r^{2} \phi(r)^{3 / 2} \tag{2.4.1}
\end{equation*}
$$

and the delta function can be converted to a boundary condition. Indeed, integrating from 0 to (small) $r$ gives:

$$
(r \phi(r))^{\prime} r-r \phi(r)=\kappa \int_{0}^{r}\left(\frac{\phi\left(r^{\prime}\right)}{C}\right)^{3 / 2} 4 \pi r^{\prime 2} d r^{\prime}-\kappa Z
$$

where we used the fact that $r(r \phi(r))^{\prime \prime}=\left(r(r \phi(r))^{\prime}\right)^{\prime}-(r \phi(r))^{\prime}$. We observe that taking the condition

$$
\begin{equation*}
\lim _{r \rightarrow 0} r \phi(r)=\kappa Z \tag{2.4.2}
\end{equation*}
$$

makes both sides of the equation equal in the limit $r \rightarrow 0$. Hence, what we need to solve is Eq. 2.4.1 under the boundary condition Eq. 2.4.2 and $\lim _{r \rightarrow \infty} \phi(r)=0$.

Fermi simplified this equation by changing variable $x=\alpha r$ and:

$$
\begin{equation*}
r \phi(r)=\kappa \psi_{Z}(x) \tag{2.4.3}
\end{equation*}
$$

where

$$
\begin{equation*}
\alpha=(4 \pi)^{2 / 3} \frac{\kappa}{C} . \tag{2.4.4}
\end{equation*}
$$

This gives the following "clean" equation by Fermi to be solved:

$$
\begin{equation*}
\psi_{Z}^{\prime \prime}(x)=\frac{\psi_{Z}(x)^{3 / 2}}{\sqrt{x}} \tag{2.4.5}
\end{equation*}
$$

$$
\begin{align*}
\lim _{x \rightarrow 0} \psi_{Z}(x) & =Z  \tag{2.4.6}\\
\lim _{x \rightarrow \infty} \psi(x) & =0
\end{align*}
$$

This equation assumes that $\psi_{Z}(x)$ is never negative, in accordance with the definition Eq. (2.3.6). We can show that one needs to solve the Fermi equation only for $Z=1$. The other $Z \mathrm{~s}$ are obtained by a simple transformation:

Exercise 18. The Fermi equation for atom with charge $Z$, the solution for $Z=1$ generates the solution for all $Z$ :

$$
\begin{equation*}
\psi_{Z}(x)=Z \psi_{1}\left(Z^{1 / 3} x\right) \tag{2.4.7}
\end{equation*}
$$

Show that $\psi_{Z}(x)$ obeys Eq. (2.4.5) and the boundary condition

$$
\lim _{x \rightarrow 0} \psi_{Z}(x)=Z
$$

Solution. The second derivative of $\psi_{Z}(x), \psi_{Z}^{\prime \prime}(x)=Z^{5 / 3} \frac{\psi_{1}\left(Z^{1 / 3} x\right)^{3 / 2}}{\sqrt{Z^{1 / 3} x}}$ is easily shown to equal to $\frac{\psi_{Z}(x)^{3 / 2}}{\sqrt{x}}$, establishing Eq. (2.4.5). The boundary condition too is a direct result if the $x \rightarrow 0$ boundary condition of $\psi(x)$.

Exercise 19. Assuming $\psi_{Z}(x)=\beta x^{\alpha}$ for large $x$, show that:

$$
\begin{equation*}
\lim _{x \rightarrow \infty} x^{3} \psi_{Z}(x)=144 \tag{2.4.8}
\end{equation*}
$$

What is the $r$-dependence of the long range density in TF theory of the atom?

Solution. Plug $\psi_{Z}(x)=\beta x^{\alpha}$ into Eq. (2.4.5), and find $\beta \alpha(\alpha-1) x^{\alpha-2}=\beta^{3 / 2} x^{3 \alpha / 2-1 / 2}$. From the $x$-dependence we find $\alpha=-3$ and then $\beta^{3 / 2}=\beta(-3)(-3-1) \rightarrow \beta=144$; hence Eq. (2.4.8). As for the density dependence on $r$ : From Eqs. (2.3.6) and (2.4.3) we have: $n(r) \propto \phi(r)^{3 / 2} \propto\left(\psi_{Z}(\alpha r) / r\right)^{3 / 2} \propto$ $\left(\frac{1}{r^{4}}\right)^{3 / 2}=\frac{1}{r^{6}}$.

EXercise 20. Show that the kinetic energy of the electrons is proportional to $\int \frac{\psi_{Z}(x)^{5 / 2}}{\sqrt{x}} d x$. Find the proportionality constant.

Solution. From Eqs. (2.1.10) and (2.3.6), the kinetic energy is $T_{T F}=\frac{3}{5} C \int\left(\frac{\phi(\boldsymbol{r})}{C}\right)^{5 / 2} d^{3} r$ which in spherical coordinates gives: $T_{T F}=\frac{3}{5} 4 \pi C^{-3 / 2} \int_{0}^{\infty}(r \phi(r))^{5 / 2} \frac{d r}{\sqrt{r}}$. We change integration variable $\alpha x=r$ and use the definition Eq. (2.4.3) and we find after some manipulations:

$$
\begin{align*}
T_{T F} & =\frac{3}{5} C^{-3 / 2} 4 \pi \sqrt{\alpha} \int_{0}^{\infty}\left(\kappa \psi_{Z}(x)\right)^{5 / 2} \frac{d x}{\sqrt{x}} \\
& =\frac{3}{5} C^{-1}(4 \pi)^{2 / 3} \kappa^{2} \int_{0}^{\infty} \psi_{Z}(x)^{5 / 2} \frac{d x}{\sqrt{x}} \\
& =0.67776 E_{h} \int_{0}^{\infty} \psi_{Z}(x)^{5 / 2} \frac{d x}{\sqrt{x}} \tag{2.4.9}
\end{align*}
$$

Now use Eq. (2.4.7) so: $T_{T F}(Z)=0.67776 E_{h} \frac{Z^{5 / 2}}{\sqrt{Z^{1 / 3}}} \int \frac{\left(\psi_{1}(y)\right)^{5 / 2}}{\sqrt{y}} d y$ which leads to $T_{T F}(Z)=T_{T F}(1) Z^{7 / 3}$. From the virial theorem $E=-T_{T F}$ and thus $E_{Z}=E_{1} Z^{7 / 3}$. There is no analytical way to compute $E_{1}$. A numerical calculation gives: $E_{1}=-0.7688 E_{h}$. Hence the energy of atomic number $Z$ is

$$
\begin{equation*}
E_{Z}=-0.7688 E_{h} Z^{7 / 3} . \tag{2.4.10}
\end{equation*}
$$



Figure 2.4.1. The density functional (within the local density approximation) energy of rare gas atoms vs the Thomas-Fermi value of Eq. 2.4.10. It should be noted that the HF energies are almost indistinguishable from the LDA values on the scale of this graph.

This should be compared with the exact energy of the hydrogen atom, which is $-0.5 E_{h}$. The TF energy is much lower. But for larger atoms the estimate improves. In fact Lieb proved that the TF energy approaches the exact energy of atoms as $Z$ grows. [4] This is demonstrated in Fig. (2.4.1) where the TF energy is compared to higher-level ab initio theories (the local density approximation for DFT [2]). In that respect TF theory is an "exact" theory of the many electron atom.

EXercise 21. Atoms shrink in size as $Z$ grows: show that the atomic radius is proportional to $Z^{-1 / 3}$.

Solution. The radius of the neutral atom can be defined by the moment over the density

$$
R=\frac{1}{Z} \int_{0}^{\infty} n(r) r \times 4 \pi r^{2} d r
$$

where from Eq. (2.4.3) $n(r)=\left(\frac{\phi(r)}{C}\right)^{3 / 2}$. Now, from Eqs. (2.4.3), (2.4.4) and (2.4.7):

$$
\begin{aligned}
r \phi(r) & =\kappa \psi_{Z}(x) \\
& =Z \kappa \psi_{1}(y)
\end{aligned}
$$

where $y=x Z^{1 / 3}=\alpha Z^{1 / 3} r$ and $\alpha=(4 \pi)^{2 / 3} \frac{\kappa}{C}$. Hence:

$$
\begin{aligned}
n(r) r^{3} d r & =\left(\frac{r \phi(r)}{C}\right)^{3 / 2} r^{3 / 2} d r \\
& =\left(\frac{Z \kappa \psi_{1}(y)}{C}\right)^{3 / 2} \frac{y^{3 / 2} d y}{\left(\alpha Z^{1 / 3}\right)^{5 / 2}}
\end{aligned}
$$

so: $R_{Z}=Z^{-1 / 3} R_{1}$, where $R_{1}=(4 \pi)^{-2 / 3} \frac{C}{\kappa} \int_{0}^{\infty}\left[y \psi_{1}(y)\right]^{3 / 2} d y$

### 2.5. Thomas-Fermi theory cannot explain chemical bonds

What is the physical meaning of $\phi(\boldsymbol{r})$ of Eq. 2.3.6? Suppose we change the positive charge at a point $\boldsymbol{r}$ by $\delta n_{+}(\boldsymbol{r})$ simultaneously changing $N_{e}$ by the same amount $\delta N_{e}=\int \delta n_{+}(\boldsymbol{r}) d^{3} r$, so that the total charge of the system is conserved. What is the change in the TF energy $E_{T F *}$ of Eq. 2.3.3? There are two contributions. First, from the change $\delta N_{e}$ there is a change due to the chemical potential, Eq. 2.3.4. Then, from the change in $\delta n_{+}(\boldsymbol{r})$ there is a change by taking the derivative of the Coulomb term of the TF energy functional in Eq. 2.1.12: $\left(\delta E_{T F *} / \delta n_{+}(\boldsymbol{r})\right)_{2}=\int\left(n_{+}\left(\boldsymbol{r}_{2}\right)-n\left(\boldsymbol{r}_{2}\right)\right) u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}_{2}\right|\right) d^{3} r_{2}$. Adding the two contributions, taking into account Eq. 2.3 .5 we see
that the total change is

$$
\left.\frac{\delta E_{T F *}}{\delta n_{+}(\boldsymbol{r})}\right|_{\text {const total charge }}=\phi(\boldsymbol{r}) \delta n_{+}(\boldsymbol{r})
$$

Hence, $\phi(\boldsymbol{r})$ describes the rate of change of the TF energy when a positive charge perturbation is applied keeping the total amount of charge constant. Notice that from the very definition of $\phi(\boldsymbol{r})$,Eq. 2.3.6 the potential is positive everywhere, meaning that the TF energy increases whenever we add a neutral system to our system. Hence whenever two atoms approach each other the energy grows: any kind of chemical bonding is precluded by TF theory! This result was discovered by Edward Teller[6].

### 2.6. Linear response theory and Thomas-Fermi screening

We now examine the electrostatic potential which develops when a positive charge density $\delta n_{+}(\boldsymbol{r})=\delta q \delta(\boldsymbol{r})$ is imposed on the homogeneous electron gas (HEG). In vacuum the electrostatic potential by the added charge would be Coulombic, namely $\delta \phi(\boldsymbol{r})=\frac{\kappa \delta q}{r}$. However in the HEG, which is a model for a simple metal, we expect that an added positive (negative) charge will attract (repel) the negatively charged electrons and thus the overall potential $\delta \phi(\boldsymbol{r})$ will be smaller and have maybe a different spatial dependence. TF theory offers a quantitative account of this effect. It predicts a length scale for this.

We will assume that upon the addition of the external charge $\delta n_{+}(\boldsymbol{r})$ we add the same amount of electronic charge, so the system stays neutral. Under these conditions, we have from Eq. (2.3.5):

$$
\frac{2}{3} C n_{*}(\boldsymbol{r})^{-1 / 3} \delta n_{*}(\boldsymbol{r})=\int\left(\delta n_{+}\left(\boldsymbol{r}^{\prime}\right)-\delta n_{*}\left(\boldsymbol{r}^{\prime}\right)\right) u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|\right) d^{3} r^{\prime}
$$

Where we ignore the change in $\mu$. In the end of the calculation we show that the result anyway preserves the total charge of the system so $\delta \mu=0$. The right-hand side of this equation is the perturbation potential, hence:

$$
\begin{equation*}
\delta \phi(\boldsymbol{r})=\frac{2}{3} C n_{*}^{-1 / 3} \delta n_{*}(\boldsymbol{r}) . \tag{2.6.1}
\end{equation*}
$$

For the electron gas we have the simplification that $n_{*}(\boldsymbol{r})=n$ so the linear response equation is:

$$
\frac{2}{3} C n^{-1 / 3} \delta n_{*}(\boldsymbol{r})=\int\left(\delta n_{+}\left(\boldsymbol{r}^{\prime}\right)-\delta n_{*}\left(\boldsymbol{r}^{\prime}\right)\right) u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|\right) d^{3} r^{\prime}
$$

In order to solve this equation we use the Fourier transform:

$$
\begin{aligned}
& \tilde{f}(\boldsymbol{k})=\int f(\boldsymbol{r}) e^{i \boldsymbol{k} \cdot \boldsymbol{r}} d^{3} r \\
& f(\boldsymbol{r})=\frac{1}{(2 \pi)^{3}} \int \tilde{f}(\boldsymbol{k}) e^{-i \boldsymbol{k} \cdot \boldsymbol{r}} d^{3} k .
\end{aligned}
$$

We use the fact that the Fourier transform of a convolution $f(\boldsymbol{r})=\int h\left(\boldsymbol{r}^{\prime}\right) u\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) d^{3} r^{\prime}$ is the product of the Fourier transforms: $\tilde{f}(\boldsymbol{k})=\tilde{u}(\boldsymbol{k}) \tilde{h}(\boldsymbol{k}),{ }^{1}$ Thus: $\frac{2}{3} C n^{-1 / 3} \delta \tilde{n}_{*}(\boldsymbol{k})=$ $\left(\delta \tilde{n}_{+}(\boldsymbol{k})-\delta \tilde{n}_{*}(\boldsymbol{k})\right) \frac{4 \pi \kappa}{k^{2}}$ leading to $\left[\frac{2}{3}\left(\frac{k^{2}}{4 \pi \kappa}\right) C n^{-1 / 3}+1\right] \delta \tilde{n}_{*}(\boldsymbol{k})=$ $\delta \tilde{n}_{+}(\boldsymbol{k})$, from which:

$$
\begin{equation*}
\delta \tilde{n}_{*}(\boldsymbol{k})=\gamma_{T F}^{2} \frac{\delta \tilde{n}_{+}(\boldsymbol{k})}{k^{2}+\gamma_{T F}^{2}} \tag{2.6.2}
\end{equation*}
$$

where:

$$
\begin{equation*}
\gamma_{T F}=\sqrt{\frac{6 \pi \kappa}{C} n^{1 / 3}} \tag{2.6.3}
\end{equation*}
$$

[^1]is the Thomas-Fermi screening constant. One can see that $\delta \tilde{n}_{*}(\boldsymbol{k}=0)=\delta \tilde{n}_{+}(\boldsymbol{k}=0)$ so the total charge is unchanged by the perturbation, as we required. For a point charge in the origin $\delta n_{+}(\boldsymbol{r})=\delta q \delta(\boldsymbol{r})$, and $\delta \tilde{n}_{+}(\boldsymbol{k})=\delta q$ the perturbed charge distribution is $\delta \tilde{n}_{*}(k)=\gamma_{T F}^{2} \frac{\delta q}{k^{2}+\gamma_{T F}^{2}}$ and a reciprocal Fourier transform gives
\[

$$
\begin{aligned}
\delta n_{*}(r) & =\frac{1}{(2 \pi)^{3}} \int \delta n_{*}(\boldsymbol{k}) e^{-i \boldsymbol{k} \cdot \boldsymbol{r}} d^{3} k \\
& =\frac{1}{4 \pi} \gamma_{T F}^{2} \delta q \frac{e^{-\gamma_{T F} r}}{r}
\end{aligned}
$$
\]

The potential is now obtained from Eq. (2.6.1):

$$
\begin{equation*}
\delta \phi(\boldsymbol{r})=\frac{\kappa}{r} \delta q e^{-\gamma_{T F} r} \tag{2.6.4}
\end{equation*}
$$

and is a Yukawa potential. It is similar to the vacuum potential $\frac{\kappa \delta q}{r}$ only when very close to the charge perturbation, i.e. when $\gamma r \ll 1$. As one moves away, the potential acts as if it is due to a smaller charge $\delta q e^{-\gamma r}$. The potential is nearsighted: beyond the distance $\gamma_{T F}^{-1}$ the charge is completely screened away.

## 2.7. von-Weizsäcker kinetic energy

The Thomas-Fermi kinetic energy density functional is exact in the limit of non-interacting homogeneous gas of electrons in an infinite box and also in the limit of an atom at very high density ( $Z \rightarrow \infty$ ). We would like to mention here another density functional which is exact in a certain limit, i.e the limit of a single electron. In this case, if the particle's wave function is $\psi(\boldsymbol{r})$ the kinetic energy is: $T=\frac{\hbar^{2}}{2 m_{e}} \int|\nabla \psi(\boldsymbol{r})|^{2} d^{3} r$ the density of the particle is given by $n(\boldsymbol{r})=|\psi(\boldsymbol{r})|^{2}$. Given the density $n(\boldsymbol{r})$ a kinetic energy can be assigned to it, by imagining the relation $\psi(\boldsymbol{r})=\sqrt{n(\boldsymbol{r})}$ :

$$
T_{v W}[n]=\frac{\hbar^{2}}{2 m_{e}} \int(\nabla \sqrt{n(\boldsymbol{r})})^{2} d^{3} r
$$

Since $\nabla \sqrt{n(\boldsymbol{r})}=\frac{\nabla n(\boldsymbol{r})}{2 \sqrt{n(\boldsymbol{r})}}$ we obtain:

$$
T_{v W}[n]=\frac{\hbar^{2}}{8 m_{e}} \int \frac{(\nabla n(\boldsymbol{r}))^{2}}{n(\boldsymbol{r})} d^{3} r
$$

Defining a wave vector

$$
\boldsymbol{k}_{v W}(\boldsymbol{r})=\frac{1}{2} \frac{\nabla n(\boldsymbol{r})}{n(\boldsymbol{r})}=\frac{1}{2} \nabla \log n(\boldsymbol{r})
$$

we have

$$
T_{v W}[n]=\int \frac{\hbar^{2} \boldsymbol{k}_{v W}(\boldsymbol{r})^{2}}{2 m_{e}} n(\boldsymbol{r}) d^{3} r
$$

Exercise 22. Show that:
The functional derivative of $v_{v W}(\boldsymbol{r})=\frac{\delta T_{v W}}{\delta n(\boldsymbol{r})}$ is:

$$
v_{v W}(\boldsymbol{r})=-\frac{\hbar^{2}}{8 m_{e}}\left(\boldsymbol{k}_{v W}^{2}(\boldsymbol{r})+\nabla \cdot \boldsymbol{k}_{v W}(\boldsymbol{r})\right) .
$$

Given a density $n(\boldsymbol{r})$, what is the potential for which this density is groundstate density?

