

# Electron Density Functional Theory

Lecture notes (rough draft)

October 2009

**Roi Baer**

Institute of Chemistry,  
The Fritz Haber Center for Molecular Dynamics  
The Hebrew University of Jerusalem,  
Jerusalem, 91904 Israel

© Roi Baer

## Contents

<b>I. INTRODUCTORY BACKGROUND TOPICS .....</b>	<b>7</b>
<b>A. THE SCHRÖDINGER EQUATION .....</b>	<b>7</b>
<b>B. BOSONS AND FERMIONS .....</b>	<b>7</b>
I. GROUND-STATE WAVE FUNCTION OF BOSONS .....	7
II. GROUND-STATE WAVE FUNCTION OF FERMIONS .....	10
<b>C. WHY ELECTRONIC STRUCTURE IS AN IMPORTANT DIFFICULT PROBLEM .....</b>	<b>10</b>
<b>D. THE BORN-OPPENHEIMER THEORY .....</b>	<b>12</b>
I. THE ADIABATIC THEOREM.....	12
II. MOTIVATION FOR THE BORN-OPPENHEIMER APPROXIMATION: CLASSICAL NUCLEI.....	16
III. THE BORN-OPPENHEIMER APPROXIMATION IN QUANTUM NUCLEAR CASE .....	18
<b>E. ELECTRON CORRELATION .....</b>	<b>21</b>
I. THE ELECTRONIC WAVE FUNCTION OF TWO NON-INTERACTING ELECTRONS .....	21
II. CORRELATION IN ACTION: A WAVE FUNCTION OF 2 <i>INTERACTING</i> PARTICLES IN AN HARMONIC TRAP .....	24
<b>F. THE ELECTRON DENSITY IS A MUCH SIMPLER OBJECT THAN THE WAVE FUNCTION.</b>	<b>31</b>
<b>G. THE VARIATIONAL PRINCIPLE .....</b>	<b>34</b>
<b>H. DILATION RELATIONS.....</b>	<b>36</b>
I. THE CONCEPT OF THE VIRIAL IN CLASSICAL MECHANICS.....	37
II. THE VIRIAL THEOREM IN QUANTUM MECHANICS.....	39
III. SLATER'S THEORY OF THE CHEMICAL BOND USING THE VIRAL THEOREM .....	40
IV. OTHER DILATION FACTS.....	47
<b>I. INTRODUCTION TO FUNCTIONALS .....</b>	<b>48</b>
I. DEFINITION OF A FUNCTIONAL AND SOME EXAMPLES .....	48
II. LINEAR FUNCTIONALS .....	49
III. FUNCTIONAL DERIVATIVES.....	50
IV. INVERTIBLE FUNCTION FUNCTIONALS AND THEIR FUNCTIONAL DERIVATIVES.....	53
V. CONVEX FUNCTIONALS .....	54
<b>J. MINIMIZING FUNCTIONALS IS AN IMPORTANT PART OF DEVELOPING A DFT .....</b>	<b>58</b>
I. MINIMIZATION OF FUNCTIONS.....	58
II. CONSTRAINED MINIMIZATION: LAGRANGE MULTIPLIERS.....	60
III. MINIMIZATION OF FUNCTIONALS .....	64

<b>II. <u>MY FIRST DENSITY FUNCTIONAL: THOMAS-FERMI THEORY</u></b> .....	<b>68</b>
A. BASIC CONCEPTS IN THE ELECTRON GAS AND THE THOMAS-FERMI THEORY .....	68
B. MINIMIZATION OF THE THOMAS-FERMI ENERGY .....	73
C. THOMAS-FERMI DOES NOT ACCOUNT FOR MOLECULES .....	75
D. THOMAS-FERMI SCREENING .....	76
E. VON WEIZSÄCKER KINETIC ENERGY .....	80
<b>III. <u>MANY-ELECTRON WAVE FUNCTIONS</u></b> .....	<b>81</b>
A. THE ELECTRON SPIN .....	81
B. THE PAULI PRINCIPLE.....	82
C. THE EXCITED STATES OF THE HELIUM ATOM .....	83
D. THE SLATER WAVE FUNCTION IS THE BASIC ANTI-SYMMETRIC FUNCTION DESCRIBING N ELECTRONS IN N ORBITALS.....	83
E. WITHOUT LOSS OF GENERALITY, WE MAY ASSUME THE ORBITALS OF A SLATER WAVE FUNCTION ARE ORTHOGONAL .....	84
F. ANY ANTISYMMETRIC FUNCTION CAN BE EXPANDED AS A SUM OF BASIC SLATER (DETERMINANTAL) FUNCTIONS.....	85
G. DETERMINANT EXPECTATION VALUES.....	87
I. ONE-BODY OPERATORS .....	87
II. TWO-BODY OPERATORS .....	88
<b>IV. <u>THE HARTREE-FOCK THEORY</u></b> .....	<b>89</b>
A. THE HARTREE-FOCK ENERGY AND EQUATIONS.....	89
B. RESTRICTED CLOSED-SHELL HARTREE-FOCK .....	95
C. ATOMIC ORBITALS AND GAUSSIAN BASIS SETS.....	98
D. VARIATIONAL-ALGEBRAIC APPROACH HARTREE-FOCK .....	99
E. THE ALGEBRAIC DENSITY MATRIX AND CHARGE ANALYSIS.....	105
F. SOLVING THE HARTREE-FOCK EQUATIONS.....	107
I. DIRECT INVERSION IN ITERATIVE SPACE (DIIS).....	107
II. DIRECT MINIMIZATION .....	108
G. PERFORMANCE OF THE HARTREE-FOCK APPROXIMATION .....	109

<b>H. BEYOND HARTREE-FOCK.....</b>	<b>110</b>
<b><u>V. ADVANCED TOPICS IN HARTREE-FOCK THEORY .....</u></b>	<b><u>112</u></b>
<b>A. LOW-LYING EXCITATIONS AND THE STABILITY OF THE HARTREE-FOCK GROUND STATE .....</b>	<b>112</b>
<b>I. CI-SINGLES AND BRILLOUIN'S THEOREM .....</b>	<b>112</b>
<b>II. HARTREE-FOCK STABILITY.....</b>	<b>114</b>
<b>B. KOOPMANS' THEOREM.....</b>	<b>116</b>
<b>C. FRACTIONAL OCCUPATION NUMBERS, THE HF ORBITAL FUNCTIONAL AND THE GENERALIZED KOOPMANS' THEOREM.....</b>	<b>119</b>
<b>D. HARTREE-FOCK FOR THE HOMOGENEOUS ELECTRON GAS.....</b>	<b>121</b>
<b>I. HARTREE-FOCK ORBITALS AND ORBITAL ENERGIES OF HEG.....</b>	<b>121</b>
<b>II. THE DENSITY OF STATES OF THE HEG .....</b>	<b>127</b>
<b>III. STABILITY OF THE HEG IN HARTREE-FOCK THEORY .....</b>	<b>129</b>
<b><u>VI. THE HOHENBERG-KOHN DENSITY THEORY .....</u></b>	<b><u>131</u></b>
<b>A. THE FIRST HK THEOREM .....</b>	<b>131</b>
<b>B. THE HK FUNCTIONAL .....</b>	<b>133</b>
<b>C. MINIMUM PRINCIPLE FOR DENSITY FUNCTIONAL THEORY .....</b>	<b>135</b>
<b>D. AN INTERESTING OBSERVATION ON THE VARIATIONAL PRINCIPLE OF NON-INTERACTING ELECTRONS.....</b>	<b>137</b>
<b>E. THE SET OF V-REPRESENTABLE DENSITIES.....</b>	<b>138</b>
<b>I. V-REP DENSITIES CORRESPOND TO GROUND STATES WAVE FUNCTION OF SOME POTENTIAL WELL.....</b>	<b>138</b>
<b>II. SOME NON-V-REPRESENTABILIT ISSUES.....</b>	<b>139</b>
<b>III. THE SET OF V-REP DENSITIES OF A GIVEN ELECTRON NUMBER IS NOT CONVEX .....</b>	<b>140</b>
<b>F. LEVY-LIEB GENERALIZATION OF THE HK FUNCTIONAL .....</b>	<b>143</b>
<b>G. THE DILATION INEQUALITY FOR THE HK FUNCTIONAL .....</b>	<b>148</b>
<b><u>VII. THE KOHN-SHAM METHOD .....</u></b>	<b><u>149</u></b>
<b>A. NON-INTERACTING ELECTRONS.....</b>	<b>149</b>