## CHAPTER 3

## Hartree-Fock Theory

We examine some properties of the many electron wave function with spin in Appendix D. Here we describe only closed shell spin systems.

### 3.1. Slater functions

The many electron wave function is a antisymmetrized product of a spatial $N$ electron wave function and a spin $N$ electron wave function given in Eq.D.3.10 in appendix D. The simplest form is when the spatial part is a product of single particle wave functions $\varphi_{n}(\boldsymbol{r}), n=1, \ldots, N$ and the spin part is a closed shell product:
(3.1.1)
$\Phi(1, \ldots, N)=\hat{A} \varphi_{1}\left(\boldsymbol{r}_{1}\right) \varphi_{1}\left(\boldsymbol{r}_{2}\right) \cdots \varphi_{N / 2}\left(\boldsymbol{r}_{N-1}\right) \varphi_{N / 2}\left(\boldsymbol{r}_{N}\right) \alpha(1) \beta(2) \cdots \alpha(N-1) \beta(N)$
This form reduces to the closed shell "Slater function":
(3.1.2)

$$
\Phi(1, \ldots, N)=\frac{1}{\sqrt{N!}} \operatorname{det}\left[\begin{array}{ccccc}
\varphi_{1}\left(\boldsymbol{r}_{1}\right) & \bar{\varphi}_{1}\left(\boldsymbol{r}_{1}\right) & \cdots & \varphi_{N / 2}\left(\boldsymbol{r}_{1}\right) & \bar{\varphi}_{N / 2}\left(\boldsymbol{r}_{1}\right) \\
\varphi_{1}\left(\boldsymbol{r}_{2}\right) & \bar{\varphi}_{1}\left(\boldsymbol{r}_{2}\right) & \cdots & \varphi_{N / 2}\left(\boldsymbol{r}_{2}\right) & \bar{\varphi}_{N / 2}\left(\boldsymbol{r}_{2}\right) \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\varphi_{1}\left(\boldsymbol{r}_{2}\right) & \bar{\varphi}_{1}\left(\boldsymbol{r}_{2}\right) & \cdots & \varphi_{N / 2}\left(\boldsymbol{r}_{N-1}\right) & \bar{\varphi}_{N / 2}\left(\boldsymbol{r}_{N-1}\right) \\
\varphi_{1}\left(\boldsymbol{r}_{2}\right) & \bar{\varphi}_{1}\left(\boldsymbol{r}_{2}\right) & \cdots & \varphi_{N / 2}\left(\boldsymbol{r}_{N}\right) & \bar{\varphi}_{N / 2}\left(\boldsymbol{r}_{N}\right)
\end{array}\right]
$$

For example, for $N=2$ :
(3.1.3) $\quad \Phi(1,2)=\frac{1}{\sqrt{2}} \operatorname{det}\left[\varphi_{1}\left(\boldsymbol{r}_{1}\right) \bar{\varphi}_{1}\left(\boldsymbol{r}_{2}\right)\right]$

$$
\begin{equation*}
=\frac{1}{\sqrt{2}}\left(\varphi_{1}\left(\boldsymbol{r}_{1}\right) \bar{\varphi}_{1}\left(\boldsymbol{r}_{2}\right)-\varphi_{1}\left(\boldsymbol{r}_{2}\right) \bar{\varphi}_{1}\left(\boldsymbol{r}_{1}\right)\right) \tag{3.1.4}
\end{equation*}
$$

$$
\begin{equation*}
=\varphi_{1}\left(\boldsymbol{r}_{1}\right) \varphi_{1}\left(\boldsymbol{r}_{2}\right) \frac{1}{\sqrt{2}}(\alpha \beta-\beta \alpha) \tag{3.1.5}
\end{equation*}
$$

Lemma 23. In a nonzero Slater function $\Phi(1, \ldots, N)=$ $\operatorname{det}\left[\varphi_{1}(1) \varphi_{2}(2) \cdots \varphi_{N}(N)\right]$, if $\varphi_{n}(\boldsymbol{x})=\sum_{m=1}^{N} \phi_{m}(\boldsymbol{x}) A_{m n}$, $n=1, \ldots, N$ then $\Phi(1, \ldots, N)=\operatorname{det} A \times \operatorname{det}\left[\phi_{1}(1) \phi_{2}(2) \cdots \phi_{N}(N)\right]$.

Proof. Simply a property of determinants: $\operatorname{det}[A B]=$ $\operatorname{det}[A] \operatorname{det}[B]$.

Corollary 24. We may assume without loss of generality that the single particle wave functions $\phi_{n}(\boldsymbol{x})$ in a Slater function $\Phi(1, \ldots, N)=\operatorname{det}\left[\phi_{1}(1) \phi_{2}(2) \cdots \phi_{N}(N)\right]$ are orthogonal.

Proof. If they are not, we can choose $\varphi_{n}(\boldsymbol{x})=\sum_{m=1}^{N} \phi_{m}(\boldsymbol{x}) A_{m n}$ where $A$ is a $N \times N$ matrix which are selected in a way that renders $\varphi_{n}(\boldsymbol{x})$ an orthonormal set. Then, since $\Phi(1, \ldots, N)=$ $\operatorname{det} A \times \operatorname{det}\left[\varphi_{1}(1) \varphi_{2}(2) \cdots \varphi_{N}(N)\right]$ the two Slater functions $\operatorname{det}\left[\phi_{1}(1) \phi_{2}(2) \cdots \phi_{N}(N)\right]$ and $\operatorname{det}\left[\varphi_{1}(1) \varphi_{2}(2) \cdots \varphi_{N}(N)\right]$ are identical to within a multiplicative factor $\operatorname{det} A$.

Exercise 25. Find the matrix $A$ which transforms the $N$ wave functions $\phi_{m}(\boldsymbol{x})$ to a set $\varphi_{n}(\boldsymbol{x})=\sum_{m=1}^{N} \phi_{m}(\boldsymbol{x}) A_{m n}$ which is orthonormal, i.e.: $\delta_{n n^{\prime}}=\left\langle\varphi_{n} \mid \varphi_{n^{\prime}}\right\rangle$

Solution. To obtain an equation for $A$, we expand this condition:

$$
\begin{align*}
\delta_{n n^{\prime}} & =\left\langle\sum_{m=1}^{N} \phi_{m} A_{m n} \mid \sum_{m^{\prime}=1}^{N} \phi_{m^{\prime}} A_{m^{\prime} n^{\prime}}\right\rangle \\
& =\sum_{m=1}^{N} A_{m n}^{*} \sum_{m^{\prime}=1}^{N} A_{m^{\prime} n^{\prime}} S_{m m^{\prime}} \tag{3.1.6}
\end{align*}
$$

where $S_{m m^{\prime}}=\left\langle\phi_{m} \mid \phi_{m^{\prime}}\right\rangle$ is the $N \times N$ overlap matrix of the $N$ $\phi$ orbitals from which $\Phi_{S}$ is composed. Eq. (3.1.6)can be written in matrix form as: $I=A^{\dagger} S A$, where $I$ is the $N \times N$ identity matrix, and $A$ can now be expressed formally as: $A A^{\dagger}=S^{-1}$. If $S$ is invertible, this equation has many solutions. For example, since $S$ is Hermitean ( $S=S^{\dagger}$ ) we can diagonalize: $S=U s U^{\dagger}$ where $U$ is the diagonalizing unitary matrix $\left(U U^{\dagger}=\right.$
$\left.U^{\dagger} U=I\right)$ and $s$ the diagonal matrix of eigenvalues of $S$. Then, one can select $A=U s^{1 / 2}$. The Hermitean overlap matrix $S$ is positive definite, i.e. all it's eigenvalues are positive. For suppose the normal column vector $u$, is an eigenvector of $S$, corresponding to the eigenvalue $s$, then $s=u^{\dagger} S u$. Indeed, for t we have:

$$
\begin{aligned}
s & =u^{\dagger} S u \\
& \sum_{m m^{\prime}=1}^{N} u_{m}^{*} S_{m m^{\prime}} u \\
& =\sum_{m m^{\prime}=1}^{N} u_{m}^{*}\left\langle\phi_{m} \mid \phi_{m^{\prime}}\right\rangle u_{m^{\prime}} \\
& =\left\langle\sum_{m=1}^{N} \phi_{m} u_{m} \mid \sum_{m^{\prime}=1}^{N} \phi_{m^{\prime}} u_{m^{\prime}}\right\rangle \\
& >0
\end{aligned}
$$

the last line is correct because the overlap of a wave function $\sum_{m=1}^{N} \phi_{m}(\boldsymbol{x}) u_{m}$ with itself is always positive.

### 3.2. Density, reduced density matrix of Slater functions

We defined the density and reduced density matrix of a Fermionic wave function in Eqs. (1.5.2) and (1.5.4). For Slater functions, $\Phi=\operatorname{det}\left[\phi_{1} \ldots \phi_{N}\right]$ these objects have more specific form, in terms of the one particles orbitals. The RDM of Eq. (1.5.4) becomes:

$$
\begin{equation*}
\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)=\sum_{n=1}^{N} \phi_{n}(\boldsymbol{r}) \phi_{n}\left(\boldsymbol{r}^{\prime}\right)^{*} \tag{3.2.1}
\end{equation*}
$$

and the density, $n_{\Phi}(\boldsymbol{r})=\langle\Phi| \hat{n}(\boldsymbol{r})|\Phi\rangle=\rho_{\Phi}(\boldsymbol{r}, \boldsymbol{r})$ is

$$
\begin{equation*}
n_{\Phi}(\boldsymbol{r})=\sum_{n=1}^{N}\left|\phi_{n}(\boldsymbol{r})\right|^{2} \tag{3.2.2}
\end{equation*}
$$

Exercise 26. Prove that $\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)$ is idempotent, i.e.

$$
\begin{equation*}
\int \rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime \prime}\right) \rho_{\Phi}\left(\boldsymbol{r}^{\prime \prime}, \boldsymbol{r}^{\prime}\right) d \boldsymbol{r}^{\prime \prime}=\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right) \tag{3.2.3}
\end{equation*}
$$

Solution. Simply use the definition of a density matrix in Eq. (3.2.1) in terms of single particle wave functions twice and then use the orthogonality of these wave functions.

EXERCISE 27. Calculate the density matrix of homogeneous electron gas particles of density $n=\frac{N}{V}$ where $N$ is the number of electrons and $V$ is the box size in the thermodynamic limit i.e. when both $N$ and $V$ go to $\infty$. Hint: Assume that the single particle states are $\varphi_{\boldsymbol{k}}(\boldsymbol{r})=\frac{e^{i \boldsymbol{k} \cdot \boldsymbol{r}}}{\sqrt{V}}$ and $\boldsymbol{k}$ is such that $\boldsymbol{k}<k_{F}$ where $k_{F}$ is the Fermi wave number (see Eq. (2.1.7)).

Solution. The DM $\rho\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)$ is a sum over products of the type $\varphi_{\boldsymbol{k}}(\boldsymbol{r}) \varphi_{\boldsymbol{k}}\left(\boldsymbol{r}^{\prime}\right)^{*}=\frac{e^{i \boldsymbol{k} \cdot\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right)}}{V}$ and therefore the DM is a function of $s \equiv r-r^{\prime}$, but since there is no preferred direction in space, we can say that $s=s \hat{\boldsymbol{z}}$ is in the z direction and the DM becomes a function of $s \equiv\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|: \rho(s)=2 \sum_{|\boldsymbol{k}|<k_{F}} \frac{e^{i k \cdot s}}{V}$. We can replace the sum by integral (see Eq. (2.1.5)): $\rho(s)=$ $\frac{2}{(2 \pi)^{3}} \int_{0}^{k_{F}} k^{2} d k \int_{0}^{\pi} \sin \theta d \theta \int_{0}^{2 \pi} d \phi e^{i k s \cos \theta}$, where we used the definition of a scalar product $\boldsymbol{k} \cdot \boldsymbol{s}=k s \cos \theta$. Evaluating the integrals we find: $\rho(s)=\frac{2}{(2 \pi)^{2}} \int_{0}^{k_{F}} k^{2} d k \times \frac{2 \sin k s}{k s}=\frac{k_{F}^{3}}{\pi^{2}} \frac{\sin k_{F} s-k_{F} s \cos k_{F} s}{\left(k_{F} s\right)^{3}}$ :

$$
\begin{equation*}
\rho(s)=\frac{k_{F}^{3}}{\pi^{2}} \frac{j_{1}\left(k_{F} s\right)}{k_{F} s} \tag{3.2.4}
\end{equation*}
$$

where $j_{1}(x)=\frac{\sin x-x \cos x}{x^{2}}$ is the spherical Bessel function of order 1 . The density matrix $\rho(s)$ is plotted in Fig. 3.2.1.


Figure 3.2.1. The density matrix $\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)=$ $2 \sum_{|\boldsymbol{k}|<k_{F}} \varphi_{\boldsymbol{k}}(\boldsymbol{r}) \varphi_{\boldsymbol{k}}\left(\boldsymbol{r}^{\prime}\right)^{*}=\rho(s)$, where $s=$ $\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|$, corresponding to a Slater determinant $\Phi$ composed of the wave vector plane waves $\varphi_{\boldsymbol{k}}(\boldsymbol{r})=e^{i \boldsymbol{k} \cdot \boldsymbol{r}} / \sqrt{V}$ where $|\boldsymbol{k}|<k_{F}$ in a box of volume $V$.

### 3.3. Expectation values in Slater functions

Consider a 1-particle operator of the form: $\hat{O}=\sum_{n=1}^{N} \hat{o}(n)$. The expectation value in a Slater function with orthonormal orbitals is:

$$
\begin{equation*}
\langle\Phi| \hat{O}|\Phi\rangle=\sum_{n=1}^{N}\langle n| \hat{o}|n\rangle \tag{3.3.1}
\end{equation*}
$$

where the notation $\langle n| \hat{o}|n\rangle \equiv\left\langle\varphi_{n}\right| \hat{o}\left|\varphi_{n}\right\rangle$ is used. For closed shell:

$$
\begin{equation*}
\langle\Phi| \hat{O}|\Phi\rangle=2 \sum_{n=1}^{N / 2}\langle n| \hat{o}|n\rangle \tag{3.3.2}
\end{equation*}
$$

For two body operators: $\hat{U}=\frac{1}{2} \sum_{n \neq m}^{N} \hat{u}(n, m)$ we have then (using the notation $\langle n m| \hat{u}|k l\rangle \equiv\left\langle\varphi_{n} \varphi_{m}\right| \hat{u}\left|\varphi_{k} \varphi_{l}\right\rangle$ ):

$$
\begin{align*}
\langle\Phi| \hat{U}|\Phi\rangle & =\frac{1}{2} \sum_{n \neq m}^{N}[\langle n m| \hat{u}|n m\rangle-\langle n m| \hat{u}|m n\rangle]  \tag{3.3.3}\\
& =\frac{1}{2} \sum_{n, m}^{N}[\langle n m| \hat{u}|n m\rangle-\langle n m| \hat{u}|m n\rangle] .
\end{align*}
$$

Note the second equality, where the condition $n \neq m$ was removed because the term $m=n$ is canceled between the two contributions.

For the Coulomb interaction, $\hat{u}_{C}(n, m)=\frac{\kappa}{\left|\boldsymbol{r}_{n}-\boldsymbol{r}_{m}\right|}$ where $\boldsymbol{r}_{n}$ and $\boldsymbol{r}_{m}$ are the positions of electrons $n$ and $m$, the expectation value $\langle\Phi| \hat{U}|\Phi\rangle$ is thus written as:

$$
\begin{equation*}
\langle\Phi| \hat{U}_{C}|\Phi\rangle=J[\Phi]+K[\Phi] \tag{3.3.5}
\end{equation*}
$$

where $J$ and $K$ are the Coulomb and exchange energies,respectively defined by:

$$
\begin{align*}
J[\Phi] & \equiv \frac{1}{2} \sum_{n, m}^{N}\langle n m| \hat{u}_{C}|n m\rangle  \tag{3.3.6}\\
K[\Phi] & \equiv-\frac{1}{2} \sum_{n, m}^{N}\langle n m| \hat{u}_{C}|m n\rangle \tag{3.3.7}
\end{align*}
$$

The Coulomb energy can be written as the Hartree-energy functional:

$$
\begin{equation*}
J[\Phi]=E_{H}\left[n_{\Phi}\right] \tag{3.3.8}
\end{equation*}
$$

where

$$
\begin{equation*}
E_{H}[n]=\frac{1}{2} \iint n(\boldsymbol{r}) u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|\right) n\left(\boldsymbol{r}^{\prime}\right) d \boldsymbol{r} d \boldsymbol{r}^{\prime} \tag{3.3.9}
\end{equation*}
$$

and where $n_{\Phi}(\boldsymbol{r})$ is the electron density of the Slater function $\Phi$, given in Eq. (3.2.2). The exchange energy can be written in
terms of the density matrix

$$
\begin{equation*}
K[\Phi]=-\frac{1}{2} \iint\left|\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)\right|^{2} u_{C}\left(\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|\right) d \boldsymbol{r} d \boldsymbol{r}^{\prime} \tag{3.3.10}
\end{equation*}
$$

where $\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)$ is the reduced density matrix (RDM) of the Slater function $\Phi$, given in Eq. (3.2.1).

For closed shell we still use these definitions but $N$ in Eqs. 3.3.6 and 3.3.7is now replaced by $N / 2$ and the integrals do not need to include spin. The expectation value is then:

$$
\begin{equation*}
\left\langle\Phi_{S}\right| \hat{U}_{C}\left|\Phi_{S}\right\rangle=4 J+2 K(\text { closed shell }) . \tag{3.3.11}
\end{equation*}
$$

EXERCISE 28. Calculate the exchange energy of a Slater determinant

### 3.4. Hartree-Fock theory

For electrons in an external potential $v_{\text {ext }}(\boldsymbol{r})$, the Hamiltonian, in the non-relativistic approximation, for $N$ electrons is given by:

$$
\begin{aligned}
\hat{H} & =\hat{h}+\hat{U} \\
& =\sum_{n=1}^{N} \hat{h}(n)+\frac{1}{2} \sum_{n \neq m}^{N} u_{C}\left(\left|\boldsymbol{r}_{n}-\boldsymbol{r}_{m}\right|\right)
\end{aligned}
$$

where $u_{C}(r)=\frac{\kappa}{r}$ is the Coulomb repulsion potential and $\hat{h}=-\frac{\hbar^{2}}{2 m_{e}} \nabla^{2}+v_{\text {ext }}(\boldsymbol{r})$. The ground state wave function and energy of this Hamiltonian obey: $\hat{H} \Psi_{G S}=E_{G S} \Psi_{G S}$. The wave function $\Psi_{G S}(1, \ldots, N)$ is extremely complicated and cannot be found to high accuracy except perhaps for very small systems of two electrons. Hence we must resort to approximations. The basic approximation is the Hartree-Fock approximation discussed now.