<b>B. ORBITALS FOR THE NON-INTERACTING ELECTRONS.....</b>	<b>153</b>
<b>C. THE CORRELATION ENERGY FUNCTIONAL: DEFINITION AND SOME FORMAL PROPERTIES.....</b>	<b>154</b>
<b>D. THE KOHN SHAM EQUATIONS.....</b>	<b>157</b>
I. THE KOHN-SHAM EQUATION FROM A SYSTEM OF NON-INTERACTING PROBLEM .....	157
II. SYSTEMS WITH PARTIALLY OCCUPIED ORBITALS.....	159
III. IS THE GROUND STATE WAVE FUNCTION OF NON-INTERACTING PARTICLES ALWAYS A SLATER WAVE FUNCTION .....	161
IV. JANAK'S THEOREM .....	163
<b>E. "VIRIAL THEOREM" RELATED IDENTITIES IN DFT.....</b>	<b>164</b>
<b>F. GALILEAN INVARIANCE.....</b>	<b>168</b>
<b>G. HOLES AND THE ADIABATIC CONNECTION.....</b>	<b>170</b>
I. THE EXCHANGE AND CORRELATION HOLES .....	170
II. THE FERMI-COULOMB HOLE FOR HARMONIC ELECTRONS.....	175
III. THE FERMI HOLE IN THE NON-INTERACTING SYSTEM.....	178
IV. THE ADIABATIC CONNECTION .....	181
<b>H. DERIVATIVE DISCONTINUITY IN THE EXCHANGE CORRELATION POTENTIAL FUNCTIONAL.....</b>	<b>184</b>
<b><u>VIII. APPROXIMATE CORRELATION ENERGY FUNCTIONALS.....</u></b>	<b><u>185</u></b>
<b>A. THE LOCAL DENSITY APPROXIMATION (LDA).....</b>	<b>185</b>
I. THE EXCHANGE ENERGY PER ELECTRON IN THE HEG .....	186
II. CORRELATION ENERGY OF THE HEG: THE HIGH DENSITY LIMIT .....	187
III. CORRELATION ENERGY OF THE HEG: THE LOW DENSITY LIMIT AND THE WIGNER CRYSTAL .....	189
IV. MONTE-CARLO DETERMINATION OF THE CORRELATION ENERGY FOR THE HEG .....	192
V. THE POLARIZED HEG; LOCAL SPIN-DENSITY APPROXIMATION (LSDA) .....	192
VI. SUCCESSES AND FAILURES OF LSDA.....	193
VII. PLAUSIBLE REASONS FOR THE SUCCESS OF LSDA .....	193
<b>B. SEMILOCAL FUNCTIONALS AND THE GENERALIZED GRADIENT APPROXIMATION ..</b>	<b>196</b>
<b><u>IX. GENERALIZED KOHN-SHAM APPROACHES.....</u></b>	<b><u>196</u></b>

A.	THE GENERALIZED KOHN-SHAM FRAMEWORK .....	196
B.	KOHN-SHAM FROM GENERALIZED KS .....	199
C.	THE HYBRID FUNCTIONAL OF BECKE .....	199
D.	.....	200
E.	LONG-RANGE SELF-REPULSION AND LACK OF DERIVATIVE DISCONTINUITY.....	200
F.	RANGE SEPARATED HYBRIDS .....	200
G.	ORBITAL FUNCTIONALS AND OPTIMIZED EFFECTIVE POTENTIALS.....	200
H.	APPROXIMATE CORRELATION FUNCTIONALS AND THE BORN-OPPENHEIMER FORCE ON NUCLEI.....	200
<b>X.</b>	<b><u>MORE ON THE DFT CORRELATION ENERGY .....</u></b>	<b>203</b>
A.	APPROXIMATIONS TO $E_x$ ARE APPROXIMATION TO $E_C$ .....	203
B.	THE NECESSITY OF LONG-RANGE PARAMETER TUNING .....	209
I.	THE SYMMETRIC RADICAL CATION .....	209
II.	AROMATIC DONOR -TCNE ACCEPTOR CHARGE TRANSFER EXCITATION.....	211
<b>XI.</b>	<b><u>TDDFT .....</u></b>	<b>214</b>
A.	TIME-DEPENDENT LINEAR RESPONSE THEORY .....	214
B.	ALGEBRAIC APPROACH TO LINEAR RESPONSE.....	216
C.	FREQUENCY-DOMAIN RESPONSE.....	219
D.	EXCITATION ENERGIES FROM THE ALGEBRAIC TREATMENT .....	220

# I. Introductory background topics

## A. The Schrödinger equation

XXX

## B. Bosons and Fermions

### i. Ground-state wave function of Bosons

Consider a one dimensional Schrödinger equation for the ground state, which may be taken real. We will ask: is this wave function of the same sign everywhere? What we mean is: is it all non-negative or all non positive? Of course, if it is all non-positive then we can multiply it by -1 and ask again: is it all non-negative? Saying it is of a constant sign is the same as saying it has no nodes, i.e. it does not go through the x- axis anywhere.

**Exercise:** Prove the following theorem:

**Theorem:** The ground-state wave function of a 1D wave non-degenerate Schrödinger equation has no nodes.

Proof. Let us assume  $\psi(x)$  is the (real) normalized ground state. Then it is the wavefunction minimizes the functional:

$$E[\psi] = T[\psi] + V[\psi] = \frac{\hbar^2 \int_{-\infty}^{\infty} \psi'(x)^2 dx + \int_{-\infty}^{\infty} \psi(x)^2 V(x) dx}{\int_{-\infty}^{\infty} \psi(x)^2 dx} \quad (1.1.1)$$

### A robust expression for the kinetic energy expectation value

The kinetic energy is usually defined as  $T[\psi] = -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} \psi(x)^* \psi''(x) dx$ . Let us call this  $T_1$ . For continuous and twice differentiable functions that go to zero at  $\pm\infty$  an equivalent definition, is  $T[\psi] = \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} |\psi'(x)|^2 dx$ . Let us call this

definition  $T_2$ . We now show that  $T_2$  is more robust as it holds for wave functions that are discontinuous. To show that  $T_1$  fails, consider the wave function  $\psi_0(x) = e^{-|x|}$ . Notice, it is normalized  $\langle \psi_0 | \psi_0 \rangle = 1$ . Let us also assume for this example that  $\frac{\hbar^2}{2m} = 1$ . Now, for  $x \neq 0$   $\psi'_0(x) = -e^{-|x|} \frac{|x|}{x}$  (there is a discontinuity of  $\psi'_0(x)$  at  $x = 0$ ) and  $\psi''_0(x) = \psi_0(x)$ . Now  $T_1 = -\int_{-\infty}^{\infty} \psi_0(x)^* \psi''_0(x) dx = -1$  which is physically absurd! Kinetic energy must be positive! On the other hand,  $T_2 = \int_{-\infty}^{\infty} |\psi'_0(x)|^2 dx = 1$ . To see that the second result is proper, let us smooth the cusp by defining  $\psi_\alpha(x) = e^{-\left(\frac{x^2}{|x|+\alpha}\right)}$ . Clearly, when  $\alpha = 0$  we have are back at the original function  $\psi_0(x)$ , but for positive  $\alpha$ , no matter how small, the first derivative  $\psi'_\alpha(x) = -\frac{|x|(2\alpha+|x|)}{(|x|+\alpha)^2} e^{-\left(\frac{x^2}{|x|+\alpha}\right)}$  is everywhere continuous while  $\psi'_0(x)$  has a discontinuity at  $x = 0$ . By using nonzero  $\alpha$  we regularize the discontinuity. Now, we can calculate the integrals numerically and we:

$\alpha = 0.1$	$\langle \psi_\alpha   \psi_\alpha \rangle = 1.152$	$T_1[\psi_\alpha] = 0.912$	$T_2[\psi_\alpha] = 0.912$
$\alpha = 0.01$	$\langle \psi_\alpha   \psi_\alpha \rangle = 1.019$	$T_1[\psi_\alpha] = 0.988$	$T_2[\psi_\alpha] = 0.988$
$\alpha = 0.001$	$\langle \psi_\alpha   \psi_\alpha \rangle = 1.002$	$T_1[\psi_\alpha] = 0.999$	$T_2[\psi_\alpha] = 0.999$

Both  $T_1$  and  $T_2$  give here the same result now since  $\psi'_\alpha$  is continuous.

From the table it is clear that  $\lim_{\alpha \rightarrow 0} T_1[\psi_\alpha] = \lim_{\alpha \rightarrow 0} T_2[\psi_\alpha] = T_2[\psi_0] \neq T_1[\psi_0]$ . Clearly, the expression  $T_1$  is not suited for discontinuities in  $\psi'$ .

Using the  $T_2$  expression, we infer that  $T[\psi] = T[|\psi|]$ . Explain why. Is this true if  $\psi(x)$  has a node at  $x_0$ ? Now explain why  $V[\psi] = V[|\psi|]$ .

Now there are two options. Either  $|\psi| = \pm\psi$ . In this case there is no sign change and there is no node. Otherwise there is a node. Assume it is in  $x_0$ .



Let us compute the energy change when one makes a small perturbation to  $|\psi|$ :

$$\delta E[|\psi|] = \delta T[|\psi|] + 2 \int_{-\infty}^{\infty} |\psi(x)|(V(x) - E[|\psi|])\delta|\psi(x)|dx \quad (1.1.2)$$

Verify this and show we explicitly assumed that the norm of  $\psi$  is 1.

Assume the node is at  $x_0 = 0$  so the wave function changes sign there.

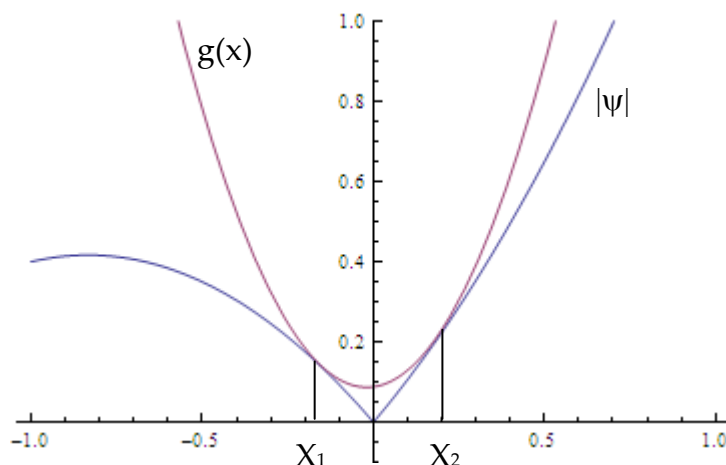


Figure I-1: The absolute value of the wave function and the parabola in it.

We plot in Figure I-1 a typical situation. We determine a large enough parameter  $a$  and the parabola:

$$p(x) = \frac{1}{2}a(x - x'_0)^2 + c$$

which has the properties that it is tangent at some  $x_1$  and  $x_2$  to  $|\psi(x)|$  (and where  $x_1 < x'_0 < x_2$ ). Then one defines a new wave function:

$$\chi(x) = \begin{cases} |\psi(x)| & x < x_1 \text{ or } x > x_2 \\ p(x) & x_1 < x < x_2 \end{cases} \quad (1.1.3)$$

By increasing  $a$  the parabola becomes narrow; adjusting  $x'_0$  accordingly we can cause  $x_1$  and  $x_2$  to uniformly approach 0 as close as needed. Under these conditions, it is possible to show that

$$-x_1 \approx x_2 \approx \frac{|\psi'(0)|}{a}$$

$$c \approx \frac{1}{2} \frac{|\psi'(0)|^2}{a}$$

Now,  $V[\chi] - V[\psi] \propto a^{-2}$ . One power of  $a^{-1}$  comes from integration over the interval of length  $x_2 - x_1$  where the two functions differ. The second power comes from the difference  $p(x_0) - \psi(x_0) = c \propto a^{-1}$ . The kinetic energy difference is *negative*:  $T[\chi] - T[\psi] \approx -\frac{\hbar^2}{2m} \frac{|\psi'(0)|^3}{a}$ . As  $a \rightarrow \infty$  the total change in energy is dominated by the change in the kinetic energy and is thus negative, showing that  $E[\chi] < E[|\psi|] = E[\psi]$ . This is a contradiction, since  $\psi$  is assumed the ground state. ♦

The proof can be extended to any number of dimensions. It can be used to prove that the ground state of a many-boson wave function has no nodes.

Prove an important immediate corollary from this theorem: the ground state of boson systems is non-degenerate.

## ii. Ground-state wave function of Fermions

XXX

### ***C. Why electronic structure is an important difficult problem***

In this course, we will study methods to treat electronic structure. This problem is often considerably more complicated than nuclear dynamics of molecules. The most obvious reason is that usually there are many more electrons than nuclei in a molecule. However, there are more reasons: electrons in molecules have extreme "quantum mechanical" character, while nuclei are more "classical". This last statement requires probably some

explanation. We will discuss in this course ways of mapping the electrons in a molecule onto a system of non-interacting particles. In such a picture, each electron "has" its own orbital or wave-function. Generally, the orbitals of different electrons in molecules strongly overlap. Furthermore, most of the valence orbitals are widely spread over space, encapsulating many molecular nuclei. Nuclei on the other hand, are massive, so their wave functions are strongly localized and hardly overlap once again, they behave much more like classical particles.

Describing electronic structure is therefore a quantum mechanical problem. Now, the electrons interact, via the attractive Coulomb force, not only with the stationary nuclei but also with each other: each electron repels each other electron via the repulsive Coulomb force. This makes the electronic structure problem a *many-body* problem.

Quantum many-body problems are very difficult to track down, much more difficult than classical mechanical ones. The reason is that for  $N_e$  electrons the many-body wave function is then a function of  $3N_e$  variables,  $\psi(\mathbf{r}_1 \dots \mathbf{r}_{N_e})$ , where  $\mathbf{r}_i$  ( $i=1 \dots N_e$ ) is 3 dimensional the position vector. Highly accurate numerical representation of such functions is next to impossible for  $N_e > 1$ . Thus, electronic structure methods invariably include approximate methods. Typically, the wave function encapsulates much more information than we care to know about. For example,  $|\psi(\mathbf{r}_1, \dots, \mathbf{r}_{N_e})|^2 d^3r_1 \dots d^3r_{N_e}$  gives the probability for finding an electron at  $\mathbf{r}_1$  and an electron at  $\mathbf{r}_2$  and an electron at  $\mathbf{r}_3$  etc. Indeed, the total probability is 1:

$$\int |\psi(\mathbf{r}_1, \dots, \mathbf{r}_{N_e})|^2 d^3r_1 \dots d^3r_{N_e} = 1 \quad (1.3.1)$$

## D. The Born-Oppenheimer theory

### i. The adiabatic theorem

Suppose a Hamiltonian is dependent parameterically on  $R = (R_1, R_2, \dots)$ . We write this as  $\hat{H}[R]$ . Denote the eigenstates of this Hamiltonian by

$$\hat{H}[R]\psi_n[R] = E_n[R]\psi_n[R] \quad (1.4.1)$$

Now, suppose we change  $R$  with time, so that we have a trajectory,  $R(t)$ . The Hamiltonian will become time dependent:  $\hat{H}(t) = \hat{H}[R(t)]$ . Suppose the system is placed in its ground state at time  $t = 0$ ,  $\phi(t = 0) = \psi_0[R(0)]$ . It will evolve according to the TDSE:

$$i\hbar\dot{\phi}(t) = \hat{H}(t)\phi(t) \quad (1.4.2)$$

Now, the adiabatic theorem says that if  $R(t)$  changes *very slowly* then:

$$\phi(t) = e^{i\theta(t)}\psi_0[R(t)] \quad (1.4.3)$$

Namely, except for a purely TD phase, the evolving state stays the instantaneous ground state of the Hamiltonian  $\hat{H}(t)$ .

To show this, we use the instantaneous eigenstates to expand  $\phi(t)$ :

$$\phi(t) = \sum_{n=0}^{\infty} a_n(t) e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} \psi_n[R(t)] \quad (1.4.4)$$

We want to plug this expansion into the SE. So let us calculate:

$$\begin{aligned} i\hbar\dot{\phi}(t) &= \sum_{n=0}^{\infty} [i\hbar\dot{a}_n + E_n a_n] e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} \psi_n[R(t)] \\ &\quad + \sum_{n=0}^{\infty} a_n e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} i\hbar\dot{\psi}_n[R(t)] \end{aligned} \quad (1.4.5)$$

And:

$$\hat{H}(t)\phi = \sum_{n=0}^{\infty} a_n e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} E_n \psi_n[R(t)] \quad (1.4.6)$$

Equating the two expressions gives:]

$$0 = \sum_{n=0}^{\infty} i\hbar \dot{a}_n e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} \psi_n[R(t)] + \sum_{n=0}^{\infty} a_n e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} i\hbar \dot{\psi}_n[R(t)] \quad (1.4.7)$$

Now let us multiply by  $\psi_m$  and integrate (remember that  $\langle \psi_n | \psi_m \rangle = \delta_{nm}$ ):

$$0 = i\hbar \dot{a}_m e^{-\frac{i}{\hbar} \int_0^t E_m(\tau) d\tau} + \sum_{n=0}^{\infty} a_n e^{-\frac{i}{\hbar} \int_0^t E_n(\tau) d\tau} i\hbar \langle \psi_m | \dot{\psi}_n \rangle \quad (1.4.8)$$

Now, we have:

$$0 = \frac{d}{dt} \langle \psi_m | \psi_n \rangle = \langle \psi_m | \dot{\psi}_n \rangle + \langle \dot{\psi}_m | \psi_n \rangle \quad (1.4.9)$$

And since  $\langle \dot{\psi}_m | \psi_n \rangle = \langle \psi_n | \dot{\psi}_m \rangle^*$ , we find:

$$\tau_{mn} = \langle \psi_m | \dot{\psi}_n \rangle = -\langle \psi_n | \dot{\psi}_m \rangle^* = -\tau_{nm}^* \quad (1.4.10)$$

The matrix  $\tau_{mn} = \langle \psi_m | \dot{\psi}_n \rangle$  is the matrix of “time-dependent non-adiabatic couplings (TDNACs)” and it is an anti-Hermitean matrix. Furthermore, for  $n = m$

$$Re \langle \psi_n | \dot{\psi}_n \rangle = 0 \quad (1.4.11)$$

Thus:

$$\dot{a}_m = -\tau_{mm} a_m - \sum_{n \neq m} e^{i \int_0^t \omega_{mn}(t') dt'} \tau_{mn} a_n \quad (1.4.12)$$

We see  $\tau_{mn}$  creates the non-adiabatic transitions, i.e. the transitions out of the ground state.

Let us play a bit with  $\tau$ . Take the time derivative of the TISE:

$$\hat{H}(t)\psi_n(t) + \dot{\hat{H}}(t)\psi_n(t) = \dot{E}_n(t)\psi_n(t) + E_n(t)\dot{\psi}_n(t) \quad (1.4.13)$$

Multiply by  $\langle \psi_m |$  and get:

$$\langle \psi_m | \dot{\hat{H}} | \psi_n \rangle + \langle \psi_m | \hat{H} | \dot{\psi}_n \rangle = \dot{E}_n(t) \langle \psi_m | \psi_n \rangle + E_n \langle \psi_m | \dot{\psi}_n \rangle. \quad (1.4.14)$$

Since:  $\langle \psi_m | \hat{H} | \psi_n \rangle = E_m \langle \psi_m | \psi_n \rangle$  we have:

$$\langle \psi_m | \dot{\hat{H}} | \psi_n \rangle = \dot{E}_n(t) \langle \psi_m | \psi_n \rangle + (E_n - E_m) \langle \psi_m | \dot{\psi}_n \rangle. \quad (1.4.15)$$

If  $n = m$  then:

$$\dot{E}_n(t) = \langle \psi_n | \dot{\hat{H}} | \psi_n \rangle. \quad (1.4.16)$$

This is called the **Hellmann-Feynman theorem**, showing that the power is the expectation value of the rate of change of the Hamiltonian. If  $n \neq m$ :

$$\tau_{mn} = \langle \psi_m | \dot{\psi}_n \rangle = \frac{\langle \psi_m | \dot{\hat{H}} | \psi_n \rangle}{E_n - E_m}. \quad (1.4.17)$$

This is called **Epstein's theorem**, giving an expression for the TDNACs. Remember that  $\hat{H}[R(t)]$  depends on  $t$  through the positions of the nuclei.