To obtain a theory for the ground state energy and wave function we use the variational principle, which says that for
any Slater wave function $\Phi, E_{G S} \leq\langle\Phi| \hat{H}|\Phi\rangle$. We thus search for the Slater function $\Phi_{H F}$ that minimizes the right hand side, giving the HF energy $E_{H F}=\left\langle\Phi_{H F}\right| \hat{H}\left|\Phi_{H F}\right\rangle$. This energy will be the best Slater function approximation to the true ground state energy $E_{G S}$.

In order to find the best Slater function, let us express the expectation value of the Hamiltonian in terms of the orbitals: $\langle\Phi| \hat{H}|\Phi\rangle=\sum_{m=1}^{N}\left\langle\phi_{m}\right| \hat{h}\left|\phi_{m}\right\rangle+J[\Phi]+K[\Phi]$ (see Eqs. (3.3.6)-(3.3.7)), and for a closed shell system:

$$
\begin{equation*}
\langle\Phi| \hat{H}|\Phi\rangle=2 \sum_{m=1}^{N / 2}\left\langle\phi_{m}\right| \hat{h}\left|\phi_{m}\right\rangle+4 J[\Phi]+2 K[\Phi] . \tag{3.4.1}
\end{equation*}
$$

Equations for the orbitals $\phi_{m}(\boldsymbol{r})$ are found by minimizing $\langle\Phi| \hat{H}|\Phi\rangle$ under the constrained of orthonormality. This involves the Lagrangian

$$
L\left[\phi_{1}, \ldots, \phi_{N}\right]=\langle\Phi| \hat{H}|\Phi\rangle-\sum_{n m=1}^{N} \varepsilon_{n m}\left(\langle n \mid m\rangle-\delta_{n m}\right),
$$

where $\varepsilon_{n m}$ are Lagrange multipliers. Since we can assume without loss of generality that $\phi_{n}(\boldsymbol{r})$ are orthogonal, we will obtain the first simplification by demanding only orthonormality using the Lagrangian: $L\left[\phi_{1}, \ldots, \phi_{N}\right]=\langle\Phi| \hat{H}|\Phi\rangle-$ $\sum_{m=1}^{N} \varepsilon_{m}(\langle m \mid m\rangle-1)$. The variation to first order is:
$\delta L=\delta\langle\Phi| \hat{H}|\Phi\rangle-\sum_{m=1}^{N} \delta \varepsilon_{m}\left(\left\langle\phi_{m} \mid \phi_{m}\right\rangle-1\right)-\sum_{m=1}^{N} \varepsilon_{m} \delta\left\langle\phi_{m} \mid \phi_{m}\right\rangle$.
At the constrained minimum $\langle m \mid m\rangle-1$ and

$$
\begin{equation*}
\delta\langle\Phi| \hat{H}|\Phi\rangle=\sum_{m=1}^{N} \varepsilon_{m} \delta\langle m \mid m\rangle \tag{3.4.2}
\end{equation*}
$$

The variation of the bra can be assumed separate from the variation in the ket. Hence

$$
\begin{equation*}
\langle\delta \Phi| \hat{H}|\Phi\rangle=\sum_{m=1}^{N} \varepsilon_{m}\left\langle\delta \phi_{m} \mid \phi_{m}\right\rangle \tag{3.4.3}
\end{equation*}
$$

And since we can use any variation we wish, let us change only $\phi_{i}(\boldsymbol{r})$ for a certain $i$ and leave other orbitals untouched. Then, using $\frac{\delta}{\delta \phi_{i}(\boldsymbol{r})^{*}}\langle m \mid m\rangle=\delta_{i m} \phi_{i}(\boldsymbol{r})$, we have:

$$
\begin{equation*}
\left\langle\frac{\delta \Phi}{\delta \phi_{i}(\boldsymbol{r})}\right| \hat{H}|\Phi\rangle=\varepsilon_{i} \phi_{i}(\boldsymbol{r}) \tag{3.4.4}
\end{equation*}
$$

We now need to evaluate the left hand side. We use:

$$
\begin{equation*}
\frac{\delta}{\delta \phi_{i}(\boldsymbol{r})^{*}} \sum_{n=1}^{N}\langle n| \hat{h}|n\rangle=\hat{h} \phi_{i}(\boldsymbol{r}) \tag{3.4.5}
\end{equation*}
$$

for the one body part and for the direct interaction.
Exercise 29. Show that:

$$
\begin{align*}
\frac{\delta}{\delta \phi_{i}(\boldsymbol{r})^{*}} J[\Phi] & =\left(\int u_{C}\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) n_{\Phi}\left(\boldsymbol{r}^{\prime}\right) d \boldsymbol{r}^{\prime}\right) \phi_{i}(\boldsymbol{r})  \tag{3.4.6}\\
& =v_{H}\left[n_{\Phi}\right](\boldsymbol{r}) \phi_{i}(\boldsymbol{r})
\end{align*}
$$

where

$$
\begin{equation*}
v_{H}[n](\boldsymbol{r}) \equiv \int u_{C}\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) n\left(\boldsymbol{r}^{\prime}\right) d \boldsymbol{r}^{\prime} \tag{3.4.7}
\end{equation*}
$$

is called the Hartree potential, and it is the classical electrostatic potential associated with the electron density $n(\boldsymbol{r})$. Furthermore, show that:
(3.4.8) $\frac{\delta}{\delta \phi_{i}(\boldsymbol{r})^{*}} K[\Phi]=-\int u_{C}\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) \rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right) \phi_{i}\left(\boldsymbol{r}^{\prime}\right) d \boldsymbol{r}^{\prime}$

Solution. Hint: we use the facts that

$$
\begin{aligned}
\frac{\delta}{\delta \phi_{i}(\boldsymbol{r})^{*}} n_{\Phi}\left(\boldsymbol{r}^{\prime}\right) & =\delta\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) \phi_{i}(\boldsymbol{r}) . \\
\frac{\delta}{\delta \phi_{i}(\boldsymbol{r})^{*}} \rho_{\Phi}\left(\boldsymbol{r}^{\prime}, \boldsymbol{r}^{\prime \prime}\right) & =\delta\left(\boldsymbol{r}-\boldsymbol{r}^{\prime \prime}\right) \phi_{i}\left(\boldsymbol{r}^{\prime}\right) .
\end{aligned}
$$

Using Eqs. (3.4.5), (3.4.6) and (3.4.8) in Eq. (3.4.4) we find the celebrated HF equations:

$$
\begin{equation*}
\hat{F} \phi_{i}(\boldsymbol{r})=\varepsilon_{i} \phi_{i}(\boldsymbol{r}) \tag{3.4.9}
\end{equation*}
$$

where $\hat{F}$ is the Fockian:

$$
\begin{equation*}
\hat{F}=\hat{h}+v_{H}(\boldsymbol{r})+\hat{k} \tag{3.4.10}
\end{equation*}
$$

and where

$$
\begin{equation*}
\hat{k} \phi(\boldsymbol{r}) \equiv-\int u_{C}\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) \rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right) \phi\left(\boldsymbol{r}^{\prime}\right) d \boldsymbol{r}^{\prime} \tag{3.4.11}
\end{equation*}
$$

is the exchange operator. This equation should be solved together with Eqs. (3.2.2) and (3.2.1). They are of the eigenvalue type, yet the Fockian itself depends on the density $n_{\Phi}(\boldsymbol{r})$ and the RDM $\rho_{\Phi}\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)$, which both depend on the orbitals $\phi_{i}(\boldsymbol{r})$. Hence, the HF equations are called self-consistent field equations. They should be solved by demanding that $\phi_{i}(\boldsymbol{r})$ should be the eigenstates of $\hat{F}$ and at the same time $\hat{F}$ is built out of the $\phi_{i}$.

Once we solve the HF equations, we have the HF function $\Phi_{H F}$ which is the Slater function minimizing the expectation value of $\hat{H}$, Eq. (3.4.1), hence the HF energy is

$$
\begin{align*}
E_{H F} & =\left\langle\Phi_{H F}\right| \hat{H}\left|\Phi_{H F}\right\rangle \\
& =\sum_{m=1}^{N}\langle m| \hat{h}|m\rangle+J\left[\Phi_{H F}\right]+K\left[\Phi_{H F}\right] . \tag{3.4.12}
\end{align*}
$$

where we use the shorthand notation $\left\langle\phi_{m}\right| \hat{h}\left|\phi_{m}\right\rangle \equiv\langle m| \hat{h}|m\rangle$.

Exercise 30. Show the HF energy can also be written as

$$
\begin{equation*}
E_{H F}=\sum_{m=1}^{N} \varepsilon_{m}-\left(J\left[\Phi_{H F}\right]+K\left[\Phi_{H F}\right]\right) \tag{3.4.13}
\end{equation*}
$$

Solution. Hint: show that, for all $m=1,2, \ldots$ :
$\varepsilon_{m}=\langle m| \hat{F}|m\rangle=\langle m| \hat{h}|m\rangle+\sum_{i=1}^{N}\left[\langle m i| u_{C}|m i\rangle-\langle m i| u_{C}|i m\rangle\right]$.
Then sum over $m=1 \ldots, N$ and compare.
For the closed shell case, called restricted HF (RHF), we have
(3.4.14) $\quad E_{R H F}=2 \sum_{m=1}^{N / 2}\langle m| \hat{h}|m\rangle+4 J\left[\Phi_{H F}\right]+2 K\left[\Phi_{H F}\right]$

Unlike the TF theory the HF theory can describe chemical bonds. For example, application of RHF to the $H_{2}$ molecule in the typical bond length ( $1.4 a_{0}$ gives the value $E_{R H F}=$ $-1.134 E_{h}$. Compared to the energy of 2 H atoms, each with ground state energy of $-\frac{1}{2} E_{h}$, this result indicates that the $H_{2}$ atomization energy is $0.134 E_{h}=3.65 \mathrm{eV}$. This calculated value should be compared to the $H_{2}$ atomization energy estimate based on experimental data, which is about 4.75 eV , showing that the RHF approximation does not give high quality atomization energies, since the deficit exceeding 1 eV is substantial in Chemical energy terms.

Linear response. Here, we consider perturbations of the system which moves it from one ground state to another ground state. This is called an adiabatic or time-independent perturbation. In the most general form the Hamiltonian $\hat{H}$ is thought to depend on a parameter $\lambda$ and the perturbation is a change in this parameter: $\lambda \rightarrow \lambda+\delta \lambda$. The HF energy and wave function depend on $\lambda$ so: $E_{H F}(\lambda)=\langle\Phi(\lambda)| \hat{H}(\lambda)|\Phi(\lambda)\rangle$. The change
is $\delta E_{H F}=\langle\delta \Phi(\lambda)| \hat{H}(\lambda)|\Phi(\lambda)\rangle+\langle\Phi(\lambda)| \hat{H}(\lambda)|\delta \Phi(\lambda)\rangle+$ $\langle\Phi(\lambda)| \delta \hat{H}(\lambda)|\Phi(\lambda)\rangle$. Taking into account Eq. 3.4.2 the sum of the first two terms on the right-hand side is $\sum_{m=1}^{N} \varepsilon_{m} \delta\left\langle\phi_{m} \mid \phi_{m}\right\rangle$ and since the orbitals stay orthonormal when $\lambda$ changes, this is zero ( $\delta\left\langle\phi_{m} \mid \phi_{m}\right\rangle=0$ ). Hence, $\delta E_{H F}=\langle\Phi(\lambda)| \delta \hat{H}(\lambda)|\Phi(\lambda)\rangle$ or:

$$
\begin{equation*}
\delta E_{H F}=\langle\Phi(\lambda)| \hat{H}^{\prime}(\lambda)|\Phi(\lambda)\rangle \delta \lambda \tag{3.4.15}
\end{equation*}
$$

This result is the HF analog of the Feynman-Hellman theorem which applies to the exact GS wave function and eigenvalue $\Psi_{G S}(\lambda)$ and $E_{G S}(\lambda)$ :

$$
\begin{equation*}
\delta E_{G S}=\left\langle\Psi_{G S}(\lambda)\right| \hat{H}^{\prime}(\lambda)\left|\Psi_{G S}(\lambda)\right\rangle \delta \lambda \tag{3.4.16}
\end{equation*}
$$

We thus see that the effect of a perturbation on the ground state energy is, to first order, simply the expectation value of $\hat{H}^{\prime}(\lambda)$. When the perturbation is a one-body operator $\lambda \sum_{n} \hat{p}(n)$ then, taking the perturbation around $\lambda=0$ we find:

$$
\begin{equation*}
\delta E_{H F}=\delta \lambda \sum_{n=1}^{N}\left\langle\phi_{n}\right| \hat{p}\left|\phi_{n}\right\rangle . \tag{3.4.17}
\end{equation*}
$$

We see that the derivative of the HF energy with respect to the perturbation parameter $\lambda$ gives the ground state expectation value of the operator in HF theory.

What about the wave functions themselves? How do they change? First order perturbation theory states that if we assume the Fockian does not change

$$
\begin{equation*}
\delta \phi_{n}(\boldsymbol{r})=\sum_{k \neq n} \phi_{k}(\boldsymbol{r}) \frac{\left\langle\phi_{k}\right| \hat{p}\left|\phi_{n}\right\rangle}{\varepsilon_{n}-\varepsilon_{k}} . \tag{3.4.18}
\end{equation*}
$$

### 3.5. The HF theory of the homogeneous electron gas

For the electron gas we know the solution of the HF equations, so all we need is to write down the energy: $E_{H F}=$


Figure 3.5.1. The Hartree-Fock energy of the homogeneous electron gas of density $n$.
$E_{T F}+E_{X}$, where the $E_{T F}$ is the Thomas Fermi energy given in 2.1.12. For the electron gas all the charge energies cancel each other so $E_{T F}=T_{T F}$ and thus the HF energy is $\varepsilon_{H F}(n)=$ $t_{S}(n)+\varepsilon_{X}(n)$ where $t_{S}=\frac{3}{5} C_{S} n^{2 / 3}$ is the Thomas Fermi kinetic energy per particle where $C=4.785 \frac{\hbar^{2}}{m_{e}}$ (see Eq. (2.1.9)). The exchange energy per particle is $\varepsilon_{X}(n)=\frac{K_{X}}{N}$, which can be estimated by evaluating $K_{X}=-\frac{1}{4} \iint \frac{\rho\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)^{2}}{\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|} d \boldsymbol{r} d \boldsymbol{r}^{\prime}$. For this case the density matrix was discussed in Exercise (27), where $\rho\left(\boldsymbol{r}, \boldsymbol{r}^{\prime}\right)$ was shown a function of $s=\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|$, so: $K_{X}=$ $-\frac{1}{4} \kappa V \int_{V} \frac{\rho(s)^{2}}{s} d s$. The integrand is a function of $s$ only and not the direction of $s$ so the integral can be accomplished by spherical integration, we further assume the box size is so large that we can take the integration to infinity: $K_{X}=$ $-\frac{1}{4} V \int_{0}^{\infty} \frac{\rho(s)^{2}}{s} 4 \pi s^{2} d s$. Using the form of $\rho(s)$ in Eq. (3.2.4) the integral can be calculated analytically, giving: $K_{X}=-\frac{1}{4} \kappa V \times$ $\frac{k_{F}^{4}}{\pi^{3}}$. Since $k_{F}^{3}=3 \pi^{2} n$ (Eq. (2.1.7)), we find:

$$
\begin{equation*}
\varepsilon_{X}=\frac{K_{X}}{N}=-\frac{3}{4 \pi} \kappa k_{F}=-\frac{3}{4} C_{x} \times n^{1 / 3} \tag{3.5.1}
\end{equation*}
$$

where $C_{x}=\left(\frac{3}{\pi}\right)^{1 / 3} E_{h} a_{0}=0.9845 E_{h} a_{0}$. Together with the kinetic energy, HF energy per electron is thus: $\varepsilon_{H F}=\frac{3}{5} C_{S} \times$ $n^{2 / 3}-\frac{3}{4} C_{x} \times n^{1 / 3}$, plotted in 3.5.1. If we assume that the electron gas could determine its own density, i.e. given a large number of electrons, they would self-determine their own volume, then they would choose to minimize $\varepsilon_{H F}$. This minimum is obtained (solving $\varepsilon_{H F}^{\prime}\left(n_{*}\right)=0$ ), at $n_{*}=\left(\frac{5}{8} \frac{C_{x}}{C_{S}}\right)^{3}=$ $0.0021 a_{0}^{-3}$ and $\varepsilon_{H F}^{*}=-\frac{15}{64} \frac{C_{x}^{2}}{C_{s}}=-1.29 \mathrm{eV}$, both values are close to the density and the energy per electron of sodium, $0.004 a_{0}^{-3}$ and -1.13 eV , respectively. The pressure of the HEG is $P=n^{2} \varepsilon_{H F}^{\prime}$ and the bulk modulus is $Y=n P^{\prime}(n)=\frac{1}{3} n^{4 / 3}\left(C_{x}-2 C_{s} n^{1 / 3}\right)$ gives $Y=n_{*} P^{\prime}\left(n_{*}\right)=0.66 G P a$ (which is much lower than that of sodium, 6.3 GPa ).

The HF eigenvalues are $\varepsilon(k)=\varepsilon_{\boldsymbol{k}}=\frac{\hbar^{2} k^{2}}{2 m_{e}}+\left\langle\psi_{\boldsymbol{k}}\right| \hat{k}\left|\psi_{\boldsymbol{k}}\right\rangle$ where:

$$
\left\langle\psi_{\boldsymbol{k}}\right| \hat{k}\left|\psi_{\boldsymbol{k}}\right\rangle=-\kappa V^{-1} \iint e^{-i \boldsymbol{k} \cdot \boldsymbol{r}} \int \frac{\rho\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right)}{\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|} e^{i \boldsymbol{k} \cdot \boldsymbol{r}^{\prime}} d \boldsymbol{r}^{\prime} d \boldsymbol{r}
$$

which, using assumptions similar to the calculation of exchange energy (see Exercise 27),

$$
\begin{aligned}
\left\langle\psi_{\boldsymbol{k}}\right| \hat{k}\left|\psi_{\boldsymbol{k}}\right\rangle & =-2 \pi \kappa \int_{0}^{\infty} s^{2} d s \frac{\rho(s)}{s} \int_{-1}^{1} e^{i k s x} d x \\
& =-4 \pi \kappa \int_{0}^{\infty} \rho(s) \frac{\sin k s}{k s} \times s d s
\end{aligned}
$$

leads to:

$$
\begin{equation*}
\varepsilon(k)=\frac{\hbar^{2} k^{2}}{2 m_{e}}\left[1+\frac{1}{k_{F} a_{0}} F\left(\frac{k}{k_{F}}\right)\right] . \tag{3.5.2}
\end{equation*}
$$

where $F(x)=\frac{4}{\pi x^{2}}\left(\left(x-x^{-1}\right) \tanh ^{-1}\left[e^{-|\ln x|}\right]-1\right)$. The eigenvalues are given in the upper panel of Fig. 3.5.2 and it is seen that at $k=k_{F}$ there is a "feature" (within the ellipse) in the eigenvalue function. This is seen in the density of states
$\rho(\varepsilon(k))=4 \pi \frac{V k^{2}}{\varepsilon^{\prime}(k)}$, plotted in the lower panel of the figure, which diverges to zero at $k=k_{F}$. As we learned, the occupied eigenvalues are ionization potentials, i.e. "hole" energies of the system and the drop to zero of density of hole states predicted by HF theory contradicts experiments on simple metals, which show a high DOS of hole states near the Fermi level. This failure is associated with the tendency of the HF band gap (the difference between the lowest unoccupied and the highest occupied energy eigenvalues) to be much larger than the experimental hole-electron energy gaps.