Thus:

$$\frac{d}{dt} \hat{H}[R(t)] = \dot{R}_N \frac{\partial}{\partial R_N} \hat{H}[R(t)] \quad (1.4.18)$$

Thus:

$$\tau_{mn} = \sum_N \frac{\langle \psi_m | \frac{\partial}{\partial R_N} \hat{H}[R(t)] | \psi_n \rangle}{E_n[R] - E_m[R]} \dot{R}_N(t) = \sum_N \tau_{mn}^N \dot{R}_N(t) \quad (1.4.19)$$

where:

$$\tau_{mn}^N[R] = \langle \psi_m | \frac{\partial}{\partial R_N} \psi_n \rangle = \frac{\langle \psi_m[R] | \frac{\partial}{\partial R_N} \hat{H}[R] | \psi_n[R] \rangle}{E_n[R] - E_m[R]} \quad (1.4.20)$$

We see that the TDNACs depend on the velocity of the nuclei. When nuclei are small the couplings out of the ground state are small. The  $\tau_{nm}^N$  are called the “non-adiabatic couplings (NACs)”. We find that as long as  $E_n - E_m$  are not zero on the trajectory, one can always find small enough  $\dot{R}$  that makes the TDNACs as small as we wish. All we need for states  $n$  and  $m$  to stay decoupled is for the following conditions to be met:

$$|\tau_{mn}| = |\dot{R}_N(t)\tau_{mn}^N(R_N(t))| \ll 1 \quad (1.4.21)$$

Thus we can make the NACs small. So all that is left in Eq. (1.4.12) is  $\dot{a}_m = -\tau_{mm}a_m$ . Define the non-dynamical phase as:  $\theta_{nd}(t) = i \int_0^t \tau_{mm}(t') dt'$  (note  $\theta$  is real) and then:

$$a_m(t) = e^{i\theta_{nd}(t)} \quad (\text{slow processes} = \text{adiabatic process}). \quad (1.4.22)$$

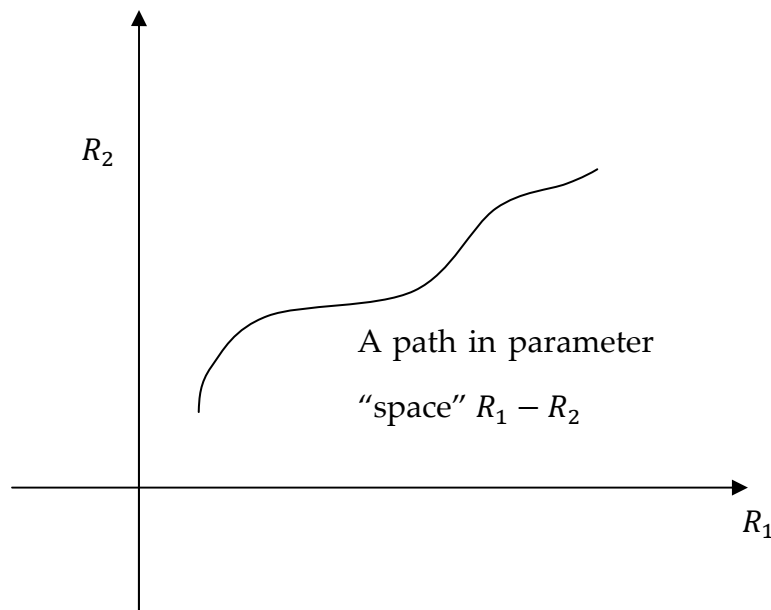
The total state is

$$\phi(t) = a_m(t)e^{-\frac{i}{\hbar}\int_0^t E_m(\tau)d\tau}\psi_m[R(t)] = e^{i(\theta_{nd}(t)+\theta_d(t))}\psi_m[R(t)] \quad (1.4.23)$$

where  $\theta_d(t) = -\frac{1}{\hbar}\int_0^t E_m(\tau)d\tau$  is called the dynamic phase. It is easy to prove that if  $R(t)$  traverses a closed loop, the non-dynamical phase depends only on that loop and not on the way it is traversed. For example, if we traverse the same loop using different velocities, the dynamic phase may change but the geometric phase will not. The closed loop geometric angle is called the **Berry phase**. The independence of path is a result of the fact that the non-dynamical phase is a **line integral**, and can be made with no reference to time:

$$\begin{aligned}
\theta_{nd}(t) &= i \int_0^t \langle \psi_m(R[\tau]) | \dot{\psi}_m(R[\tau]) \rangle d\tau \\
&= i \int_0^t \sum_N \langle \psi_m[R] | \frac{\partial}{\partial R_N} \psi_m[R] \rangle \dot{R}_N d\tau \quad (1.4.24) \\
&= i \int_{R(0)}^{R(t)} \sum_N \tau_{mm}^N(R) dR_N
\end{aligned}$$

The dynamical phase for example is not a line integral and its value depends not only on the path itself but also on the velocity taken along the way. This observation makes the non-dynamical phase a special quantity. Berry has shown that this quantity can give us information on the way the Hamiltonian is dependent on its parameters. For a real Hamiltonian, for example  $e^{i\theta_{nd}}$  around a closed path always equals either 1 or -1. If there is an even number of degeneracies enclosed by the path it is 1 and if an odd number – it is -1.



## ii. Motivation for the Born-Oppenheimer approximation: classical nuclei

In a molecule, we can think of the nuclei (having coordinates  $R$ ) as heavy and therefore classical particles which are slowly moving. The electrons are light and therefore quantum. They have coordinates  $r$  ( $r$  includes spin) and they



feel the effect of slowly moving nuclei. The electron dynamics is controlled by the Hamiltonian:

$$\hat{H}_e[R] = \hat{T}_r + V(r, R). \quad (1.4.25)$$

The potential  $V(r, R)$  describes the Coulombic interactions between electrons and nuclei (in atomic units):

$$V(r, R) = - \sum_{n,N} \frac{Z_N}{|\vec{R}_N - \vec{r}_n|} + \frac{1}{2} \sum_{n \neq m} \frac{1}{r_{nm}} + \frac{1}{2} \sum_{N \neq M} \frac{Z_N Z_M}{R_{NM}}. \quad (1.4.26)$$

In general, one wants to assume that the total energy of the molecule in this classical approximation is:

$$E = T_N + \langle \psi_e[R] | \hat{H}_e[R(t)] | \psi_e[R] \rangle \quad (1.4.27)$$

Where  $\psi_e[R]$  is the ground state of the electrons at nuclear configuration  $R$ . The adiabatic theorem states that this is reasonable as long as nuclei move slowly. Thus, the adiabatic theorem allows us to write:

$$E = T_N + V_e(R) \quad (1.4.28)$$

where  $V_e(R)$  is the ground state energy of the electrons at nuclear configuration  $R$ . This energy is a usual classical mechanical energy and the Newton equations of motion apply:

$$M_N \ddot{R}_N = F_N = - \frac{\partial}{\partial R_N} V_e(R) \quad (1.4.29)$$

We see, that the adiabatic theorem allows us to consider the nuclei as moving in a potential well which is essentially the eigensvalue of the electrons. This in essence is the BO approximation when the nuclei are classical.

### iii. The Born-Oppenheimer approximation in quantum nuclear case

The classical approach motivates a quantum treatment. We are expecting that nuclei will not excite electrons very efficiently. That is the motivation for the BO approximation.

The Born and Oppenheimer development is similar to that of the adiabatic theorem, although there are no “external” fields. Suppose we have a system of fast particles, with coordinates  $r = (r_1, r_2, \dots)$  and slow particles, with coordinates  $R = (R_1, R_2, \dots)$ . The Hamiltonian can be written as:

$$\hat{H} = \hat{T}_R + \hat{T}_r + V(r, R) \quad (1.4.30)$$

The Schrödinger equation is:

$$\hat{H}\psi_n(r, R) = E_n\psi_n(r, R) \quad (1.4.31)$$

Note that we can assume these wave functions are orthogonal:

$$\langle \psi_n | \psi_m \rangle = \int \int \psi_n^*(r, R) \psi_m(r, R) dr dR = \delta_{nm} \quad (1.4.32)$$

Now, to proceed, let us first consider the fast r-Hamiltonian

$$\hat{F}[R] = \hat{T}_r + V(r, R) \quad (1.4.33)$$

In this “fast Hamiltonian”  $\hat{F}[R]$ , the slow variables are simply parameters (it contains no derivatives with respect to  $R$ ). Thus, it depends on  $R$  *parametrically*. The fast eigenvalue problem is:

$$\hat{F}[R]\phi_k(r; R) = W_k(R)\phi_k(r; R) \quad (1.4.34)$$

The eigenvalues are functions of the parameters  $R$ . They are called the adiabatic (or Born-Oppenheimer) surfaces. The notation with the semicolon between  $r$  and  $R$  is designed to emphasize that  $\phi_k$  are wave functions in  $r$  but they depend only parametrically on  $R$ . This means, for example that the

overlap of the fast eigenstates (called adiabatic states) involves integration only over  $r$ , the dynamical coordinates, while  $R$  is held fixed, since it is only a parameter:

$$\langle \phi_k(R) | \phi_j(R) \rangle = \int \phi_k^*(r; R) \phi_j(r; R) dr = \delta_{kj} \quad (1.4.35)$$

We cannot in fact say anything really meaningful about  $\langle \phi_k(R) | \phi_j(R') \rangle$  when  $R \neq R'$  (except when  $R' - R$  is infinitesimal, but we will not go into this issue here). Now, we can expand the “real” wave-function as a linear combination of the adiabatic functions:

$$\psi_n(r, R) = \sum_k \phi_k(r; R) \Phi_{kn}(R) \quad (1.4.36)$$

We can do this because for any give  $R$   $\phi_k(r; R)$  span the space of wave functions dependent on  $r$ . In fact, the expansion coefficients,  $\Phi_{nk}(R)$  are given by:

$$\Phi_{kn}(R) = \int \phi_k^*(r; R) \psi_n(r, R) dr \quad (1.4.37)$$

Now, let us plug Eq. (1.4.36) into the SE (1.4.31):

$$\begin{aligned} E_n \psi_n(r, R) &= (\hat{T}_R + \hat{F}) \psi_n(r, R) \\ &= \sum_k \hat{T}_R [\phi_k(r; R) \Phi_{kn}(R)] + \sum_k W_k(R) \phi_k(r; R) \Phi_{kn}(R) \end{aligned} \quad (1.4.38)$$

Note that, since  $\hat{T}_R = \sum_N \frac{\hat{P}_N^2}{2M_N}$  where:  $\hat{P}_N = -i\hbar \frac{\partial}{\partial R_N}$ . We have:

$$\begin{aligned} \hat{T}_R [\phi_k(r; R) \Phi_{kn}(R)] &= \Phi_{kn}(R) \hat{T}_R [\phi_k(r; R)] \\ &+ \sum_N \frac{1}{M_N} (\hat{P}_N \phi_k(r; R)) (\hat{P}_N \Phi_{kn}(R)) \\ &+ \phi_k(r; R) \hat{T}_R \Phi_{kn}(R) \end{aligned} \quad (1.4.39)$$

Multiplying Eq. (1.4.38) by  $\phi_j(r; R)^*$  and integrating over  $r$  gives:

$$\sum_k (A_{jk} + \hat{B}_{jk}^N) \Phi_{kn}(R) + (\hat{T}_R + W_j(R)) \Phi_{jn}(R) = E_n \Phi_{jn}(R) \quad (1.4.40)$$

Where:

$$\begin{aligned} A_{jk} &= \int \phi_j(r; R)^* [\hat{T}_R \phi_k(r; R)] dr = \langle \phi_j | \hat{T}_R \phi_k \rangle \\ \hat{B}_{jk}^N &= \sum_N \frac{\int \phi_j(r; R)^* [\hat{P}_N \phi_k(r; R)] dr}{M_N} \hat{P}_N = \sum_N \frac{\langle \phi_j | \hat{P}_N \phi_k \rangle}{M_N} \hat{P}_N \\ &= - \sum_N \frac{\hbar^2}{M_N} \langle \phi_j | \frac{\partial}{\partial R_N} \phi_k \rangle \frac{\partial}{\partial R_N} = - \sum_N \frac{\hbar^2}{M_N} \tau_{jk}^N \frac{\partial}{\partial R_N} \end{aligned} \quad (1.4.41)$$

The matrices  $\tau_{jk}^N = \langle \phi_j | \frac{\partial}{\partial R_N} \phi_k \rangle$  are the non-adiabatic couplings in the “fast” system. These are exactly the non-adiabatic coefficients in the adiabatic theory (Eq. (1.4.20)).

It is possible to show that:

$$\begin{aligned} \sum_k (A_{jk} + \hat{B}_{jk}^N) \Phi_{kn}(R) + \hat{T}_R \Phi_{jn}(R) \\ = \sum_k \sum_N - \frac{\hbar^2}{2M_N} \left( \frac{\partial}{\partial R_N} + \tau_{jk}^N \right)^2 \Phi_{kn}(R) \end{aligned} \quad (1.4.42)$$

Thus Eq. (1.4.40) becomes:

$$\sum_k \sum_N - \frac{\hbar^2}{2M_N} \left( \frac{\partial}{\partial R_N} + \tau_{jk}^N \right)^2 \Phi_{kn}(R) + W_j(R) \Phi_{jn}(R) = E_n \Phi_{jn}(R) \quad (1.4.43)$$

This is a Schrödinger –like equation which determines the coefficients  $\Phi_{jn}(R)$ . These are called the “slow eigenfunctions”. Once they are computed one has an exact solution to the Schrödinger equation. However, we do not really want to solve this infinite set of coupled differential equations. Thus we assume that the quantities  $\tau_{jk}^N$  for  $j \neq k$  can be neglected. Note that here there

is no  $\dot{R}_N$  which can be taken as small as needed to make the effect of  $\tau_{jk}^N$  as minute as we need. Still we can hope that the fact that we chose  $R$  to be slow degrees of freedom allow us to make just this approximation! This was the idea of Born and Oppenheimer who neglected  $\tau_{jk}^N$ :

$$\sum_N -\frac{\hbar^2}{2M_N} \frac{\partial^2}{\partial R_N^2} \Phi_{jn}(R) + W_j(R) \Phi_{jn}(R) = E_n \Phi_{jn}(R) \quad (1.4.44)$$

The resulting equation is a Schrödinger equation for the slow degrees of freedom, which move under the force of electrons derived from a potential  $W_j(R)$ . When applied to molecules the slow degrees of freedom are usually the nuclei and the fast - the electrons. . There is a problem with neglecting  $\tau_{jj}^N$  because of their non-dynamical effect. Taking them into account results in treating them as a magnetic field:

$$\sum_N \frac{\hbar^2}{2M_N} \left( i \frac{\partial}{\partial R_N} - \tau_{mm} \right)^2 \Phi_{jn}(R) + W_j(R) \Phi_{jn}(R) = E_n \Phi_{jn}(R) \quad (1.4.45)$$

The BO approximation breaks the molecular SE into two consecutive tasks. First, the electronic SE Eq. (1.4.34) must be solved for any relevant clamped position of nuclei  $R$ . Then, the nuclear equation (1.4.26).

**Further reading on this subject:** M. Baer, Beyond Born-Oppenheimer: electronic non-adiabatic coupling terms and conical intersections (Wiley, Hoboken, N.J., 2006).

## ***E. Electron correlation***

### **i. The electronic wave function of two non-interacting electrons**

In order to appreciate the 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:

$$\hat{h} = -\frac{\hbar^2}{2m_e}\nabla^2 + v_{ext}(x) \quad (1.5.1)$$

Now consider a 2-electron problem. Assume we have two electrons, Fermions, which are non-interacting in the well. The Hamiltonian is

$$\hat{H} = \hat{h}(1) + \hat{h}(2) \quad (1.5.2)$$

The notation  $\hat{h}(i)$  means the Hamiltonian of Eq. (1.5.1) with the coordinate of the  $i$ -th electron.

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 spin-polarized, i.e. that the total spine is 1, both electrons are in spin-up orbitals. We try the following form as a wave function:

$$\Psi(x_1, x_2) = \frac{1}{\sqrt{2}}[\psi_0(x_1)\psi_1(x_2) - \psi_0(x_2)\psi_1(x_1)]\alpha(1)\alpha(2) \quad (1.5.3)$$

Notice that the spatial part is anti-symmetric while the spin part is symmetric. This renders the entire wave-function anti-symmetric, in accordance with the Pauli principle. The notation  $\alpha(1)\alpha(2)$  means both electrons have spin projection up ( $\alpha$ ). We do not yet know if and under what conditions this wave function can actually describe eigenstates of the two electrons the Hamiltonian (1.5.2). We assume that  $\psi_0(x)$  and  $\psi_1(x)$  are orthonormal. This makes the state normalized, since:

$$\begin{aligned}
\int \Psi(x_1, x_2)^2 dx_1 dx_2 &= \frac{1}{2} \int (\psi_0(x_1)\psi_1(x_2) - \psi_0(x_2)\psi_1(x_1))^2 dx_1 dx_2 \\
&= \frac{1}{2} \int [(\psi_0(x_1)\psi_1(x_2))^2 + (\psi_0(x_2)\psi_1(x_1))^2 \\
&\quad - 2\psi_0(x_1)\psi_1(x_2)\psi_0(x_2)\psi_1(x_1)] dx_1 dx_2
\end{aligned} \tag{1.5.4}$$

The first and second terms are both equal to:

$$\begin{aligned}
\int (\psi_0(x_1)\psi_1(x_2))^2 dx_1 dx_2 &= \int \psi_0(x_1)^2 dx_1 \int \psi_1(x_2)^2 dx_2 = 1 \times 1 \\
&= 1
\end{aligned} \tag{1.5.5}$$

The third term in (1.5.4) is:

$$\begin{aligned}
\int \psi_0(x_1)\psi_1(x_2)\psi_0(x_2)\psi_1(x_1) dx_1 dx_2 \\
= \int \psi_0(x_1)\psi_1(x_1) dx_1 \int \psi_0(x_2)\psi_1(x_2) dx_2 = 0 \times 0 = 0
\end{aligned} \tag{1.5.6}$$

Indeed the entire wave function is orthonormal (thanks to the factor  $1/\sqrt{2}$  in (1.5.3)).