Figure 3.5.2. The Hartree-Fock wave number dependent eigenvalues $\varepsilon_{k}$ (upper panel) and density of states (DOS, lower panel) for the homogeneous electron gas at the density $n_{*}$ minimizing the HF energy (see Fig. 3.5.1).

## CHAPTER 4

## Hohenberg-Kohn theory

In view of the poor predictions of chemical bonds and molecular properties afforded by HF approximation and the high numerical price of wave function approaches, it is beneficial to seek out methods that circumvent the need to represent the many-body electronic wave function. We studied in detail two theories. One was based on the density, but had no real rigorous basis. The other was a method that assumed the electronic wave function is of the form applicable only for non-interacting electrons. We now want to describe a rigorous method that combines ideas of this type in a new way which is both rigorous and leads to very accurate approximations.

### 4.1. Hohenberg-Kohn theorem

In electronic structure theory the Hamiltonian is given as:

$$
\begin{equation*}
\hat{H}=\hat{T}+\hat{U}+\int v(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d \boldsymbol{r} \tag{4.1.1}
\end{equation*}
$$

Where $\hat{T}=\sum_{n=1}^{N} \frac{-\hbar^{2}}{2 m_{e}} \nabla_{n}^{2}$ is the kinetic energy and $\hat{U}=\frac{1}{2} \sum_{m \neq n} u_{C}\left(\mid \boldsymbol{r}_{n}-\boldsymbol{r}_{n}\right.$ is the Coulomb repulsion and $v(\boldsymbol{r})$ is the external potential acting on the electrons, represented by their density function $\hat{n}(\boldsymbol{r})=\sum_{n=1}^{N} \delta\left(\boldsymbol{r}-\hat{\boldsymbol{r}}_{n}\right)$ where $\hat{\boldsymbol{r}}_{n}$ is the position of electron $n$. It is essential to appreciate that the terms $\hat{F}=\hat{T}+\hat{U}$ and $\hat{n}(\boldsymbol{r})$ appearing in $\hat{H}$ are the same for all $N$-electron Hamiltonians. Thus what differentiates one system $\hat{H}$ from another system $\hat{H}^{\prime}$ is only the difference in the external potentials $v(\boldsymbol{r})$ and
$v^{\prime}(\boldsymbol{r}):$

$$
\hat{H}-\hat{H}^{\prime}=\int\left(v(\boldsymbol{r})-v^{\prime}(\boldsymbol{r})\right) \hat{n}(\boldsymbol{r}) d \boldsymbol{r}=\sum_{n=1}^{N}\left(v\left(\hat{\boldsymbol{r}}_{n}\right)-v^{\prime}\left(\hat{\boldsymbol{r}}_{n}\right)\right)
$$

As an application of this, we show that there are no eigenstates of $\hat{H}$, which are also eigenstates of $\hat{H}^{\prime}$. For suppose that $\Psi_{m}$ is such an eigenstate:

$$
\begin{equation*}
\left(E_{m}-E_{m}^{\prime}\right) \Psi_{m}\left(\boldsymbol{r}_{1}, \ldots, \boldsymbol{r}_{N}\right)=\sum_{n=1}^{N}\left(v\left(\boldsymbol{r}_{n}\right)-v^{\prime}\left(\boldsymbol{r}_{n}\right)\right) \Psi_{m}\left(\boldsymbol{r}_{1}, \ldots, \boldsymbol{r}_{N}\right) \tag{4.1.2}
\end{equation*}
$$

We can divide by the wave functions (at points it does not zero) and we obtain $\left(E_{m}-E_{m}^{\prime}\right)=\sum_{n=1}^{N}\left(v\left(\boldsymbol{r}_{n}\right)-v^{\prime}\left(\boldsymbol{r}_{n}\right)\right)$ This clearly cannot hold at all positions unless $v(\boldsymbol{r})-v^{\prime}(\boldsymbol{r})=0$ at any point.

Next, let us observe that if $\Psi_{1}$ and $\Psi_{2}$ are two wave functions having the same density expectation value at each point in space: $\left\langle\Psi_{1}\right| \hat{n}(\boldsymbol{r})\left|\Psi_{1}\right\rangle=\left\langle\Psi_{2}\right| \hat{n}(\boldsymbol{r})\left|\Psi_{2}\right\rangle$ and $\left\langle\Psi_{1}\right| \hat{H}\left|\Psi_{1}\right\rangle<$ $\left\langle\Psi_{2}\right| \hat{H}\left|\Psi_{2}\right\rangle$, then: $\left\langle\Psi_{1}\right| \hat{F}\left|\Psi_{1}\right\rangle<\left\langle\Psi_{2}\right| \hat{F}\left|\Psi_{2}\right\rangle$ and for any other Hamiltonian, $\hat{H}^{\prime}$ this inequality holds as well:

$$
\begin{equation*}
\left\langle\Psi_{1}\right| \hat{H}^{\prime}\left|\Psi_{1}\right\rangle<\left\langle\Psi_{2}\right| \hat{H}^{\prime}\left|\Psi_{2}\right\rangle \tag{4.1.3}
\end{equation*}
$$

These two fact will now be used to prove the first theorem of DFT, due to Hohenberg and Kohn:

THEOREM 31. (Hohenberg-Kohn): If two (non-degenerate) electronic systems exhibit the same ground-state density function then they are identical: their external potentials are the same up to a constant.

Proof. Proof by reductio ad absurdum. Assume $\hat{H}$ and $\hat{H}^{\prime}$ are not identical to within a constant and the two corresponding ground states $\Psi$ and $\Psi^{\prime}$ have the same density expectation value. Then according to the variational principle $\left\langle\Psi^{\prime}\right| \hat{H}^{\prime}\left|\Psi^{\prime}\right\rangle<\langle\Psi| \hat{H}^{\prime}|\Psi\rangle$ and since this inequality holds for any Hamiltonian, we obtain $\left\langle\Psi^{\prime}\right| \hat{H}\left|\Psi^{\prime}\right\rangle<\langle\Psi| \hat{H}|\Psi\rangle$, an absurd result since it contradicts the variational principle.

The Hohenberg-Kohn theorem is a uniqueness theorem. It shows that the ground state density of Hamiltonian $\hat{H}$ is cannot be ground state density of any other Hamiltonian. But does any density $n(\boldsymbol{r})$ correspond to a ground state density of something?

Definition 32. Pure state v-representable densities: A density $n(\boldsymbol{r})$ is called pure state v-representable if it is the ground state density of $N=\int n(\boldsymbol{r} d \boldsymbol{r})$ electrons in some potential $v(\boldsymbol{r})$. In other words, if there exists a potential $v(\boldsymbol{r})$ such that the ground state wave function $\Psi_{g s}$ of $\hat{H}=\hat{T}+\hat{U}+$ $\int v(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d \boldsymbol{r}$ yields this density: $n(\boldsymbol{r})=\left\langle\Psi_{g s}\right| \hat{n}(\boldsymbol{r})\left|\Psi_{g s}\right\rangle$.

### 4.2. Hohenberg-Kohn variational principle

The Hohenberg-Kohn functional. From the HK theorem, it is clear that a v-representable density $n(\boldsymbol{r})$ determines uniquely (up to a constant) the potential $v(\boldsymbol{r})$. This means that we may think of the potential as a functional of the density. We write $v[n](\boldsymbol{r})$. We may then think of the Hamiltonian is a functional of v-representable densities: $\hat{H}[n]$ and they ground state wave function too: $\Psi_{g s}[n]$. We can also see that the 3rd excited state of $\hat{H}[n]$ is a functional of $n(\boldsymbol{r})$. In short, all properties of the system are functionals of v-representable densities.

Since "everything" is a functional of v-representable densities, we can assert that the ground-state kinetic energy $T[n]$ and electron repulsion energy $U[n]$ are such functional and


Figure 4.2.1. Illustration of Definition 33: the left domain is convex while the right domain is not, since a segment joining to of its points is not fully contained within it.
define the Hohenberg-Kohn functional:
(4.2.1) $\quad F_{H K}[n]=T[n]+U[n]=\left\langle\Psi_{g s}[n]\right| \hat{T}+\hat{U}\left|\Psi_{g s}[n]\right\rangle$.
$F_{H K}[n]$ is a universal functional, not limited to any particular molecular system. We have of course no practical way to calculate $F_{H K}[n]$ in general, and despite this progress can be made with the discovery of the Hohenberg-Kohn minimum principle.

The domain of definition of $F_{H K}$, the domain of v-representable densities, is difficult to characterize in a simple way. In fact, and perhaps unexpectedly, this domain is not even convex.[3] Roughly speaking, a convex domain has no holes or insertions in it. If our domain is not convex, we cannot easily search in it or work with it. Here's a formal definition of a convex domain.

Definition 33. Convex domain: a set which includes the segments joining any pair of its points. See Fig. 4.2.1

This means, that if $n_{0}$ and $n_{1}$ are two v-representable densities domain it is not guaranteed that the convex sum $n_{\lambda} \equiv$ $\lambda n_{0}+(1-\lambda) n_{1}$ (where $\lambda$ is some parameter in the range $[0,1]$ ) is in it as well. An example where convexity fails is
produced by considering the case where the two densities are different and yet correspond to degenerate eigenstates of the same Hamiltonian, as we show later.

We ignore for the moment the practical problem of v-representability and present the HK minimum principle, which allows us, in principle, to set up a ground state energy and density search as a minimum, without the need to explicitly solve the Schrödinger equation:

Theorem 34. (Hohenberg-Kohn minimum principle): The ground state density $n_{v}(\boldsymbol{r})$ of $N$ electrons in a potential $v(\boldsymbol{r})$ minimizes the functional:

$$
\begin{equation*}
E_{v, N}[n]=F_{H K}[n]+\int v(\boldsymbol{r}) n(\boldsymbol{r}) d \boldsymbol{r} \tag{4.2.2}
\end{equation*}
$$

defined for all $v$-representable densities $n(\boldsymbol{r})$ with $\int n(\boldsymbol{r}) d \boldsymbol{r}=$ $N$. The ground state energy $E_{g s}$ of this system is equal to the minimum value $E_{v, N}\left[n_{v}\right]$.

Proof. Denote $\hat{H}_{v}=\hat{T}+\hat{U}+\int v(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d \boldsymbol{r}$. Then for any v-rep density $n(\boldsymbol{r}) E_{v, N}[n]=\left\langle\Psi_{g s}[n]\right| \hat{H}_{v}\left|\Psi_{g s}[n]\right\rangle$. Since $\Psi_{g s}\left[n_{v}\right]$ is the ground state of $\hat{H}_{v}$, we find from the variational principle of quantum mechanics:

$$
\begin{equation*}
E_{v, N}[n]=\left\langle\Psi_{g s}[n]\right| \hat{H}_{v}\left|\Psi_{g s}[n]\right\rangle \geq E_{g s}=\left\langle\Psi_{g s}\left[n_{v}\right]\right| \hat{H}_{v}\left|\Psi_{g s}\left[n_{v}\right]\right\rangle=E_{v, N}[n \tag{4.2.3}
\end{equation*}
$$

The Levy-Lieb functional and the variational principle. We would like to use the HK minimum principle as a variational principle. However, since the HK functional is only defined for v-rep densities and the space of v-rep densities is not convex, this proves to be impossible. However we can formally remedy this situation by defining a new functional, denoted $F_{L L}[n]$ within a broader domain yet coinciding with the HK functional for v-rep densities.

The approach of Lieb, followed the Fenchel-Legendre theory (see Appendix A.3) and defined the Legendre transform of the ground state energy function $E_{g s}[v]=\min _{\Psi}\langle\Psi| \hat{T}+\hat{U}+\int v(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d \boldsymbol{r}$

$$
\begin{equation*}
F_{L}[n]=\max _{v}\left\{E_{g s}[v]-\int n(\boldsymbol{r}) v(\boldsymbol{r}) d \boldsymbol{r}\right\} \tag{4.2.4}
\end{equation*}
$$

The functional $F_{L}$ is convex and does not require $n$ to be v-rep. Furthermore, one also has:

$$
\begin{equation*}
E_{g s}[v]=\min _{n}\left\{F_{L}[n]+\int n(\boldsymbol{r}) v(\boldsymbol{r}) d \boldsymbol{r}\right\} \tag{4.2.5}
\end{equation*}
$$

(see proof in sectionComnbination of the two leads to the Levy approach, whiThis functional is defined as:

$$
\begin{equation*}
F_{L L}[n]=\min _{\Psi \rightarrow n}\langle\Psi| \hat{T}+\hat{U}|\Psi\rangle, \tag{4.2.6}
\end{equation*}
$$

where the minimum is taken over $N$-body Fermionic wave functions $\Psi(1, \ldots, N)$ (i.e. antisymmetric to exchange of two particles) that have the density expectation value equal to $n(\boldsymbol{r})$ at each point $\boldsymbol{r}$. Since for v-representable densities $n(\boldsymbol{r})$ the minimizing wave function is $\Psi_{g s}[n]$, we have for this case $F_{L L}[n]=F_{H K}[n]$. However, $F_{L L}$ is defined over a much wider set of densities.

In the sequel we will use the symbol $F[n]$ to mean $F_{L L}$. To minimize $E_{v, N}[n]$, we set up the Lagrangian $L_{v, N}[n]=$ $E_{v, N}[n]-\mu\left(\int n(\boldsymbol{r}) d \boldsymbol{r}-N\right)$, so that the constraint $\int n(\boldsymbol{r}) d \boldsymbol{r}=$ $N$ is incorporated. The variational principle of DFT now becomes an implicit equation for the density $n_{g s}(\boldsymbol{r})$ :

$$
\begin{equation*}
\left(\frac{\delta F}{\delta n(\boldsymbol{r})}\right)_{n_{g s}}+v(\boldsymbol{r})=\mu \tag{4.2.7}
\end{equation*}
$$

Furthermore, derivative of $L$ by the constraint $N$ gives the physical meaning of $\mu$ :

$$
\begin{equation*}
\mu=\frac{\partial E_{v, N}\left[n_{g s}\right]}{\partial N}, \tag{4.2.8}
\end{equation*}
$$

showing that $\mu$ is the chemical potential of the system.
Direct density functional theory. Suppose we have an approximation $F_{\text {app }}[n]$ for $F_{L L}[n]$. Then the ground state energy and density for electrons in a potential $v(\boldsymbol{r})$ can be found by minimizing

$$
E_{v, N}[n]=F_{a p p}[n]+\int v(\boldsymbol{r}) n(\boldsymbol{r}) d \boldsymbol{r}
$$

under the constraint $\int n(\boldsymbol{r}) d^{3} \boldsymbol{r}=N$. The constraint introduces a Lagrange multiplier and the Euler-Lagrange equation for the approximate ground state density $n_{*}(\boldsymbol{r})$ is

$$
\left(\frac{\delta F_{a p p}}{\delta n(\boldsymbol{r})}\right)_{n_{*}}=\mu-v(\boldsymbol{r})
$$

EXERCISE 35. (DFT description of strongly interacting Bosons at high density) Suppose the particles are bosons, interacting through a contact potential $u_{C}\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) \rightarrow c \delta\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right)$ and that due to the high density of the Bosons, the kinetic energy can be neglected. Then, approximately: $F_{\text {app }}[n]=\frac{1}{2} c \int n(\boldsymbol{r})^{2} d r$. Derive the density of the system of bosons. Apply the theory to such bosons in a harmonic potential.

Solution. The Euler-Lagrange equation is

$$
n_{*}(\boldsymbol{r})=s\left(\frac{\mu-v(\boldsymbol{r})}{c}\right) .
$$

where $s(x)=\theta(x) x$.
For $N$ 1D bosons in the potential $v(x)=\frac{1}{2} k x^{2}$ we have, for $-a<x<a$ :

$$
n_{*}(x)=\frac{\mu-\frac{1}{2} k x^{2}}{c}=\frac{k a^{2}}{2} c^{-1}\left(1-\left(\frac{x}{a}\right)^{2}\right)
$$

where $\frac{1}{2} k a^{2}=\mu$ and $a$ is obtained from the condition: $\int_{-a}^{a} n_{*}(x) d x=$ $N$ which gives $a=\left(\frac{3 c}{2 k} N\right)^{1 / 3}$. An example of how well this DFT


Figure 4.2.2. Diffusion Monte Carlo and the DFT estimate of the density of $N=32$ Bose particles having unit mass in a harmonic well with $k=1$ and interacting through the potential $u(r)=c \delta(r)$, with $c=4$. Under these conditions $a=5.768$ and the density maximum is $n_{\max }=\frac{\mu}{c}=4.16$.
approximation does is shown in Fig. 4.2.2 where the DFT density is seen to be close to an essentially exact diffusion Monte Carlo calculation.

Non-interacting electrons. Non-interacting electrons are Fermions which have a density $n(\boldsymbol{r})$ in the ground state of a single body potential well $v_{s}(\boldsymbol{r})$ ( $s$ stands for "single"), i.e. they are ground states of the Hamiltonian of the form: $\hat{H}_{s}=$ $\hat{T}+\int v(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d \boldsymbol{r}$. The ground states of such non-interacting electrons are Slater wave functions $\Phi[n]=\operatorname{det}\left[\phi_{1}, \ldots, \phi_{N}\right]$ where $\phi_{n}(\boldsymbol{r})$ are the $N$ lowest energy eigenstate $\hat{h}_{s} \phi_{n}(\boldsymbol{r})=$ $\varepsilon_{n} \phi_{n}(\boldsymbol{r})$ of the single-particle Hamiltonian $\hat{h}_{s}=-\frac{1}{2} \nabla^{2}+v_{s}(\boldsymbol{r})$. The kinetic energy of the non-interacting electrons in their
ground states is thus

$$
\begin{aligned}
T_{s}[n] & =\langle\Phi[n]| \hat{T}|\Phi[n]\rangle \\
& =\sum_{n=1}^{N}\left\langle\phi_{n}\right|-\frac{1}{2} \nabla^{2}\left|\phi_{n}\right\rangle
\end{aligned}
$$

From the HK minimum principle, we have

$$
\begin{equation*}
\frac{\delta T_{s}}{\delta n(\boldsymbol{r})}+v_{s}(\boldsymbol{r})=\mu_{s} \tag{4.2.9}
\end{equation*}
$$

The functional $T_{s}$ defined in this way is defined for noninteracting v-representable densities. We can extend its domain of definition by defining a LL extension:

$$
T_{s}[n]=\min _{\Psi \rightarrow n}\langle\Psi| \hat{T}|\Psi\rangle
$$

where $\Psi$ is a antisymmetric N -body wave function constraint to reproduce the density expectation value $n(\boldsymbol{r})$. In order to perform the constraint minimization, we introduce Lagrange multipliers $v(\boldsymbol{r})$ to impose the constraints:

$$
L[\Psi, v]=\langle\Psi| \hat{T}|\Psi\rangle+\int v(\boldsymbol{r})[\langle\Psi| \hat{n}(\boldsymbol{r})|\Psi\rangle-n(\boldsymbol{r})] d \boldsymbol{r}
$$

The necessary condition for the minimum is that the following equation be solved $\left(\hat{T}+\int v_{s}(\boldsymbol{r}) \hat{n}(\boldsymbol{r}) d \boldsymbol{r}\right) \Psi_{s}=0$ and then $T_{s}=\left\langle\Psi_{s}\right| \hat{T}\left|\Psi_{s}\right\rangle$. Here's how to find $\Psi_{s}$. For any given $v(\boldsymbol{r})$ we first solve the single particle eigenvalue problem:

$$
\begin{align*}
\hat{h} \phi_{n}(\boldsymbol{r}) & =\varepsilon_{n} \phi_{n}(\boldsymbol{r}) \\
\hat{h} & =-\frac{1}{2} \nabla^{2}+v_{s}(\boldsymbol{r}) \tag{4.2.10}
\end{align*}
$$

We then vary $v_{s}(\boldsymbol{r})$ until we find the potential $v_{s}(\boldsymbol{r})$ and $N$ eigenstates $\phi_{p_{n}}(\boldsymbol{r})$ (not necessarily the lowest energy) for which

$$
\begin{equation*}
n(\boldsymbol{r})=\sum_{n} \phi_{p_{n}}(\boldsymbol{r})^{2} \tag{4.2.11}
\end{equation*}
$$

We then set $\Psi_{s}=\operatorname{det}\left[\phi_{p_{1}} \ldots, \phi_{p_{N}}\right]$ and then $\left[\hat{T}+\int\left[v_{K S}(\boldsymbol{r})-\sum_{n=1}^{N} \varepsilon_{p_{n}}\right] \hat{n}\right.$ 0 . Clearly there are many such solutions. The minimal kinetic energy we are looking for is then found by selecting the solution which gives the the smallest possible $T_{s}=\left\langle\Psi_{s}\right| \hat{T}\left|\Psi_{s}\right\rangle=$ $\sum_{n=1}^{N}\left\langle\phi_{p_{n}}\right|-\frac{1}{2} \nabla^{2}\left|\phi_{p_{n}}\right\rangle$. The Lagrange multipliers $v_{s}(\boldsymbol{r})$ are thus potentials of a non-interacting Schrodinger equation. The main difference in this approach is that we do not select, necessarily the lowest energy solution, but the lowest kinetic energy ones. Finally the Lagrange multipliers are the derivative of the Lagrangian $L$ with respect to the constraint value $n$, analogous to Eq. 4.2.9

$$
\begin{equation*}
\frac{\delta T_{S}[n]}{\delta n(\boldsymbol{r})}=\frac{\delta L\left[\Psi_{K S}, v_{K S}\right]}{\delta n(\boldsymbol{r})}=-v_{s}(\boldsymbol{r}) \tag{4.2.12}
\end{equation*}
$$

### 4.3. The Kohn-Sham density functional theory

The HK functional can be approximated by two functionals we can relatively easily calculate: the kinetic energy of non interacting electrons $T_{s}[n]$ and the Hartree energy. We thus define a new functional $E_{x c}[n]$ which holds all other, as yet unknown quantities in $F$ :

$$
\begin{equation*}
F[n]=T_{s}[n]+E_{H}[n]+E_{x c}[n] \tag{4.3.1}
\end{equation*}
$$

This formalism is based on the hunch that $E_{x c}[n]$ is expected to be a small quantity and much easier to approximated than $F[n]$. The DFT equation (Eq. 4.2.7) becomes: $\frac{\delta T_{s}}{\delta n(\boldsymbol{r})}+v_{H}(\boldsymbol{r})+v_{x c}(\boldsymbol{r})+$ $v(\boldsymbol{r})=\mu$, where $v_{H}(\boldsymbol{r})=\frac{\delta E_{H}}{\delta n(\boldsymbol{r})}$ and $v_{x c}(\boldsymbol{r})=\frac{\delta E_{x c}}{\delta n(\boldsymbol{r})}$. Taking Eq. 4.2 .12 we find that the non-interacting potential is, up to a constant:

$$
\begin{equation*}
v_{s}(\boldsymbol{r})=v_{H}(\boldsymbol{r})+v_{x c}(\boldsymbol{r})+v(\boldsymbol{r}) \tag{4.3.2}
\end{equation*}
$$

This equation leads to the following scheme for solving the DFT equation Eq. 4.2.7: One solves the non-interacting particle eigenstate equation Eq. 4.2 .10 with $v_{s}$ given in the above equation and the density given in Eq. 4.2.11.

Linear response. Linear response of non-interacting electrons. Linear response of interacting electrons.

CHAPTER 5

## Time-dependent DFT

## APPENDIX A

## Befriending Functionals

## A.1. Functionals, linear functionals

Definition 36. A mapping function $n(\boldsymbol{r})$ of a given set to a real number $F[n]$ is called a functional. A functional is called linear if for any two functions $n(\boldsymbol{r})$ and $m(\boldsymbol{r}): F[n+m]=$ $F[n]+F[m]$ and if $F[\alpha n]=\alpha F[n]$ for any number $\alpha$.

Example 37. Here are some examples:
(1) The simplest functional: $F[n]=0$ maps any function $n(r)$ to the number zero. This functional is linear.
(2) A more interesting functional is one that maps any function to its value at a given point $\boldsymbol{r}_{0}$ :

$$
\begin{equation*}
F_{\boldsymbol{r}_{0}}[n]=n\left(\boldsymbol{r}_{0}\right) . \tag{A.1.1}
\end{equation*}
$$

this functional is also linear.
(3) Next, consider the functional

$$
\begin{equation*}
I_{V}[n]=\int_{V} n(\boldsymbol{r}) d^{3} r \tag{A.1.2}
\end{equation*}
$$

mapping each function to its average value on the volume $V$. This functional is linear.
(4) Another familiar example is the kinetic energy functional for a particle of mass $m_{e}$ in some wave function $\psi(r)$ :

$$
T_{1}[\psi]=-\frac{\hbar^{2}}{2 m} \int \psi(\boldsymbol{r})^{*} \nabla^{2} \psi(\boldsymbol{r}) d \boldsymbol{r}
$$

(5) The potential energy:

$$
V[\psi]=\int|\psi(\boldsymbol{r})|^{2} V(\boldsymbol{r}) d \boldsymbol{r}
$$

Theorem 38. A linear functional is generally of the form:

$$
\begin{equation*}
F[n]=\int f(\boldsymbol{r}) n(\boldsymbol{r}) d \boldsymbol{r} \tag{A.1.3}
\end{equation*}
$$

Proof. Define:

$$
\begin{equation*}
f(\boldsymbol{r}) \equiv F\left[\delta_{\boldsymbol{r}}\right] \tag{A.1.4}
\end{equation*}
$$

where

$$
\delta_{r}\left(\boldsymbol{r}^{\prime}\right) \equiv \delta\left(\boldsymbol{r}^{\prime}-\boldsymbol{r}\right)
$$

is the $\delta$-function at $\boldsymbol{r}$. Now write the functional $n\left(\boldsymbol{r}^{\prime}\right)$ as a linear combination of $\delta$-functions $\delta_{\boldsymbol{r}}\left(\boldsymbol{r}^{\prime}\right)$ :

$$
\begin{aligned}
n\left(\boldsymbol{r}^{\prime}\right) & =\int n(\boldsymbol{r}) \delta_{\boldsymbol{r}^{\prime}}(\boldsymbol{r}) d \boldsymbol{r} \\
& =\int n(\boldsymbol{r}) \delta_{\boldsymbol{r}}\left(\boldsymbol{r}^{\prime}\right) d \boldsymbol{r}
\end{aligned}
$$

then, considering $F$ as operating on functions of $\boldsymbol{r}^{\prime}$, from linearity:

$$
\begin{aligned}
F[n] & =\int n(\boldsymbol{r}) F\left[\delta_{\boldsymbol{r}}\right] d \boldsymbol{r} \\
& =\int n(\boldsymbol{r}) f(\boldsymbol{r}) d \boldsymbol{r}
\end{aligned}
$$

Definition 39. The function $f(\boldsymbol{r})$ in Eq. (A.1.3) is called the linear kernel of $F$.

## A.2. Functional derivatives

The concept of functional derivative is highly analogous to the concept of function derivatives. In functions, $f(x)$ we known that near a point $x_{0}$ the function behaves very similar
to a straight line, i.e. $f\left(x_{0}+\delta x\right)-f\left(x_{0}\right)$ is a linear function and the slope of this function is the derivative $\left(\frac{d f}{d x}\right)_{x_{0}}$ of $f$ at $x_{0}$ :

$$
f\left(x_{0}+\delta x\right)-f\left(x_{0}\right) \rightarrow\left(\frac{d f}{d x}\right)_{x_{0}} \delta x \quad(\text { as } \delta x \rightarrow 0) .
$$

In functional analysis we want to achieve a similar feat. Hence for some functional $F[n]$ we take a function $n_{0}(\boldsymbol{r})$ and we want to explore all functions $n=n_{0}+\delta n$ which are close to this function:

Definition 40. Given a functional $F[n]$, we say it is derivable with respect to $n$ at the $n_{0}$, where $n_{0}(\boldsymbol{r})$ is given function, if for any function $\delta n(\boldsymbol{r})$ for which $\|\delta n\|$ goes to zero the functional $\delta F_{n_{0}}[\delta n] \equiv F\left[n_{0}+\delta n\right]-F\left[n_{0}\right]$ goes to a linear functional. Using the standard form of linear functionals, Eq. (A.1.3) we write:
(A.2.1) $\delta F_{n_{0}}[\delta n] \rightarrow \int\left(\frac{\delta F}{\delta n(\boldsymbol{r})}\right)_{n_{0}} \delta n(\boldsymbol{r}) d \boldsymbol{r} \quad($ as $\|\delta n\| \rightarrow 0)$ where $\left(\frac{\delta F}{\delta n(\boldsymbol{r})}\right)_{n_{0}}$ is the linear kernel of $\delta F_{n_{0}}[\delta n]$.

Since the function $\delta n(\boldsymbol{r})$ is arbitrary (except for being very small) it is often convenient to take it as a delta function at $r$, i.e. $\delta n(\boldsymbol{r})=\varepsilon \delta(\boldsymbol{r})$ (and use the $L_{1}$ norm) where $\varepsilon>0$ is as small as needed. Then:

$$
\begin{equation*}
\left(\frac{\delta F}{\delta n(\boldsymbol{r})}\right)_{n_{0}}=\lim _{\varepsilon \rightarrow 0} \varepsilon^{-1} \delta F_{n_{0}}\left[\varepsilon \delta_{\boldsymbol{r}}\right] \tag{A.2.2}
\end{equation*}
$$

EXAMPLE 41. The functional derivative of $I_{\text {all space }}[n]$ (Eq. (A.1.2)):

$$
\begin{array}{rl}
\left(\frac{\delta F}{\delta n(\boldsymbol{r})}\right)_{n_{0}} & N \\
=\lim _{\varepsilon \rightarrow 0} \varepsilon^{-1}\left(F\left[n_{0}+\varepsilon \delta_{\boldsymbol{r}}\right]-F\left[n_{0}\right]\right) \\
& =\lim _{\varepsilon \rightarrow 0} \varepsilon^{-1} \int_{I}\left(n_{0}\left(\boldsymbol{r}^{\prime}\right)+\varepsilon \delta_{\boldsymbol{r}}\left(\boldsymbol{r}^{\prime}\right)-n_{0}\left(\boldsymbol{r}^{\prime}\right)\right) d \boldsymbol{r}^{\prime} \\
& =\int_{I} \delta_{\boldsymbol{r}}\left(\boldsymbol{r}^{\prime}-\boldsymbol{r}\right) d \boldsymbol{r}^{\prime}=1
\end{array}
$$

Exercise 42. Find the functional derivative of the Hartree energy functional:

$$
E_{H}[n]=\frac{1}{2} \iint \frac{n(\boldsymbol{r}) n\left(\boldsymbol{r}^{\prime}\right)}{\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|} d \boldsymbol{r} d \boldsymbol{r}^{\prime} .
$$

Solution. We have applying Eq. (A.2.2) we find:

$$
\begin{aligned}
\left(\frac{\delta E_{H}}{\delta n\left(\boldsymbol{r}_{1}\right)}\right)_{n_{0}} & =\frac{1}{2} \iint \frac{\delta_{\boldsymbol{r}_{1}}(\boldsymbol{r}) n_{0}\left(\boldsymbol{r}^{\prime}\right)+\delta_{\boldsymbol{r}_{1}}\left(\boldsymbol{r}^{\prime}\right) n_{0}(\boldsymbol{r})}{\left|\boldsymbol{r}-\boldsymbol{r}^{\prime}\right|} d \boldsymbol{r} d \boldsymbol{r}^{\prime} \\
& =\int \frac{n_{0}\left(\boldsymbol{r}^{\prime}\right)}{\left|\boldsymbol{r}_{1}-\boldsymbol{r}^{\prime}\right|} d \boldsymbol{r}^{\prime}
\end{aligned}
$$

## A.3. Convex functionals and Legendre transforms

Convex functions. A function $f(x)$ is said to be convex if for any pair of abscissas $x_{1}$ and $x_{2}$ the curve $(x, f(x))$ described by $f(x)$ in the interval $x_{1} \leq x \leq x_{2}$ is always below the straight line connecting $\left(x_{1}, f\left(x_{1}\right)\right)$ and $\left(x_{2}, f\left(x_{2}\right)\right)$. A more formal definition is

Definition 43. A function $f(x)$ is convex in a simply connected region $D^{1}$ if for any two points $x_{1}, x_{2}$ in $D$ and for any $0 \leq \lambda \leq 1$ the following inequality holds:

$$
f\left(\lambda x_{2}+(1-\lambda) x_{1}\right) \leq \lambda f\left(x_{2}\right)+(1-\lambda) f\left(x_{1}\right)
$$

[^2]A useful interpretation of the above definition for convexity is Jensen's inequality which is stated in terms of averaged quantities:

$$
\begin{equation*}
f(\langle x\rangle) \leq\langle f(x)\rangle \tag{A.3.1}
\end{equation*}
$$

where $\langle x\rangle=\sum_{n=1}^{N} c_{n} x_{n}$ is an average over $N$ points $x_{1}, \ldots, x_{N}$ in the interval and $\langle f(x)\rangle=\sum_{n=1}^{N} c_{n} f\left(x_{n}\right)$ an average of the corresponding function values and the $N$ non-negative weights $c_{1}, \ldots, c_{N}$ sum to $1\left(\sum_{n} c_{n}=1\right)$.

Lemma 44. For a derivable convex function $f(x)$ in a simply connected domain $D$ :
(A.3.2)

$$
\left(x_{1}-x_{2}\right) \cdot \nabla f\left(x_{2}\right) \leq f\left(x_{1}\right)-f\left(x_{2}\right) \leq\left(x_{1}-x_{2}\right) \cdot \nabla f\left(x_{1}\right)
$$

Proof. Consider the equality

$$
f\left(\lambda x_{1}+(1-\lambda) x_{2}\right)=f\left(\lambda\left(x_{1}-x_{2}\right)+x_{2}\right) .
$$

Now take $0<\lambda \ll 1$. Then: $f\left(x_{2}\right)+\lambda\left(x_{1}-x_{2}\right) \cdot \nabla f\left(x_{2}\right) \leq$ $\lambda f\left(x_{1}\right)+(1-\lambda) f\left(x_{2}\right)$ from which $\left(x_{1}-x_{2}\right) \cdot \nabla f\left(x_{2}\right) \leq f\left(x_{1}\right)-$ $f\left(x_{2}\right)$. Now, switch $1 \longleftrightarrow 2$, so $\left(x_{2}-x_{1}\right) \cdot \nabla f\left(x_{1}\right) \leq f\left(x_{2}\right)-$ $f\left(x_{1}\right)$ and multiplication by -1 gives: $f\left(x_{1}\right)-f\left(x_{2}\right) \leq\left(x_{1}-x_{2}\right)$. $\nabla f\left(x_{1}\right)$. The two inequalities are combined in Eq. A.3.2

Lemma 45. A convex function which is twice derivable has a positive definite Hessian derivative.

Proof. Assume that for a some $x, f_{i j}(x)$ is not positive definite. Thus assume $\delta x_{i} f_{i j}(x) \delta x_{j}<0$ for some infinitesimal $\delta x$. Now, for any $0<\lambda<1$ use the identity $f(x+\lambda \delta x)=$ $f((x+\delta x) \lambda+(1-\lambda) x)$ to deduce:

$$
f(x+\lambda \delta x) \leq \lambda f(x+\delta x)+(1-\lambda) f(x)
$$

Then, we expand in powers of $\delta x$ and to second order we see the left hand side equals to $f(x)+\lambda f_{i}(x) \delta x_{i}+\frac{1}{2} \lambda^{2} \delta x_{i} f_{i j}(x) \delta x_{j}$ while the right hand side gives $f(x)+\lambda\left[\delta x_{i} f_{i}(x)+\frac{1}{2} \delta x_{i} f_{i j}(x) \delta x_{j}\right]$, hence the above inequality leads to $\lambda^{2} \delta x_{i} f_{i j}(x) \delta x_{j} \leq \lambda \delta x_{i} f_{i j}(x) \delta x_{j}$.

Since $\lambda>0$, divide by the negative number $\lambda \delta x_{i} f_{i j}(x) \delta x_{j}$ and obtain $\lambda>1$, which is a contradiction to the fact that $\lambda$ is smaller than 1.

One of the important properties of a convex function is that it cannot have a local minimum: either there is no minimum (for example $f(x)=e^{x}$ ) or there is just one global minimum (for example $f(x)=x^{2}$. The proof is that the derivative is always increasing as $x$ increases. Thus if we are at a minimum, where the derivative is zero, then moving away from this minimum in a certain direction the derivative will always increase or decrease and will not be zero again.

## Legendre transform of a functional.

Definition 46. The Legendre transform of a functional $f(x)$ defined in some domain $D \subset \Re^{N}$ is

$$
\begin{equation*}
f^{*}(y)=\sup _{x \in D}\left\{y^{T} x-f(x)\right\} \tag{A.3.3}
\end{equation*}
$$

defined for $y \in \Re^{N}$.
Lemma 47. (Fenchel's inequality) $y^{T} x \leq f(x)+f^{*}(y)$
Proof. By definition $f^{*}(y) \geq x^{T} y-f(x)$.
Lemma 48. The Legendre transform $f^{*}(y)$ of $f(x)$ is a convex function.