Now, let us see under what condition the wave function in Eq. (1.5.3) is an eigenstate of the Hamiltonian in (1.5.2)

$$\begin{aligned}
\hat{H}\Psi(x_1, x_2) &= (\hat{h}(1) + \hat{h}(2))\Psi(x_1, x_2) \\
&= \frac{1}{\sqrt{2}}\{\psi_1(x_2)\hat{h}(1)\psi_0(x_1) + \psi_0(x_1)\hat{h}(2)\psi_1(x_2)\} \\
&\quad - \frac{1}{\sqrt{2}}\{\psi_0(x_2)\hat{h}(1)\psi_1(x_1) + \psi_1(x_1)\hat{h}(2)\psi_0(x_2)\}
\end{aligned} \tag{1.5.7}$$

This should be equated to:

$$E\Psi(x_1, x_2) = \frac{1}{\sqrt{2}}\{E\psi_0(x_1)\psi_1(x_2) - E\psi_0(x_2)\psi_1(x_1)\} \tag{1.5.8}$$

If we choose the orbitals  $\psi_0(x)$  and  $\psi_1(x)$  to be eigenstates of  $\hat{h}$  (which are orthogonal so that is compatible with our previous assumption):

$$\hat{h}\psi_i(x) = \epsilon_i\psi_i(x), \quad i = 0,1, \dots \quad (1.5.9)$$

Thus:

$$\begin{aligned} \hat{H}\Psi(x_1, x_2) &= \frac{1}{\sqrt{2}} \{ \epsilon_0\psi_1(x_2)\psi_0(x_1) + \epsilon_1\psi_0(x_1)\psi_1(x_2) \} \\ &\quad - \frac{1}{\sqrt{2}} \{ \epsilon_1\psi_0(x_2)\psi_1(x_1) + \epsilon_0\psi_1(x_1)\psi_0(x_2) \} \\ &= (\epsilon_0 + \epsilon_1)\Psi(x_1, x_2) \end{aligned} \quad (1.5.10)$$

And we see that indeed  $\Psi(x_1, x_2)$  is an eigenstate with energy  $E = \epsilon_0 + \epsilon_1$ .

## ii. Correlation in action: a wave function of 2 *interacting* particles in an Harmonic trap

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:

$$v_{ext}(x) = \frac{1}{2}\mu\omega^2x^2 \quad (1.6.1)$$

The two lowest eigenstates of a Harmonic oscillator are:

$$\psi_0(x) = N_0 e^{-\frac{1\mu\omega}{2\hbar}x^2} \quad \psi_1(x) = N_1 \sqrt{\frac{\mu\omega}{\hbar}} x e^{-\frac{1\mu\omega}{2\hbar}x^2} \quad (1.6.2)$$

The normalization constants are:

$$N_0 = \left(\frac{\mu\omega}{\pi\hbar}\right)^{1/4}, \quad N_1 = \left(\frac{4\mu\omega}{\pi\hbar}\right)^{1/4} \quad (1.6.3)$$

And the eigenvalues are:

$$\epsilon_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad n = 0,1, \dots \quad (1.6.4)$$



The groundstate energy of the 2-electron system in the triplet state will be placing one electron in  $\psi_0(x)$  and another in  $\psi_1(x)$ :

$$E = \epsilon_0 + \epsilon_1 = 2\hbar\omega \quad (1.6.5)$$

As discussed above singlet and triplet two-electron ground state wave functions composed of two orbitals must be space-symmetric or antisymmetric, respectively. We consider below 3 wave functions. The first  $\Psi_{S,00}$  is the ground state singlet where both electrons are in  $\psi_0$ . The second and third are a singlet and a triplet made from one electron in  $\psi_0$  and the other in  $\psi_1$ :

$$\begin{aligned} \Psi_{S,00}(x_1, x_2) &= \psi_0(x_1)\psi_0(x_2) = N_{00}e^{-\frac{1\mu\omega}{2\hbar}(x_1^2+x_2^2)} \\ \Psi_{S,01}(x_1, x_2) &= \frac{1}{\sqrt{2}}[\psi_0(x_1)\psi_1(x_2) + \psi_0(x_2)\psi_1(x_1)] \\ &= N_{01}e^{-\frac{1\mu\omega}{2\hbar}(x_1^2+x_2^2)}(x_2 + x_1) \\ \Psi_{T,01}(x_1, x_2) &= \frac{1}{\sqrt{2}}[\psi_0(x_1)\psi_1(x_2) - \psi_0(x_2)\psi_1(x_1)] \\ &= N_{01}e^{-\frac{1\mu\omega}{2\hbar}(x_1^2+x_2^2)}(x_2 - x_1) \end{aligned} \quad (1.6.6)$$

The normalization factors are:

$$N_{00} = N_0^2 = \sqrt{\frac{\mu\omega}{\pi\hbar}} \quad N_{01} = \frac{1}{\sqrt{2}}\sqrt{\frac{\mu\omega}{\hbar}} N_0 N_1 = \frac{\mu\omega}{\pi\hbar} \quad (1.6.7)$$

Eq. (1.6.6) describes the distribution of positions of both electrons in their corresponding states. We now want to ask, 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 quantity  $\langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle$ . If electrons are completely unaware of each other, this quantity, called the position-autocorrelation function 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. For the triplet wave function in Eq. (1.6.6) we find:

$$\begin{aligned}
\langle x_1 \rangle &= \langle \Psi_{T,01} | x_1 | \Psi_{T,01} \rangle = \frac{1}{2} \langle \psi_0 \psi_1 - \psi_1 \psi_0 | x_1 | \psi_0 \psi_1 - \psi_1 \psi_0 \rangle \\
&= \frac{1}{2} (\langle \psi_0 | x_1 | \psi_0 \rangle \langle \psi_1 | \psi_1 \rangle + \langle \psi_1 | x_1 | \psi_1 \rangle \langle \psi_0 | \psi_0 \rangle \\
&\quad - \langle \psi_0 | x_1 | \psi_1 \rangle \langle \psi_1 | \psi_0 \rangle - \langle \psi_1 | x_1 | \psi_0 \rangle \langle \psi_0 | \psi_1 \rangle) \\
&= \frac{1}{2} (\langle \psi_0 | x_1 | \psi_0 \rangle + \langle \psi_1 | x_1 | \psi_1 \rangle) = 0
\end{aligned} \tag{1.6.8}$$

Of-course, the same result would be obtained if we calculate  $\langle x_2 \rangle$  because the electrons are equivalent. Furthermore:

$$\begin{aligned}
\langle x_1 x_2 \rangle &= \langle \Psi_{T,01} | x_1 x_2 | \Psi_{T,01} \rangle = \frac{1}{2} \langle \psi_0 \psi_1 - \psi_1 \psi_0 | x_1 x_2 | \psi_0 \psi_1 - \psi_1 \psi_0 \rangle \\
&= \frac{1}{2} (\langle \psi_0 | x_1 | \psi_0 \rangle \langle \psi_1 | x_2 | \psi_1 \rangle + \langle \psi_1 | x_2 | \psi_1 \rangle \langle \psi_0 | x_1 | \psi_0 \rangle \\
&\quad - \langle \psi_0 | x_1 | \psi_1 \rangle \langle \psi_1 | x_2 | \psi_0 \rangle - \langle \psi_1 | x_1 | \psi_0 \rangle \langle \psi_0 | x_2 | \psi_1 \rangle) \\
&= -|\langle \psi_0 | x | \psi_1 \rangle|^2 = -\frac{1}{2} \frac{\hbar}{\mu \omega}
\end{aligned} \tag{1.6.9}$$

This negative quantity is there because the Pauli principle pushes the electrons to opposite sides (when one electron has positive x coordinate the other has negative and vice versa). Let's see what happens in the singlet wave function  $\Psi_{S,00}$ . Here too  $\langle x_1 \rangle = \langle x_2 \rangle = 0$ . Then:

$$\begin{aligned}
\langle x_1 x_2 \rangle &= \langle \Psi_{S,00} | x_1 x_2 | \Psi_{S,00} \rangle = \langle \psi_0 \psi_0 | x_1 x_2 | \psi_0 \psi_0 \rangle \\
&= \langle \psi_0 | x_1 | \psi_0 \rangle \langle \psi_0 | x_2 | \psi_0 \rangle = 0
\end{aligned} \tag{1.6.10}$$

Thus, the singlet ground state shows no correlation. However, this does not mean that all singlet wave functions of 2 electrons have no correlation. Indeed, let us study the situation in  $\Psi_{S,01}$ . The development is very similar to Eq. (1.6.9), except that the minus sign is now a plus sign so:

$$\begin{aligned}
\langle x_1 x_2 \rangle &= \langle \Psi_{S,01} | x_1 x_2 | \Psi_{S,01} \rangle = \frac{1}{2} \langle \psi_0 \psi_1 + \psi_1 \psi_0 | x_1 x_2 | \psi_0 \psi_1 + \psi_1 \psi_0 \rangle \\
&= |\langle \psi_0 | x | \psi_1 \rangle|^2 = \frac{1}{2} \frac{\hbar}{\mu \omega}
\end{aligned}
\tag{1.6.11}$$

Here we find positive autocorrelation, indicative of the fact that spin-opposite non-interacting pairs of electrons want to stick together (when one has positive  $x$  the other wants to have as well and when one has negative the other one wants negative as well) like non-interacting bosons.