Proof. Relies on the obvious fact that the maximum of a sum is not larger than the sum of the maxima. Hence, for $0 \leq \lambda \leq 1$ :

$$
\begin{aligned}
f^{*}(\lambda y+(1-\lambda) z) & =\sup _{x \in D}\left\{\left(\lambda y^{T}+(1-\lambda) z^{T}\right) x-f(x)\right\} \\
& \leq \lambda \sup _{x \in D}\left\{y^{T} x-f(x)\right\}+(1-\lambda) \sup _{x \in D}\left\{\left(z^{T} x-f(x)\right)\right\} \\
& =\lambda f^{*}(y)+(1-\lambda) f^{*}(z)
\end{aligned}
$$

Lemma 49. Let $f(x)$ be differentiable, Then the Legendre transform $f^{*}(y)$ is differentiable. Furthermore, if $x(y)$ maximizes $y^{T} x-f(x)$ then: $f_{j}(x(y))=y_{j}$ and $f_{j}^{*}(y)=x_{j}(y)$.

Proof. Suppose $x(y)$ is the maximizer of $y^{T} x-f(x)$. Then $y_{j}=f_{j}(x(y))$ and from $f^{*}(y)=y^{T} x(y)-f(x(y))$ we find: $f_{j}^{*}(y)=x_{j}(y)$.

Note that if $x(y)$ is the point maximizing the RHS we have

$$
\begin{gathered}
\left(\frac{\partial f(x)}{\partial x_{i}}\right)_{x=x(y)}=y_{i} \\
f^{*}(y)=y^{T} x(y)-f(x(y))
\end{gathered}
$$

From this, we get a kind of symmetric relation:

$$
\frac{\partial f^{*}(y)}{\partial y_{j}}=x_{j}(y)
$$

Furthermore:

$$
\begin{aligned}
\delta_{i j} & =\frac{\partial}{\partial y_{j}} y_{i}=\frac{\partial}{\partial x_{i}} \frac{\partial}{\partial y_{j}} f(x(y)) \\
& =\frac{\partial}{\partial x_{i}} \frac{\partial}{\partial x_{l}} f(x(y)) \frac{\partial x_{l}}{\partial y_{j}}
\end{aligned}
$$

Showing that $\frac{\partial x_{l}}{\partial y_{j}}=\left(H^{-1}\right)_{i j}$ is the inverse of the Hessian $H_{i l}=$ $\frac{\partial}{\partial x_{i}} \frac{\partial}{\partial x_{l}} f(x)$ at $x=x(y)$.

ExAMPLE. Here are some examples:
(1) $f(x)=\frac{1}{n}|x|^{n}(n>1) . f^{*}(y)=\max _{x \in D}\left\{y^{T} x-\frac{1}{n}|x|^{n}\right\}$ . The minimizing $x^{*}$ obeys: $y=\nabla f\left(x_{*}\right)^{T}=\left|x_{*}\right|^{n-2} x_{*}$ thus: $|y|=\sqrt{y^{T} y}=\left|x_{*}\right|^{n-1}$. Hence: $f^{*}(y)=\left(1-\frac{1}{n}\right)\left|x_{*}\right|^{n}=$ $\left(1-\frac{1}{n}\right)|y|^{\frac{n}{n-1}}=\frac{1}{m}|y|^{m}$ where $\frac{1}{n}+\frac{1}{m}=1$.
(2) From this, the convex conjugate of $f(x)=|x|$ is $f^{*}(y)=\left\{\begin{array}{cc}0 & |y| \leq 1 \\ \infty & |y|>1\end{array}\right.$.
(3) $f(x)=e^{x}$. $f^{*}(y)=\sup _{x}\left\{y^{T} x-e^{x}\right\}$. Note that for $y=0$ then $f^{*}$ is the minimum of $e^{x}$ i.e. 0 . If $y<0$ then by taking $x \rightarrow-\infty$ we obtain $f^{*}=\infty$. For $y>0$ the minimizing $x_{*}$ obeys: $y=e^{x_{*}}$ and from this $x_{*}=\log y$ so $f^{*}(y)=y(\log y-1)$. Summarizing: $f^{*}(y)=\left\{\begin{array}{cc}y(\log y-1) & y>0 \\ 0 & y=0 \\ \infty & y<0\end{array}\right.$
Theorem 50. Given $f(x), f^{*}(y)=\sup _{x \in D}\left\{y^{T} x-f(x)\right\}$ the Legendre transform of $f$ and $f^{* *}(z)=\sup _{y}\left\{z^{T} y-f^{*}(y)\right\}$ then $f(x)=f^{* *}(x)$ iff $f(x)$ is a convex function.

Proof. Let $x(y)$ be the maximizing $x$ in $\sup _{x \in D}\left\{y^{T} x-f(x)\right\}$. Then $y_{j}=f_{j}(x(y))$. Also: $f^{*}(y)=y^{T} x(y)-f(x(y))$ so $f_{j}^{*}(y)=x_{j}(y)$. Finally, $y_{j}=f_{j}^{* *}(x(y)) y_{j}=f_{j}(x(y))$. Clearly $\delta y_{j}=f_{j k}(x(y)) x_{k l}(y) \delta y_{l}$ Also $z_{k}=f_{k}^{*}(y(z))$. Now, $f^{*}(y+\delta y)>$ $\sup _{x \in D}\left\{y^{T} x-f(x)\right\}+\sup _{x \in D}\left\{\delta y^{T} x-f(x)\right\}$

Then $f^{* *}(z)=z^{T} y(z)-\left(y(z)^{T} x(y(z))-f(x(y(z)))\right)=$ $\left(z^{T}-x(y(z))\right) y(z)+f(x(y(z)))$

Convex functionals and Legendre transforms. A similar definition of convexity can be applied to functionals. Functional $F[n]$ is convex if

$$
F\left[\lambda n_{2}+(1-\lambda) n_{1}\right] \quad \leq \quad \lambda F\left[n_{2}\right]+(1-\lambda) F\left[n_{1}\right]
$$

The following holds in analogy to Eq. A.3.2:
(A.3.4)
$\left.\int\left(n_{1}(x)-n_{2}(x)\right) \frac{\delta F[n]}{\delta n(x)}\right|_{n_{2}} d x \leq F\left[n_{1}\right]-F\left[n_{2}\right]$

$$
\begin{equation*}
\leq\left.\int\left(n_{1}(x)-n_{2}(x)\right) \frac{\delta F[n]}{\delta n(x)}\right|_{n_{1}} d x \tag{A.3.5}
\end{equation*}
$$

A similar condition exists for the second functional derivative. The Jensen inequality holds here as well: $F[\langle n\rangle] \leq\langle F[n]\rangle$.

EXAMPLE 51. Consider the functional $Q[V]=\int_{\Omega} e^{-\beta V(x)} d x$ where $\Omega$ is some finite spatial domain. This is convex inheriting the property from $f(V)=e^{V}$ :

$$
\begin{aligned}
Q\left[\lambda V_{1}+(1-\lambda) V_{2}\right] & =\int_{\Omega} e^{-\beta\left(\lambda V_{1}(x)+(1-\lambda) V_{2}(x)\right.} d x \\
& \leq \int_{\Omega}\left[\lambda e^{-\beta V_{1}(x)}+(1-\lambda) e^{-\beta V_{2}(x)}\right] d x \\
& =\lambda Q\left[V_{1}\right]+(1-\lambda) Q\left[V_{2}\right]
\end{aligned}
$$

Definition 52. The Legendre transform of a function $F[n]$ defined in some function domain is

$$
\begin{equation*}
F^{*}[m]=\max _{n \in D}\left\{\int n(x) m(x) d x-F(n)\right\} \tag{A.3.6}
\end{equation*}
$$

defined for functions $m$. We note that the domain of definition of $F^{*}$ may be much larger than $D$, the domain of definition of $F$.

The Legendre transform $F^{*}[m]$ is convex in its domain, even if the functional $F[n]$ is not. The proof is similar to that of Lemma 48. Also, the Fenchel inequality becomes:

$$
\begin{equation*}
\int m(x) n(x) d x \leq F(n)+F^{*}[m] \tag{A.3.7}
\end{equation*}
$$

## APPENDIX B

## The virial theorem

## B.1. Dilation relations

Consider the normalized wave function $\psi(x)$ of a particle of mass $m_{e}$ in a potential $V(x)$. From it, we can define a family of normalized wave functions:

$$
\begin{equation*}
\psi_{\gamma}(x)=\gamma^{1 / 2} \psi(\gamma x) \tag{B.1.1}
\end{equation*}
$$

We can check that each member is indeed normalized by replacing variable in the normalization integral $y=\gamma x$ :

$$
\begin{aligned}
\left\langle\psi_{\gamma} \mid \psi_{\gamma}\right\rangle & =\int \gamma \psi(\gamma x)^{2} d x \\
& =\int \psi(y)^{2} d y \\
& =1
\end{aligned}
$$

This operation of "stretching" ( $\gamma<1$ ) or "compressing" $(\gamma>1)$ the argument of a function is called $a$ dilation. Dilations affect functionals. Consider the kinetic energy functional:

$$
T[\psi]=-\frac{\hbar^{2}}{2 m} \int \psi(x)^{*} \psi^{\prime \prime}(x) d x
$$

In order to calculate the effect of the dilation on the functional prove that:

$$
\psi_{\gamma}^{\prime \prime}(x)=\gamma^{5 / 2} \psi^{\prime \prime}(\gamma x)
$$

Then, using this prove that:

$$
\begin{equation*}
T\left[\psi_{\gamma}\right]=\gamma^{2} T[\psi] \tag{B.1.2}
\end{equation*}
$$

This transform property of the kinetic energy carries over to many body wave functions, where the dilation is uniform:
(B.1.3) $\quad \Psi_{\gamma}\left(\boldsymbol{x}_{1}, \boldsymbol{x}_{2}, \ldots, \boldsymbol{x}_{N}\right)=\gamma^{3 N / 2} \Psi\left(\gamma \boldsymbol{x}_{1}, \gamma \boldsymbol{x}_{2}, \ldots, \gamma \boldsymbol{x}_{N}\right)$

Consider a potential which is a homogeneous function of order $M$, i.e. has the property:

$$
\begin{equation*}
V\left(\gamma \boldsymbol{x}_{1}, \gamma \boldsymbol{x}_{2}, \ldots, \gamma \boldsymbol{x}_{N}\right)=\gamma^{M} V\left(\boldsymbol{x}_{1}, \boldsymbol{x}_{2}, \ldots, \boldsymbol{x}_{N}\right) \tag{B.1.4}
\end{equation*}
$$

Then the corresponding potential functional
$V[\Psi]=\int \Psi\left(\boldsymbol{x}_{1}, \boldsymbol{x}_{2}, \ldots, \boldsymbol{x}_{N}\right)^{2} V\left(\boldsymbol{x}_{1}, \boldsymbol{x}_{2}, \ldots, \boldsymbol{x}_{N}\right) d^{3} x_{1} d^{3} x_{2} \cdots d^{3} x_{N}$
transforms under dilation in the following way:

$$
\begin{equation*}
V\left[\Psi_{\gamma}\right]=\gamma^{-M} V[\Psi] \tag{B.1.5}
\end{equation*}
$$

For the Harmonic oscillators $M=2$ and $V\left[\Psi_{\gamma}\right]=\gamma^{-2} V[\Psi]$. For Coulomb interaction $M=-1$ so $V\left[\Psi_{\gamma}\right]=\gamma V[\Psi]$.

The Hamiltonian of Homogeneous potential then has the interesting property:

$$
\begin{equation*}
E[\Psi]=T[\Psi]+V[\Psi] \tag{B.1.6}
\end{equation*}
$$

then:

$$
\begin{equation*}
E\left[\Psi_{\gamma}\right]=\gamma^{2} T[\Psi]+\gamma^{-M} V[\Psi] \tag{B.1.7}
\end{equation*}
$$

## B.2. The virial theorem - classical mechanics

First, let us define the virial. For a system with coordinate $x_{n}$ collectively denoted as $\boldsymbol{x}$ the virial in classical mechanics is the time average of $\boldsymbol{x} \cdot \boldsymbol{F}$ where $\boldsymbol{F}$ is the force vector:

$$
\text { virial }=\langle\boldsymbol{x} \cdot \boldsymbol{F}\rangle \equiv \frac{1}{\tau} \int_{-\tau / 2}^{+\tau / 2} \boldsymbol{x}(t) \cdot \boldsymbol{F}(t) d t
$$

Here, $\boldsymbol{x}$ are the Cartesian coordinates of the system particles and $\boldsymbol{F}$ the forces. The average is a long time trajectory. It is assumed that the trajectory is so long that it has forgot it's starting point (except that the total energy and momentum are preserved). An axiomatic assumption (called that is that
the time average is equal to the thermodynamics average over an ensemble. For example, if the kinetic energy is $\langle T\rangle$ then the ensemble average can be take as the Canonical ensemble average at inverse temperature $\beta=\frac{3 N}{2\langle T\rangle}$. For forces derived from potentials:

$$
\boldsymbol{F}=-\frac{d V(\boldsymbol{x})}{d \boldsymbol{x}}
$$

The viral is the canonical average:

$$
\begin{aligned}
\text { virial } & =\langle\boldsymbol{x} \cdot \boldsymbol{F}\rangle \\
& =\frac{1}{Q(\beta)} \int e^{-\beta V(\boldsymbol{x})}\left(\boldsymbol{x} \cdot \frac{-d V(\boldsymbol{x})}{d \boldsymbol{x}}\right) d \boldsymbol{x}
\end{aligned}
$$

where

$$
Q(\beta)=\int e^{-\beta V(\boldsymbol{x})} d \boldsymbol{x}
$$

EXERCISE 53. Prove by partial integration and using the fact that $\lim _{|\boldsymbol{x}| \rightarrow \infty} e^{-\beta V(\boldsymbol{x})}=0$ that

$$
\text { virial }=-3 N \beta^{-1}
$$

From this deduce that

$$
\text { virial }=-2\langle T\rangle .
$$

The virial is related to a dilations:
Exercise 54. Euler's Lemma: Prove that for systems with potential which is homogeneous of order $M$ (see Eq. B.1.5)

$$
\boldsymbol{x} \cdot \frac{d V(\boldsymbol{x})}{d \boldsymbol{x}}=M V(\boldsymbol{x})
$$

Deduce from this that for such systems:

$$
\text { virial }=-M\langle V\rangle
$$

It thus follows that:
Theorem 55. The virial theorem (classical mechanics): For systems with homogeneous potentials of order $M$

$$
\begin{equation*}
2\langle T\rangle-M\langle V\rangle=0 \tag{B.2.1}
\end{equation*}
$$

For purely Coulombic systems, $M=-1$ and one must have (B.2.2)

$$
2\langle T\rangle=-\langle V\rangle
$$

## B.3. The virial theorem - quantum mechanics

Consider a Hamiltonian $\hat{H}=\hat{T}+\hat{V}$ where $\hat{V}=V\left(\boldsymbol{x}_{1}, \ldots, \boldsymbol{x}_{N}\right)$ is homogeneous of order $M$. Then the GS wave function $\Psi^{G S}$ minimizes the energy functional Eq. B.1.6. Thus the scaled wave function, defined by

$$
\begin{equation*}
\Psi_{\lambda}\left(\boldsymbol{x}_{1}, \ldots, \boldsymbol{x}_{N}\right)=\gamma^{3 N / 2} \Psi^{G S}\left(\gamma \boldsymbol{x}_{1}, \ldots, \gamma \boldsymbol{x}_{N}\right) \tag{B.3.1}
\end{equation*}
$$

has the energy functional Eq. B.1.6 given by (see Eq. B.1.7

$$
\begin{equation*}
E\left[\Psi_{\gamma}\right]=\gamma^{2} T\left[\Psi_{G S}\right]+\gamma^{-M} V\left[\Psi_{G S}\right] \tag{B.3.2}
\end{equation*}
$$

Since the energy is minimized when $\gamma=1$ we have

$$
\begin{equation*}
\left(\frac{d}{d \gamma} E\left[\Psi_{\gamma}\right]\right)_{\gamma=1}=0 \tag{B.3.3}
\end{equation*}
$$

and therefore:

$$
\begin{equation*}
0=2 T\left[\Psi_{G S}\right]-M V\left[\Psi_{G S}\right] \tag{B.3.4}
\end{equation*}
$$

which shows that Theorem 55 also holds in quantum mechanics.

For Coulomb potentials $M=-1$ we have

$$
\begin{align*}
E_{G S} & =T\left[\Psi_{G S}\right]+V\left[\Psi_{G S}\right]  \tag{B.3.5}\\
& =\frac{V\left[\Psi_{G S}\right]}{2}  \tag{B.3.6}\\
& =-T\left[\Psi_{G S}\right] . \tag{B.3.7}
\end{align*}
$$

A subtle but important issue: all energies and potentials in the above expressions are absolute. We usually give energy and potentials only to within an additive constant. However, the fact that the potential is homogeneous, it cannot tolerate addition of a constant ( $x^{2}$ is homogeneous of order 2 but $x^{2}+a$ is not).

## B.4. Other Dilation Facts

Consider the Schrödinger equation:

$$
\begin{equation*}
-\frac{\hbar^{2}}{2 m_{e}} \psi^{\prime \prime}(x)+V(x) \psi(x)=E \psi(x) . \tag{B.4.1}
\end{equation*}
$$

Evaluate at the point $\lambda x$ :

$$
\begin{equation*}
-\frac{\hbar^{2}}{2 m_{e}} \psi^{\prime \prime}(\gamma x)+V(\gamma x) \psi(\gamma x)=E \psi(\gamma x) . \tag{B.4.2}
\end{equation*}
$$

Then define $\psi_{\gamma}(x)=\psi(\gamma x)$ and using $\psi^{\prime \prime}(\gamma x)=\frac{1}{\gamma^{2}} \psi_{\gamma}^{\prime \prime}(x)$ :

$$
\begin{equation*}
-\frac{\hbar^{2}}{2 m_{e}} \psi_{\gamma}^{\prime \prime}(x)+\gamma^{2} V(\gamma x) \psi_{\gamma}(x)=\gamma^{2} E \psi_{\gamma}(x) \tag{B.4.3}
\end{equation*}
$$

We thus see that the SE is invariant under the dilation transformation composed of three steps:

$$
\begin{equation*}
v(x) \rightarrow v_{\gamma}(x)=\gamma^{2} v(\gamma x) \tag{B.4.4}
\end{equation*}
$$

$$
\begin{equation*}
E \rightarrow E_{\gamma}=\gamma^{2} E \tag{B.4.5}
\end{equation*}
$$

$$
\begin{equation*}
\psi(x) \rightarrow \psi_{\gamma}(x)=\psi(\gamma x) \tag{B.4.6}
\end{equation*}
$$

Now, apply this for Coulomb systems, where the potential is homogeneous of order $M=-1$. Then

$$
\begin{align*}
v(x) & \rightarrow v_{\gamma}(x)=\gamma v(x)  \tag{B.4.7}\\
E & \rightarrow E_{\gamma}=\gamma^{2} E  \tag{B.4.8}\\
\psi(x) & \rightarrow \psi_{\gamma}(x)=\psi(\gamma x) \tag{B.4.9}
\end{align*}
$$

In other words, we have:
Theorem 56. In a purely Coulombic system if the potential is scaled by a factor $\gamma$ then the energy is scaled by a factor $\gamma^{2}$ and the wave function is dilated by a factor $\gamma$. Furthermore the kinetic energy is scaled by a factor $\gamma^{2}$ as well.

Proof. The only thing left to prove is the last statement concerning the kinetic energy. This follows from the fact that $E$ scales as $\gamma^{2}$ and from the virial theorem by which $T=-E$.

EXERCISE 57. An application of the above discussion allows an interesting exact conclusion concerning the groundstate energy per particle $\varepsilon\left(k_{F}, \kappa\right)$ of the homogeneous electron gas of density $n$. Here the parameter $k_{F}=\left(3 \pi^{2} n\right)^{1 / 3}$ is the Fermi momentum (see Eq. 2.1.7) and the Coulomb coupling strength is $\kappa=\frac{e^{2}}{4 \pi \epsilon_{0}}$. Show that:

$$
\begin{equation*}
\frac{\partial \log \varepsilon\left(k_{F}, e^{2}\right)}{\partial \log k_{F}}+\frac{\partial \log \varepsilon\left(k_{F}, \kappa\right)}{\partial \log \kappa}=2 \tag{B.4.10}
\end{equation*}
$$

Hint: Start from $\varepsilon\left(\lambda k_{F}, \lambda \kappa\right)=\lambda^{2} \varepsilon\left(k_{F}, \kappa\right)$. Then take the derivative with respect to $\lambda$.

## APPENDIX C

## Minimization Theory

The minimization of multidimensional functions is important for electronic structure because of the variational theorem of quantum mechanics and the variational theorem of density functional theory. Here we outline the main results which are used throughout the book.

## C.1. Necessary conditions for minimization

We start with the problem of minimization of a simple 1D function. Given a function $\mathrm{f}(\mathrm{x})$, we want to find a point $x_{*}$ such that the function is minimal. It is clear that the slope of $f(x)$ at $x=x_{*}$ must be zero. If the slope is positive then we can go left (decrease $x_{*}$ ) and reduce $f(x)$. Same logic (but to right) if it was negative. Thus, a necessary condition for a minimum is:

$$
f^{\prime}\left(x_{*}\right)=0
$$

Let us expand in a Taylor's series the function around the point $x_{*}$. Clearly, we have: $f(x)=f\left(x_{*}\right)+\frac{1}{2} f^{\prime \prime}\left(x_{*}\right)\left(x-x_{*}\right)^{2}+$ $\cdots$ where the linear term, proportional to $f^{\prime}\left(x_{*}\right)$ is zero. If we assume that $f^{\prime \prime}\left(x_{*}\right) \neq 0$ we find that when $x$ is extremely close to $x_{*} f(x)$ is nearly a parabola and it is "smiling" when $f^{\prime \prime}\left(x_{*}\right)>0$, i.e. has a minimum and "sad" when $f^{\prime \prime}\left(x_{*}\right)<0$ i.e. has a maximum. Thus, the 2 nd necessary condition for a minimum is:

$$
f^{\prime \prime}(x)>0 .
$$

Now, let us consider the case of functions of two variables, $f(\boldsymbol{r})$ where $\boldsymbol{r}=\binom{x}{y} \equiv\left(\begin{array}{ll}x & y\end{array}\right)^{T}$. Notice that we use the "transpose" symbol superscript- $T$ to turn a column vector into a row vector. Let us Taylor-expand $f(\boldsymbol{r})$ around a point 2D space $\boldsymbol{r}_{*}$, using the notation $\nabla f\left(\boldsymbol{r}_{*}\right) \equiv\left(\begin{array}{cc}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{array}\right)_{\boldsymbol{r}_{*}} \equiv \boldsymbol{g}_{*}^{T}$ as the gradient of $f$ and $\nabla^{T} \nabla f\left(\boldsymbol{r}_{*}\right)=\left(\begin{array}{cc}\frac{\partial^{2} f}{\partial x^{2}} & \frac{\partial^{2} f}{\partial y \partial x} \\ \frac{\partial^{2} f}{\partial x \partial y} & \frac{\partial^{2} f}{\partial y^{2}}\end{array}\right)_{\boldsymbol{r}_{*}}=K_{*}$ as the symmetric "Hessian" of $f$ :

$$
\begin{aligned}
f(\boldsymbol{r}) & =f\left(\boldsymbol{r}_{*}\right)+\nabla f\left(\boldsymbol{r}_{*}\right)\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right) \\
& +\frac{1}{2}\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)^{T} \nabla^{T} \nabla f\left(\boldsymbol{r}_{*}\right)\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)+\cdots \\
& =f\left(\boldsymbol{r}_{*}\right)+\boldsymbol{g}_{*}^{T}\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)+\frac{1}{2}\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)^{T} K_{*}\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)+\cdots
\end{aligned}
$$

When $\boldsymbol{r}_{*}$ is a minimum, moving infinitesimally away from it in any direction will not change the function in a linear manner. Why? If the function decreases moving from $\boldsymbol{r}_{*}$ in some direction $\boldsymbol{d}$, no matter how small the step size, then this contradicts that $\boldsymbol{r}_{*}$ is a minimum. If the function increases linearly in direction $\boldsymbol{d}$ then due to continuity it must decrease in the direction $-\boldsymbol{d}$, again, contradicting the assertion that $\boldsymbol{r}_{*}$ is a minimum. Hence the function does not change linearly in any direction. In other words, the gradient $\boldsymbol{g}_{*}$ must be zero. The the Taylor expansion becomes: $f(\boldsymbol{r})=f\left(\boldsymbol{r}_{*}\right)+\frac{1}{2}\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)^{T} K_{*}\left(\boldsymbol{r}-\boldsymbol{r}_{*}\right)+$ $\cdots$ which when limited to a very small surrounding of $\boldsymbol{r}_{*}$ is the equation of a parabola. In order for $f(\boldsymbol{r})$ to have a minimum at $\boldsymbol{r}_{*}$ the second term on the right hand side must always be positive.

Definition 58. A matrix $K$ for which $\boldsymbol{x}^{T} K \boldsymbol{x}$ is always positive for any $\boldsymbol{x} \neq 0$ is called a positive definite matrix in this case we write " $K>0$ "

Since our Hessian is symmetric, we infer from the following lemma that all its eigenvalues are positive:

LEMMA 59. A symmetric matrix is PD if and only if all its eigenvalues are positive.

The necessary conditions for a minimum of multivariable functions $f(\boldsymbol{r})$ are:

$$
\begin{aligned}
\boldsymbol{g}_{*}^{T} & =\nabla f\left(\boldsymbol{r}_{*}\right)=0 \\
K & =\nabla^{T} \nabla f\left(\boldsymbol{r}_{*}\right)>0
\end{aligned}
$$

ExAMPLE 60. As an example, let us take the function $f(x)=$ $x^{2}+y^{2}$. This function is zero at the origin and non-negative elsewhere. So its minimum is at $x=y=0$. The gradient is $\nabla f(x)=2\left(\begin{array}{ll}x & y\end{array}\right)$ is indeed zero at the origin. The Hessian is $K=2\left(\begin{array}{ll}1 & 0 \\ 0 & 1\end{array}\right)$, hence the eigenvalues are real and it is PD.

## C.2. Constrained minimization

Suppose we want to minimize $f(x, y)$, under a constraint that $y=g(x)$. We can transform the problem into a minimization of a function of single variable: minimize the function $F(x) \equiv f(x, g(x))$. The derivative with respect to $x$ must be zero at $x_{*}$. We use the chain rule:

$$
0=F^{\prime}\left(x_{*}\right)=\left(\frac{\partial f}{\partial x}\right)_{\left(x_{*}, g\left(x_{*}\right)\right)}+\left(\frac{\partial f}{\partial y}\right)_{\left(x_{*}, g\left(x_{*}\right)\right)} g^{\prime}\left(x_{*}\right)
$$

and we can do a similar job for the second derivative. Finding equations for $x_{*}$.

But sometimes it is not possible to write the constraint directly as $y=g(x)$. A more general form is: $h(x, y)=0$. This is also more symmetric. Consider the situation of minimizing $f(x, y)$ under the constraint $h(x, y)=0$. An example is shown in Figure C.2.1 where contours of $f$ and $h$ are shown, including the $h=0$ contour. The minimum of $f$ on this contour is
depicted by a black point at $\boldsymbol{r}_{*}=\left(x_{*}, y_{*}\right)$. This means that moving from that point slightly along the $h=0$ contour $\boldsymbol{r}_{*}$ will not change $f$ in a linear way, for if $f$ decreases this is not a minimum and if $f$ increases we will move in the opposite direction and f will necessarily decrease. Thus $\nabla f\left(\boldsymbol{r}_{*}\right)$, must be normal to the line $h=0$. Furthermore, $\nabla h\left(\boldsymbol{r}_{*}\right)$ must also be normal to this contour. Thus, the necessary condition for minimum is that both vectors point to the same direction and so are proportional. We denote the proportionality constant by $\lambda_{*}$ :

$$
\begin{equation*}
\nabla f\left(\boldsymbol{r}_{*}\right)=\lambda_{*} \nabla h\left(\boldsymbol{r}_{*}\right) \tag{C.2.1}
\end{equation*}
$$

$$
\begin{equation*}
h\left(\boldsymbol{r}_{*}\right)=0 \tag{C.2.2}
\end{equation*}
$$

Lagrange noted that both these equations are obtainable by minimizing the function

$$
\begin{equation*}
L(\boldsymbol{r}, \lambda)=f(\boldsymbol{r})-\lambda h(\boldsymbol{r}) \tag{C.2.3}
\end{equation*}
$$

with respect to $r$ and maximizing it with respect to $\lambda$.
EXAMPle 61. Consider minimizing $f(x, y)=x+y$ under the constraint function $h(x, y)=x^{2}+y^{2}-1$ (i.e. find the point on the unit circle for which $x+y$ is minimal). Then the Lagrangian is

$$
\begin{equation*}
L(x, y)=(x+y)-\lambda\left(x^{2}+y^{2}-1\right) \tag{C.2.4}
\end{equation*}
$$

Then
(C.2.5) $\quad \nabla L=\left((1-2 \lambda x),(1-2 \lambda y),\left(x^{2}+y^{2}-1\right)\right)$

Equating to $(0,0,0)$ gives

$$
\begin{equation*}
1=2 \lambda_{*} x_{*}=2 \lambda_{*} y_{*} \tag{C.2.6}
\end{equation*}
$$

and $x^{2}+y^{2}=1$ from which we deduce that $\lambda_{*}$ is finite and $x_{*}=y_{*}$ and so $x_{*}=y_{*}= \pm \frac{1}{\sqrt{2}}$ and $\lambda_{*}=\frac{1}{2 x_{*}}= \pm \frac{1}{\sqrt{2}}$,

Since these are necessary conditions, we need to consider them further (they might correspond to a maximum of $f$ ). The minimum corresponds only to the negative solution. The


Figure C.2.1. Contours of the function $f(x, y)$ and the constraints $h(x, y)$.The gradients $\nabla f$ and $\nabla g$ are orthogonal to the contours and point in direction of function ascent. The black point is the location of $\left(x_{*}, y_{*}\right)$ which minimize $f(x, y)$ under the constraint that $h(x, y)=0$.
method of Lagrange multipliers is often more convenient to work with than direct replacement. The problem is thus that of finding a minimum of $f\left(r_{1}, \ldots, r_{N}\right)$ under the $M<N$ constraints $h\left(r_{1}, \ldots, r_{N}\right)=c$, or $h_{\alpha}\left(r_{1}, \ldots, r_{N}\right)=c_{\alpha}, \alpha=$ $1, \ldots, M$ is: find $\boldsymbol{r}_{*}$ such that:

$$
\begin{equation*}
\boldsymbol{h}\left(\boldsymbol{r}_{*}\right)=\boldsymbol{c} \tag{C.2.7}
\end{equation*}
$$

$$
\begin{equation*}
\forall \boldsymbol{r}: f\left(\boldsymbol{r}_{*}\right) \boldsymbol{h}\left(\boldsymbol{r}_{*}\right) \leq f(\boldsymbol{r}), \boldsymbol{h}(\boldsymbol{r})=\boldsymbol{c} \tag{C.2.8}
\end{equation*}
$$

To facilitate such a search, we formulate the Lagrangian function:
(C.2.9)

$$
L(\boldsymbol{r}, \boldsymbol{\lambda} ; \boldsymbol{c})=f(\boldsymbol{r})-\sum_{\alpha=1}^{M} \lambda_{\alpha}\left(h_{\alpha}(\boldsymbol{r})-c_{\alpha}\right) .
$$

Our plan is to find the position of $\boldsymbol{r}(\boldsymbol{\lambda})$ which minimizes $L$ for any choice of $\boldsymbol{\lambda}$ and then change $\boldsymbol{\lambda}$ until $\boldsymbol{h}(\boldsymbol{r}(\boldsymbol{\lambda}))=\boldsymbol{c}$. At $\lambda_{*}$ we have $L$ assuming a minimum at a point $\boldsymbol{r}_{*}=\boldsymbol{r}\left(\boldsymbol{\lambda}_{*}\right)$. The necessary conditions for the constrained minimum to be achieved at the point $\boldsymbol{r}_{*}$ and with the Lagrange multipliers $\lambda_{*}$ are the so-called Euler-Lagrange equations:
(C.2.10)

$$
\begin{aligned}
& \text { (C.2.10) } 0=\frac{\partial}{\partial r_{i}} L\left(\boldsymbol{r}_{*}, \boldsymbol{\lambda}_{*} ; \boldsymbol{c}\right)=\frac{\partial}{\partial r_{i}} f(\boldsymbol{r})-\sum_{\alpha=1}^{M} \lambda_{\alpha} \frac{\partial}{\partial r_{i}} h_{\alpha}(\boldsymbol{r}) \\
& \text { (C.2.11) } 0=\frac{\partial}{\partial \lambda_{\alpha}} L\left(\boldsymbol{r}_{*}, \boldsymbol{\lambda}_{*} ; \boldsymbol{c}\right)=h_{\alpha}\left(\boldsymbol{r}_{*}\right)-c_{\alpha}
\end{aligned}
$$

Note that $\boldsymbol{\lambda}_{*}$ is not a minimizer of $L$. In fact the opposite is true: $\lambda_{*}$ is its maximizer.

It is interesting now to ask how $f\left(r_{*}\right)$ changes if we change the value of the constraint $c_{\alpha}$. Indeed, when the constraints are changed, the optimized point and Lagrange multiplier can change, so the Lagrangian is changed:

$$
\begin{equation*}
\frac{\partial L}{\partial c_{\alpha}}=\frac{\partial r_{i}}{\partial c_{\alpha}} \frac{\partial L}{\partial r_{i}}+\frac{\partial \lambda_{\beta}}{\partial c_{\alpha}} \frac{\partial L}{\partial \lambda_{\beta}}+\frac{\partial L}{\partial c_{\alpha}} \tag{C.2.12}
\end{equation*}
$$

The first two terms on the right are zero by the Euler-Lagrange equations, Eq. (C.2.10) so

$$
\begin{equation*}
\frac{\partial L}{\partial c_{\alpha}}=\frac{\partial L}{\partial c_{\alpha}}=\lambda_{\alpha} \tag{C.2.13}
\end{equation*}
$$

This equation reveals the meaning of the Lagrange multipliers $\lambda_{\alpha *}$ at the optimal point: they are equal to the rate at which the optimal value of the minimized function $f$ changes when $c_{\alpha}$, the value of the $\alpha$ constraint, is changed. This is an important result which we use below whenever we want to give physical significance to Lagrange multipliers.

## C.3. Minimization of functionals

The same considerations for functions apply for functionals. Given a functional $I[f]$, a necessary condition for its minimum is:

$$
\begin{equation*}
\frac{\delta I}{\delta f(\boldsymbol{r})}=0 . \tag{C.3.1}
\end{equation*}
$$

For example, consider a 1D classical particle of mass $m$ in a potential well $v(x)$. The action $S$ is a functional of the trajectory $x(t)$ is the integral of the Lagrangian $L=T-V$ along a trajectory where $T=\frac{1}{2} m \dot{x}(t)^{2}$ is the kinetic energy and $V=v(x(t))$ is the potential energy along the trajectory:

$$
\begin{equation*}
S[x]=\int_{t_{0}}^{t_{f}}\left[\frac{1}{2} m \dot{x}\left(t^{\prime}\right)^{2}-v\left(x\left(t^{\prime}\right)\right)\right] d t^{\prime} \tag{C.3.2}
\end{equation*}
$$

For any trajectory $x(t)$ between times $t_{0}$ and $t_{f}, S[x]$ returns a number. Lagrange showed that finding the trajectory that makes $L[x]$ stationary (although, not necessarily minimal), under the condition that $x\left(t_{0}\right)=x_{0}$ and $x\left(t_{f}\right)=x_{f}$ are given and thus not varied, is equivalent to solving Newton's equations under these same constraints. The functional differentiation of the kinetic energy is performed with:

$$
\begin{align*}
\delta T[x] & =\frac{1}{2} m \int_{t_{0}}^{t_{f}}\left[\left(\dot{x}\left(t^{\prime}\right)+\delta \dot{x}\left(t^{\prime}\right)\right)^{2}-\left(\dot{x}\left(t^{\prime}\right)\right)^{2}\right] d t^{\prime}  \tag{C.3.3}\\
& =\frac{1}{2} m \int_{t_{0}}^{t_{f}} 2 \dot{x}\left(t^{\prime}\right) \delta \dot{x}\left(t^{\prime}\right) d t^{\prime} \\
& =-m \int_{t_{0}}^{t_{f}} \ddot{x}\left(t^{\prime}\right) \delta x\left(t^{\prime}\right) d t^{\prime} \tag{C.3.5}
\end{align*}
$$

and:
(C.3.6) $-\delta V[x]=-\int_{t_{0}}^{t_{f}}\left[v\left(x\left(t^{\prime}\right)+\delta x\left(t^{\prime}\right)\right)-v\left(x\left(t^{\prime}\right)\right)\right] d t^{\prime}$
(C.3.7)

$$
=-\int_{t_{0}}^{t_{f}} v^{\prime}\left(x\left(t^{\prime}\right)\right) \delta x\left(t^{\prime}\right) d t^{\prime}
$$

Demanding $\delta S=0$ we find:
(C.3.8)

$$
-m \ddot{x}(t)-v(x(t))=0
$$

which is the Newton equation of motion.
Exercise 62. Use functional analysis and the Euler-Lagrange equation to find the shape in a plane of a closed contour encircling maximal area under the constraint of a given circumference $l_{0}$.