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) Wavefunctions that are mere products of single-particle orbitals have no correlation.
- 2) If the products are symmetrized like in the case of the excited singlet the correlation  $\langle x_1 x_2 \rangle$  is positive indicating that the particles “like to be together” i.e. both on the right or both on the left of the origin.
- 3) If the products are anti-symmetrized like in the case of the triplet the correlation  $\langle x_1 x_2 \rangle$  is negative indicating that the particles “like to be separated”.

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

$$\hat{H} = \hat{h}(1) + \hat{h}(2) + V(x_1, x_2)
\tag{1.6.12}$$

Let us take one case of coupling which is simple enough to yield to analytical analysis:

$$V(x_1, x_2) = \mu \gamma^2 x_1 x_2
\tag{1.6.13}$$

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: since if  $x_1$  and  $x_2$  are both large and of the same sign this is energy-costly; if they are both large and of opposite sign that lowers energy. In this case, the Hamiltonian is:

$$\hat{H} = \frac{p_1^2}{2\mu} + \frac{p_2^2}{2\mu} + \frac{1}{2}\mu\omega^2 x_1^2 + \frac{1}{2}\mu\omega^2 x_2^2 + \mu\gamma^2 x_1 x_2; \quad (1.6.14)$$

### Exercise I-1

Find the eigenvalues and eigenfunctions of this Hamiltonian.

**Solution:** Define new coordinates  $X$  and  $x$  by:  $\frac{X+x}{\sqrt{2}} = x_1$  and  $\frac{X-x}{\sqrt{2}} = x_2$ . The conjugate momenta are:  $p_1 = \frac{P+p}{\sqrt{2}}$  and  $p_2 = \frac{P-p}{\sqrt{2}}$  (show that this is indeed so by calculating the commutation relations  $[P, X]$ ,  $[P, x]$  etc.). Then the new Hamiltonian is:

$$\begin{aligned} \hat{H} = & \frac{(P+p)^2}{4\mu} + \frac{(P-p)^2}{4\mu} + \frac{1}{2}\mu\omega^2 \left(\frac{X+x}{\sqrt{2}}\right)^2 + \frac{1}{2}\mu\omega^2 \left(\frac{X-x}{\sqrt{2}}\right)^2 \\ & + \mu\gamma^2 \left(\frac{X+x}{\sqrt{2}}\right) \left(\frac{X-x}{\sqrt{2}}\right) \end{aligned} \quad (1.6.15)$$

Or, after rearranging:

$$\hat{H} = \left[ \frac{P^2}{2\mu} + \frac{1}{2}\mu(\omega^2 + \gamma^2)X^2 \right] + \left[ \frac{p^2}{2\mu} + \frac{1}{2}\mu(\omega^2 - \gamma^2)x^2 \right] \quad (1.6.16)$$

We see that  $X$  and  $x$  do not interact and each is a Harmonic oscillator on its own. Let us define  $\cos \theta = \left(\frac{\gamma}{\omega}\right)^2$  and  $\Omega_1 = \sqrt{\omega^2 + \gamma^2} = \sqrt{2}\omega \cos \frac{\theta}{2}$  and  $\Omega_2 = \sqrt{\omega^2 - \gamma^2} = \sqrt{2}\omega \sin \frac{\theta}{2}$ . We find:

$$\begin{aligned} \Psi_{nm}(X, x) &= \psi_n(X; \Omega_1, \mu) \psi_m(x; \Omega_2, \mu), \\ E_{nm} &= \left(n + \frac{1}{2}\right) \hbar \Omega_1 + \left(m + \frac{1}{2}\right) \hbar \Omega_2 \end{aligned} \quad (1.6.17)$$

With  $\psi_n(z; \Omega, \mu) = N_n e^{-\frac{q^2}{2}} H_n(q)$ ,  $n = 0, 1, \dots$ ,  $q = \sqrt{\frac{\mu\Omega}{\hbar}} z$ ,  $H_n(q)$  are Hermite polynomials of order  $n$  and  $N_n$  are orthonormalization constants.

### Exercise I-2

Write 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  $r = \frac{E_{gs,\gamma}}{E_{gs,\gamma=0}}$  and on the correlation function  $c = \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle$ .

**Solution:** We need to impose upon the spatial wave function of (1.6.17) to be antisymmetric. The two variables  $X = \frac{x_1+x_2}{\sqrt{2}}$  and  $x = \frac{x_1-x_2}{\sqrt{2}}$  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  $n = 0$  for  $X$  and  $m = 1$  for  $x$ :

$$\Psi_{01}(X, x) = N_{01} e^{-\frac{\mu\Omega_1 X^2}{2\hbar} - \frac{\mu\Omega_2 x^2}{2\hbar}} x$$
$$E_{gs} = \frac{1}{2} \hbar\Omega_1 + \frac{3}{2} \hbar\Omega_2 = \frac{1}{2} \hbar\omega \times \sqrt{2} \left( \cos \frac{\theta}{2} + 3 \sin \frac{\theta}{2} \right)$$
(1.6.18)

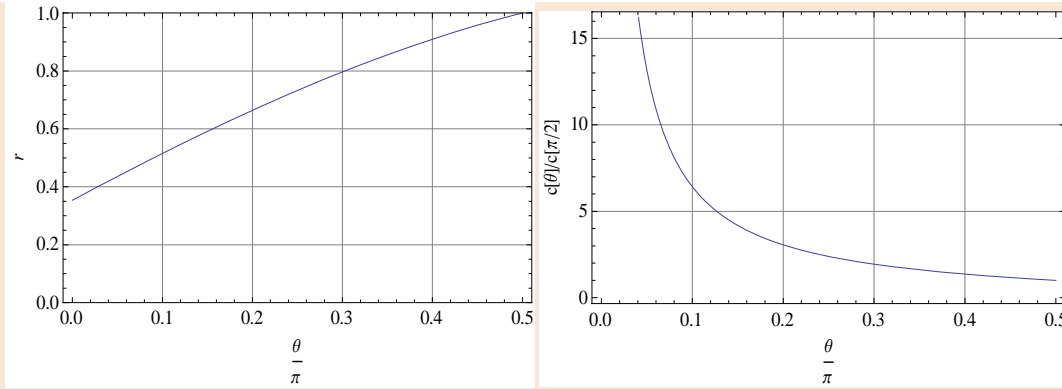
Now, let us write the wave function of Eq. (1.6.18) in terms of  $x_1$  and  $x_2$ :

$$\Psi_{01}(x_1, x_2) = N_{01} e^{-\frac{\mu\Omega_1}{4\hbar}(x_1+x_2)^2 - \frac{\mu\Omega_2}{4\hbar}(x_1-x_2)^2} \frac{(x_1 - x_2)}{\sqrt{2}}$$
$$= \frac{N_{01}}{\sqrt{2}} e^{-\frac{\mu(\Omega_1+\Omega_2)}{4\hbar}(x_1^2+x_2^2)} e^{-\frac{\mu(\Omega_1-\Omega_2)}{2\hbar}x_1x_2} (x_1 - x_2)$$
(1.6.19)

One can compare the effect of the interaction by looking at the ratio between ground-state energies of the system with and without interaction as a function of  $\theta$ , the interaction strength:

$$r = \frac{E_{gs,\gamma}}{E_{gs}} = \frac{\frac{1}{2}\Omega_1 + \frac{3}{2}\Omega_2}{2\omega} = \frac{\cos \frac{\theta}{2} + 3 \sin \frac{\theta}{2}}{2\sqrt{2}}$$
(1.6.20)

The result is shown in Figure I-2 (left panel).



**Figure I-2: (Left panel) The ratio  $r = E_{12,\gamma}/E_{12}$ ; (Right panel) the correlation ratio  $c(\theta)/c(\frac{\pi}{2})$  vs interaction strengths  $\cos \theta = (\frac{\gamma}{\omega})^2$**

The interaction lowers the energy, because now the wave function can acquire a structure that promotes the 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  $x_1x_2$ . To see this note that the expectation values of  $x$  and  $X$  are both zero and therefore  $\langle x_1 \rangle$  and  $\langle x_2 \rangle$  are zero as well. Furthermore note that  $X^2 - x^2 = 2x_1x_2$  and the auto-correlation is:

$$c(\theta) = \langle x_1x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle = \langle x_1x_2 \rangle = \frac{1}{2}(\langle X^2 \rangle - \langle x^2 \rangle) \quad (1.6.21)$$

The expectation value of the square position in harmonic oscillator is easily obtained using Hellman-Feynman theorem (rederive Eq. (1.4.16) in terms of any parameter-dependent Hamiltonian, not necessarily the time  $t$ ):

$$H = T + \frac{1}{2}\mu\omega^2 y^2 \rightarrow \left(n + \frac{1}{2}\right)\hbar = \frac{dE_n}{d\omega} = \mu \omega \langle y^2 \rangle_n \quad (1.6.22)$$

So:  $\langle X^2 \rangle_0 = \frac{\frac{1}{2}\hbar}{\mu\Omega_1}$  and similarly  $\langle x^2 \rangle_1 = \frac{\frac{3}{2}\hbar}{\mu\Omega_2}$ . Thus:

$$\begin{aligned}
 c(\theta) &= \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle = \langle x_1 x_2 \rangle = \frac{1}{2} \langle X^2 - x^2 \rangle = \frac{1}{2} \frac{\hbar}{\mu} \left( \frac{1}{2} \frac{1}{\Omega_1} - \frac{3}{2} \frac{1}{\Omega_2} \right) \\
 &= \frac{\hbar}{4\sqrt{2}\mu\omega} \left( \frac{1}{\cos\frac{\theta}{2}} - \frac{3}{\sin\frac{\theta}{2}} \right)
 \end{aligned}
 \tag{1.6.23}$$

For no interaction ( $\theta = \frac{\pi}{2}$ ), there is only correlation due to the Pauli principle: since both electrons have spin up they cannot occupy the same point in space)

We see, as derived above:

$$c\left(\frac{\pi}{2}\right) = \text{Fermi Correlation} = -\frac{\hbar}{2\mu\omega}
 \tag{1.6.24}$$

The ratio between full correlations and Fermi correlation is shown in Figure I-2 (right panel). The correlation is negative and larger (in absolute value) than the mere Pauli correlation. Indeed, the interaction pushes electrons away from each other.