Solution. Assume the origin is within the shape and the perimeter is $\boldsymbol{r}(\theta)=r(\theta)(\cos \theta \sin \theta)$. Then we want to find $r(\theta)$. (C.3.9)

$$
\boldsymbol{r}^{\prime}(\theta)=r^{\prime}(\theta)\left(\begin{array}{ll}
\cos \theta & \sin \theta
\end{array}\right)+r(\theta)(-\sin \theta \quad \cos \theta)
$$

The area of a small arc between $\theta$ and $\theta+\delta \theta$ is $\delta S=\frac{1}{2} r(\theta)^{2} d \theta$ hence the area is

$$
\begin{equation*}
S[r]=\int_{0}^{2 \pi} \frac{1}{2} r(\theta)^{2} d \theta \tag{C.3.10}
\end{equation*}
$$

The circumference is $\delta L=\left|\boldsymbol{r}^{\prime}(\theta)\right| d \theta$ so

$$
\begin{equation*}
\delta L=\sqrt{r(\theta)^{2}+r^{\prime}(\theta)^{2}} d \theta \tag{C.3.11}
\end{equation*}
$$

hence

$$
\begin{equation*}
L[r]=\int_{0}^{2 \pi} \sqrt{r(\theta)^{2}+r^{\prime}(\theta)^{2}} d \theta \tag{С.3.12}
\end{equation*}
$$

Now minimize $S$ under the constraint that $L=l_{0}$.

## APPENDIX D

## Spin states

## D.1. Spin: The Stern-Gerlach experiment

A Stern-Gerlach device is a magnet with two non-equivalent poles, $N$ and $S$ in the $z$ direction through which a beam of atoms of valence 1 shoots in the $x$ direction. As a result of the inequivalent poles the magnetic field is non-homogeneous and the magnetization of each electron is either attracted or repelled from the high magnetic field near the $N$ pole. Stern and Gerlach found that the beam splits into two well-defined


Figure D.0.1. A schematic depiction of a Stern-Gerlach experiment. An atomic beam of Ag atoms moves in the x direction between a magnet with non-homogeneous magnetic field. The atoms with spin up electrons are drawn towards the high field and those with spin down repelled by the strong field. Hence the beam splits into two.
beams. At first they rejoiced that they "proved" Bohr's theory of the H atom expecting that the angular momentum of the electron in the atom is quantized. However, it was soon realized that the $l=1$ angular momentum predicts the beam to split into three sub beams: an "up" magnetization beam, a "down" magnetization beam and a unperturbed beam having zero magnetization. Later it was understood that the situation was even more baffling as the valence electron of Ag is in the $s(l=0)$ state, and no splitting should have been seen at all! The correct interpretation of the Stern-Gerlach experiment was given several years later by Uhlenbeck and Goudsmit, who posited the existence of the "electron spin", where the motion has only two states of magnetization.

## D.2. Spin operators

We assume that the two states of the electron are a spinor of 2 components. For the SG device which splits the beam in the z direction we designate the spin up and down orthonormal states as $\alpha_{z}=\binom{1}{0}$ and $\beta_{z}=\binom{0}{1}$. The operators on such spin states must be $2 \times 2$ matrices. For example, the operator corresponding to the spin component of spin in direction $z$ is: $\hat{S}_{z}=\frac{1}{2}\left(\begin{array}{cc}1 & 0 \\ 0 & -1\end{array}\right)$ and it is easy to verify that $\alpha_{z}$ and $\beta_{z}$ are the eigenstates with eigenvalues of $M=\frac{1}{2}$ and $M=-\frac{1}{2}$ respectively. This is analogous to the angular momentum, $z$ component $L_{z}=i \hbar \frac{\partial}{\partial \phi}$ which has as eigenvalues $m \hbar$ where $m$ can assume $2 l+1$ integers $m=-l, \ldots,+l$ and $l$ is the total angular momentum quantum number which is a non-negative integer and $\hbar^{2} l(l+1)$ is the eigenvalue of the total angular momentum operator $\hat{L}^{2}$. Since in spin $M$ has only 2 possibilities then the total spin angular momentum $S$ must be equal to $\frac{1}{2}$ (since $2 \times \frac{1}{2}+1=2$ ).

Just like there are angular momentum components in the $y$ and $z$ directions, so does spin have such components. A device which splits the incoming atomic beam into two subbeams moving in the $x$ direction is the SGx apparatus and experimental observation tells us that when a beam of "up" electrons in the state $\alpha_{z}$ is passed through such $S G_{x}$ device in the the beam splits into exactly two beams of equal intensity. Hence the state $\alpha_{z}$ must be a linear combination $A \alpha_{x}+B \beta_{x}$ with $|A|^{2}=|B|^{2}=1$ of the states representing the spin-up and spin-down of the SGX apparatus, The same goes for $\beta_{z}$. Hence we write $\alpha_{z}=\frac{1}{\sqrt{2}}\left(\alpha_{x}+\beta_{x}\right)$ and $\beta_{z}=\frac{1}{\sqrt{2}}\left(\alpha_{x}-\beta_{x}\right)$. From this we find $\alpha_{x}=\frac{1}{\sqrt{2}}\binom{1}{1}$ and $\beta_{x}=\frac{1}{\sqrt{2}}\binom{1}{-1}$ with $\hat{S}_{x}=\frac{1}{2}\left(\begin{array}{ll}0 & 1 \\ 1 & 0\end{array}\right)$. Finally, we can define spin-up and -down for the $S G_{y}$ apparatus and this time the linear combination coefficients must involve imaginary numbers: $\alpha_{y}=\frac{1}{\sqrt{2}}\binom{1}{i}$ and $\beta_{y}=\frac{1}{\sqrt{2}}\binom{i}{1}$ with $\hat{S}_{y}=\frac{1}{2}\left(\begin{array}{cc}0 & -i \\ i & 0\end{array}\right)$.

It is quite straightforward to see that the spin operators obey the angular momentum commutation relations: $\left[\hat{S}_{a}, \hat{S}_{b}\right]=$ $i \hat{S}_{c}$ where $a b c$ are a cyclic permutation of xyz. For angular momentum operators $L^{2}$ has the eigenvalues $l(l+1)$ where $l$ can be any non-negative integer. Here, one has: $\hat{S}^{2}=\hat{S}_{x}^{2}+\hat{S}_{y}^{2}+$ $\hat{S}_{z}^{2}=\frac{3}{4}\left(\begin{array}{ll}1 & 0 \\ 0 & 1\end{array}\right)$ and since $\frac{3}{4}=\frac{1}{2}\left(\frac{1}{2}+1\right)$ we see once again that the total spin quantum number is $s=\frac{1}{2}$. showing that the total spin obeys Furthermore, one sees that $\hat{S}_{a} \hat{S}_{b}+\hat{S}_{b} S_{a}=0$ (when $a \neq b$ ). In orbital angular momentum it is useful to define the raising and lowering ioperators $L_{ \pm}==L_{x x} \pm i L_{y}$. The same can be done for spin, with $\hat{S}_{+}=\hat{S}_{x}+i \hat{S}_{y}=\left(\begin{array}{ll}0 & 1 \\ 0 & 0\end{array}\right)$
and $\hat{S}_{-}=\hat{S}_{x}-i \hat{S}_{y}=\left(\begin{array}{ll}0 & 0 \\ 1 & 0\end{array}\right)$. Notice that $\hat{S}_{+} \beta_{z}=\alpha_{z}$ so indeed this operator raises the spin by 1. Furthermore, $\hat{S}_{+} \alpha_{z}=0$ because one cannot raise the spin beyond $\frac{1}{2}$. Notice that $\left[\hat{S}^{2}, S_{ \pm}\right]=0$, so $S_{ \pm}$cannot change the total spin state of the system, only the spin component. It can be checked that: $\hat{S}^{2}=\hat{S}_{+} \hat{S}_{-}+\hat{S}_{z}^{2}-\hat{S}_{z}$ and $\hat{S}^{2}=\hat{S}_{-} \hat{S}_{+}+\hat{S}_{z}^{2}+\hat{S}_{z}$.

## D.3. Many electron spin states

The usefulness of these formal relations is evident when we have more than one particle. For example, suppose we have 2 particles. The total spin $\hat{S}^{2}=(\overrightarrow{\hat{S}}(1)+\overrightarrow{\hat{S}}(2))^{2}$ of the system can be written as

$$
\begin{align*}
\hat{S}^{2} & =\hat{S}^{2}(1)+\hat{S}^{2}(2) \\
& +\hat{S}_{+}(1) \hat{S}_{-}(2)+\hat{S}_{-}(1) \hat{S}_{+}(2)  \tag{D.3.1}\\
& +2 \hat{S}_{z}(1) \hat{S}_{z}(2)
\end{align*}
$$

A simple example of 2 electrons is $X(2,1,1 ; 1)=\alpha_{z}(1) \alpha_{z}(2) \equiv$ $\alpha_{z} \alpha_{z}$. It is immediate to see this state has $N=2$ electrons and $M=1$. To show it has $S=1$ We use Eq. D.3.1: $\hat{S}^{2} X(2,1,1 ; 1)=$ $\frac{3}{2} \alpha_{z} \alpha_{z}+2 \frac{1}{4} \alpha_{z} \alpha_{z}=2 \alpha_{z} \alpha_{z}$ i.e. $S(S+1)=2$ and so $S=1$. We now use the $\hat{S}_{-}$to lower the $z$ component and obtain $X(2,1,0 ; 1)=\hat{S}_{-} X(2,1,1 ; 1)=\left(\hat{S}_{-}(1)+\hat{S}_{-}(2)\right) \alpha_{z} \alpha_{z}=$ $\beta_{z} \alpha_{z}+\alpha_{z} \beta_{z}$. What about the $S=0$ state of 2 electrons? Here we show that $X(2,0,0 ; 1)=\alpha_{z} \beta_{z}-\beta_{z} \alpha_{z}$. Using Eq. D.3.1 we prove that $S=0$ this state: $\hat{S}^{2} X(2,0,0 ; 1)=\frac{3}{2} X(2,0,0 ; 1)-$ $X(2,0,0 ; 1)-\frac{1}{2} X(2,0,0 ; 1)=0$ hence $S=0$.

When we take two electrons their spins add up. The total state has spin quantum number $S$ which can be the sum of the spin of each electron, i.e $\frac{1}{2}+\frac{1}{2}=1$ or the difference $\frac{1}{2}-$ $\frac{1}{2}=0$. In general, when we add electrons can be in 4 states $\left|N_{1} S_{1} M_{1}\right\rangle\left|N_{2} S_{2} M_{2}\right\rangle \rightarrow|N, S M\rangle$ where $N=N_{1}+N_{2}, M=$
$M_{1}+M_{2}$ and $\left|S_{2}-S_{1}\right| \leq S \leq S_{1}+S_{2}$. We can build the $N e l e c t r o n s$ spin states incrementally, building from $N=1$ and adding in each stage an electron. To get a $N$ electron state of spin $S$ and component $M$ we either start from a $N-1$ electron state of spin $S+\frac{1}{2}$ and then reduce the spin by $\frac{1}{2}$ :
(D.3.2)

$$
\begin{align*}
& X(N, S, M, k)=\sqrt{\frac{S+M+1}{2(S+1)}} X\left(N-1, S+\frac{1}{2}, M+\frac{1}{2}, k^{\prime}\right) \beta(N) \\
& \text { (D.3.3) } \quad-\sqrt{\frac{S-M+1}{2(S+1)}} X\left(N-1, S+\frac{1}{2}, M-\frac{1}{2}, k^{\prime}\right) \alpha(N) \tag{D.3.3}
\end{align*}
$$

or start from a $N-1$ electron state of spin $S-\frac{1}{2}$ and then increase the spin by $\frac{1}{2}$ :
(D.3.4)

$$
\begin{align*}
& X(N, S, M, k)=\sqrt{\frac{S+M}{2 S}} X\left(N-1, S-\frac{1}{2}, M-\frac{1}{2}, k^{\prime}\right) \alpha(N) \\
& \text { (D.3.5) } \tag{D.3.5}
\end{align*}+\sqrt{\frac{S-M}{2 S}} X\left(N-1, S-\frac{1}{2}, M+\frac{1}{2}, k^{\prime}\right) \beta(N) \text { }
$$

We can then count the number of states using the branching diagram:

For example, let us construct the spin functions of $N=3$ electrons and $S=\frac{1}{2}$. We start from the 42 -electron states: $X(2,1,-1)=\beta \beta, X(2,1,0)=\frac{1}{\sqrt{2}}(\alpha \beta+\beta \alpha), X(2,1,1)=$ $\alpha \alpha$ and $X(2,0,0)=\frac{1}{\sqrt{2}}(\alpha \beta-\beta \alpha)$. When a third electron is added the states we have 4 states:

$$
\begin{equation*}
X\left(3, \frac{1}{2}, \frac{1}{2}, 1\right)=\sqrt{\frac{1}{6}}(2 \alpha \alpha \beta-\alpha \beta \alpha-\beta \alpha \alpha) \tag{D.3.6}
\end{equation*}
$$

$$
\begin{equation*}
X\left(3, \frac{1}{2},-\frac{1}{2}, 1\right)=\sqrt{\frac{1}{6}}(-2 \beta \beta \alpha+\beta \alpha \beta+\alpha \beta \beta) \tag{D.3.7}
\end{equation*}
$$

and

$$
\begin{align*}
X\left(3, \frac{1}{2}, \frac{1}{2}, 2\right) & =\sqrt{\frac{1}{2}}(\alpha \beta \alpha-\beta \alpha \alpha)  \tag{D.3.8}\\
X\left(3, \frac{1}{2},-\frac{1}{2}, 2\right) & =\sqrt{\frac{1}{2}}(\alpha \beta \beta-\beta \alpha \beta)
\end{align*}
$$

2-electron state was $X(2,0,0)$, the 3 electron state must have $S=\frac{1}{2}$, i.e. it is $X\left(3, \frac{1}{2}, \frac{1}{2}\right), X\left(3, \frac{1}{2},-\frac{1}{2}\right)$ while if the 2 electron state was $X(2,1,0)$, then the
resulting state can be: $X\left(3, \frac{1}{2}, \frac{1}{2}\right), X\left(3, \frac{1}{2},-\frac{1}{2}\right)$ while if the 2 electron state was $X(2,1,1)$ the new state can be $X\left(3, \frac{3}{2}, \frac{3}{2}\right)$ or $X\left(3, \frac{3}{2}, \frac{1}{2}\right)$ and three electron state is either $1-\frac{1}{2}=\frac{1}{2}$ or $1+\frac{1}{2}=\frac{3}{2}$. Thus there are $2 S=\frac{1}{2}$ states and one $S=\frac{3}{2}$ states. The $S=\frac{3}{2}$ has 4 states, $M=-\frac{3}{2},-\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$. The two $S=\frac{1}{2}$ states has and the angular momentum, component $M$.
t are the possible spin states? For $N$ electrons they are written as $|N, S, M k\rangle$ where $S$ is the spin quantum number $M=-S,-S+1, \ldots, S-1, S$ has $2 S+1$ values and $k$ an additional index. In principle, for a given $M$ there are $\binom{N}{M}=$ $\frac{N!}{M!(N-M)!}$ different spin states (so $\sum_{M=0}^{N}\binom{N}{M}=2^{N}$ being the total number of spin states for $N$ electrons).

Let us take $N=5$ electrons, for example, and $M=2$. There are 10 states. $S=2,3,4,5$ gives just 5 states $X(5, S, 2)$ so additional functions are needed with an additional index $(k)$ to distinguish between them.

The movement from the $|N, S, M, k\rangle$ notation to the $\alpha \alpha \beta \cdots \alpha$ notation is important for applying the Pauli principle. Given a spin state $X(N, S, M, k)$ and a spatial state $\psi(1, \ldots, N)$ we can form a function obeying the Pauli by antisymmetrization:
(D.3.10)

$$
\Psi_{N S M k}(1, \ldots, N)=\hat{A} \psi\left(\boldsymbol{r}_{1}, \ldots, \boldsymbol{r}_{N}\right) X(N, S, M, k)
$$

where $\hat{A}=\frac{1}{\sqrt{N!}} \sum_{P}(-)^{P} P$ is the Hermitean antisymmetrizing projection operator and the sum is over all permutations of the numbers $1, \ldots, N$, each permutation is symmetric (comes with sign + or antisymmetric, comes with minus sign). As an example, we take 2 electrons in the spin state $X(2,0,0,0)=$ $\alpha \beta-\beta \alpha$. In this case the antisymmetrizing operator is $A=$ $\frac{1}{\sqrt{2}}\left(1-P_{12}\right)$ where $P_{12}$ is the operator exchanging 1 and 2 : $P_{12} \psi\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right)(\alpha \beta-\beta \alpha)=\psi\left(\boldsymbol{r}_{2}, \boldsymbol{r}_{1}\right)(\beta \alpha-\alpha \beta)$, leading to

$$
\begin{aligned}
\Psi_{2000}(1,2) & =\hat{A} \psi\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right) X(2,0,0,0) \\
& =\frac{1}{\sqrt{2}}\left(\psi\left(\boldsymbol{r}_{1}, \boldsymbol{r}_{2}\right)+\psi\left(\boldsymbol{r}_{2}, \boldsymbol{r}_{1}\right)\right)(\alpha \beta-\beta \alpha)
\end{aligned}
$$

we see that the symmetric part of $\psi$ gets multiplied by the antisymmetric spin function.

## Bibliography

[1] E. Fermi. Un metodo statistico per la determinazione di alcune priorieta dell'atome. Rend. Accad. Naz., 6:602, 1927.
[2] Svetlana Kotochigova, Zachary H Levine, Eric L Shirley, MD Stiles, and Charles W Clark. Local-density-functional calculations of the energy of atoms. Phys. Rev. A, 55(1):191, 1997.
[3] M Levy and J. P. Perdew. The constrained search formulation of density functional theory. In RM Dreizler and J. da Providencia, editors, Density Functional Methods in Physics, NATO Advanced Science Institutes Series B 123, chapter 1, page 11. Plenum Press, New York, 1985.
[4] E. H. Lieb. Thomas-fermi and related theories of atoms and molecules. Rev. Mod. Phys., 53(4):603, 1981.
[5] Elliott H Lieb. The stability of matter: From atoms to stars. Springer, 2001.
[6] Edward Teller. On the stability of molecules in the thomas-fermi theory. Rev. Mod. Phys., 34(4):627, 1962.
[7] L H Thomas. The calculation of atomic fields. Proceedings of the Cambridge Philoshophical Society, 23:542, 1927.


[^0]:    ${ }^{1}$ Choosing the node in the origin simplifies the argument but is not essential.

[^1]:    ${ }^{1}$ This is easily proved: $\tilde{f}(\boldsymbol{k})=\int h\left(\boldsymbol{r}^{\prime}\right)\left[\int d^{3} r e^{i \boldsymbol{k} \cdot \boldsymbol{r}} u\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) d^{3} r\right] d^{3} r^{\prime}$ $=\int e^{i \boldsymbol{k} \cdot \boldsymbol{r}^{\prime}} h\left(\boldsymbol{r}^{\prime}\right)\left[\int d^{3} r e^{i \boldsymbol{k} \cdot\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right)} u\left(\boldsymbol{r}-\boldsymbol{r}^{\prime}\right) d^{3} r\right] d^{3} r^{\prime}=\tilde{u}(\boldsymbol{k}) \tilde{h}(\boldsymbol{k})$

[^2]:    ${ }^{1}$ In a simply connected domain $D$ if $x_{1}, x_{2} \in D$ then the straight line segment connecting the two points is also within $D$. In other words $\forall \lambda \in$ $[0,1]: \lambda x_{1}+(1-\lambda) x_{2} \in D$.