Looking at the wave function of Eq. (1.6.19), it is evident that because  $\Omega_1 \neq \Omega_2$  there is no way to write this as an antisymmetrized sum of products of 1-electron functions. From these exercises with harmonic oscillator systems we find that the issue of correlation can be quite complicated. In realistic electronic systems, when the interaction is Coulombic, not Harmonic, the situation is even more complicated because of the lack of analytical solutions.

***F. The electron density is a much simpler object than the wave function***

The complexity of the wave function is overwhelming. It includes all the information we can have on the molecule in a certain state. However, all these intricacies and details are often uninteresting for us: in many cases, we simply have no direct use for them. Take, for instance, the electronic energy - our primary target in the Born-Oppenheimer picture. It depends only on the

relative distance of pairs of particles. This is because the electron-electron Coulomb repulsion is a *pair-wise interaction*.

One interesting quantity, besides energy, that can be extracted from the electronic wave function is the electron density  $n(\mathbf{r})$ . This 3D function tells us the expectation value of the density of electrons. That is  $n(\mathbf{r})d^3r$  is the expectation value of the number of electrons at a small volume  $d^3r$  around point  $\mathbf{r}$ . Thus, we can write:

$$n(\mathbf{r}) = \langle \psi | \hat{n}(\mathbf{r}) | \psi \rangle \quad (1.7.1)$$

We use the notation:

$$\langle \psi | \phi \rangle \equiv \int \psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) \phi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) d^3r_1 d^3r_2 \dots d^3r_{N_e} \quad (1.7.2)$$

Here  $\hat{n}(\mathbf{r})$  is the operator corresponding to the electron number density. Since electrons are point particles, and the position operator for the  $i^{\text{th}}$  electron is  $\mathbf{r}_i$  this operator is defined by:

$$\hat{n}(\mathbf{r}) = \sum_{i=1}^{N_e} \delta(\hat{\mathbf{r}}_i - \mathbf{r}) \quad (1.7.3)$$

We used the definition of a  $\delta$ -function, according to which:

$$\int \delta(\mathbf{r}_1 - \mathbf{r}) f(\mathbf{r}_1) d^3r_1 = f(\mathbf{r}) \quad (1.7.4)$$

The "function"  $\delta(\hat{\mathbf{r}}_i - \mathbf{r})$  is the density of electron  $i$  at  $\mathbf{r}$ .

### Exercise I-3

Calculate the expectation value of the electron number density and show that:

$$n(\mathbf{r}) = N_e \int \psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) \psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) d^3r_2 \dots d^3r_{N_e} \quad (1.7.5)$$

**Solution:** Plugging (1.7.3) into (1.7.1) gives:



$$\begin{aligned}
n(\mathbf{r}) &= \int \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) \hat{n}(\mathbf{r}) \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) d^3r_1 d^3r_2 \dots d^3r_{N_e} \\
&= \sum_{i=1}^{N_e} \int \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) \delta(\hat{\mathbf{r}}_i \\
&\quad - \mathbf{r}) \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) d^3r_1 \dots d^3r_{N_e} = \\
&= \sum_{i=1}^{N_e} \int \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) \delta(\hat{\mathbf{r}}_1 \\
&\quad - \mathbf{r}) \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) d^3r_1 \dots d^3r_{N_e}
\end{aligned} \tag{1.7.6}$$

The last step stems from the Pauli principle: all electrons are identical, so we can replace electron  $\mathbf{r}_i$  by electron  $\mathbf{r}_1$ . The sum is now over identical numbers so it becomes a mere multiplication as in Eq. (1.7.5).

Looking at Eq. (1.7.5), we see that  $n(\mathbf{r})$  involves integrating out a huge amount of wave-function details. It is as if only the data concerning the density distribution of a single electron remains! This is multiplied by  $N_e$  in Eq. (1.7.5) so  $n(\mathbf{r})$  accounts for the combined density of all electrons. Indeed, integrating over the entire space, one obtains from (1.7.5):

$$\int n(\mathbf{r}) d^3r = N_e \tag{1.7.7}$$

Expressing the fact that the total number of electrons is  $N_e$ .

#### Exercise I-4

Calculate the 1D electron density of the triplet ground state from Exercise I-2.

**Solution:** If the wavefunction of Eq. (1.6.6) is taken then:

$$\begin{aligned}
n(x) &= N \int_{-\infty}^{\infty} e^{-\frac{\mu\omega(x^2+x_2^2)}{\hbar}} (x_2 - x)^2 dx \\
&= N e^{-\frac{\mu\omega x^2}{\hbar}} \left[ \int_{-\infty}^{\infty} e^{-\frac{\mu\omega x_2^2}{\hbar}} x_2^2 dx_2 + x^2 \int_{-\infty}^{\infty} e^{-\frac{\mu\omega x_2^2}{\hbar}} dx_2 \right] \\
&= N \sqrt{\frac{\hbar\pi}{\mu\omega}} \int_{-\infty}^{\infty} e^{-\frac{\mu\omega x_2^2}{\hbar}} dx_2 [(x^2) + x^2] \\
&= N e^{-\frac{\mu\omega x^2}{\hbar}} \sqrt{\frac{\hbar\pi}{\mu\omega}} \left[ \frac{\hbar}{2\mu\omega} + x^2 \right]
\end{aligned} \tag{1.7.8}$$

We choose  $N$  to ensure that  $\int n(\mathbf{r})d^3r = 2$ :

$$n(x) = \sqrt{\frac{\mu\omega}{\pi\hbar}} e^{-\frac{\mu\omega x^2}{\hbar}} \left[ 1 + \frac{2\mu\omega}{\hbar} x^2 \right] \tag{1.7.9}$$

Defining the “average frequency”:

$$\frac{1}{\Omega} = \frac{1}{2} \left( \frac{1}{\Omega_1} + \frac{1}{\Omega_2} \right) = \frac{1}{\sqrt{2}\omega} \frac{\cos\frac{\theta}{2} + \sin\frac{\theta}{2}}{\sin\theta} \tag{1.7.10}$$

We find the density of the state in Eq. (1.6.19) is:

$$\begin{aligned}
n(x) &= N \int_{-\infty}^{\infty} e^{-\frac{\Omega_1(x+x_2)^2}{2} - \frac{\Omega_2(x-x_2)^2}{2}} (x - x_2)^2 dx_2 \\
&= \sqrt{\frac{\Omega^3}{\pi\Omega_1^2}} e^{-\Omega x^2} (1 + 2(2\Omega_1 - \Omega)x^2)
\end{aligned} \tag{1.7.11}$$

### ***G. The variational principle***

When we look for the ground-state energy of a complicated system, with Hamiltonian  $\hat{H}$ , the variational principle is often extremely important. It says, quite simply that the ground-state energy is a minimum of a functional of wave functions. The functional is:

$$E[\phi] = \frac{\langle \phi | \hat{H} | \phi \rangle}{\langle \phi | \phi \rangle} \tag{1.8.1}$$

And the variational principle states that:

$$E_{gs} \equiv E[\phi_{gs}] \leq E[\phi] \quad (1.8.2)$$

For all  $\phi$ . What this principle allows us is to determine which of the two wave functions  $\phi_1$  and  $\phi_2$  gives an energy closer to  $E_{gs}$ , without knowing  $E_{gs}$ . Simple the one with lower energy is closer.

We can build a parameterized family of functions  $\phi_\gamma$  and determine the "optimized" parameter as the one that gives minimum to  $E(\gamma) \equiv E[\phi_\gamma]$ .

#### Exercise I-5

Consider the quantum ground-state problem for a particle in a well  $v(x) = \frac{1}{4}kx^4$ . We consider the family of functions:

$$\phi_\sigma(x) = \frac{e^{-\frac{x^2}{4\sigma^2}}}{\sqrt{\sqrt{2\pi}\sigma}} \quad (1.8.3)$$

Here  $\sigma$  is a positive parameter. What is the best function for representing the ground state?

**Solution:** The functions are normalized. The energy is  $\hat{H} = \hat{T} + \hat{V}$ . We have:

$$E(\sigma) = E[\phi_\sigma] = \langle \phi_\sigma | \hat{T} | \phi_\sigma \rangle + \langle \phi_\sigma | \hat{V} | \phi_\sigma \rangle = \frac{\hbar^2}{8m\sigma^2} + \frac{3}{4}k\sigma^4 \quad (1.8.4)$$

Where we used the fact:

$$\langle \phi_\sigma | \hat{T} | \phi_\sigma \rangle = -\frac{\hbar^2}{2m_e} \int_{-\infty}^{\infty} \phi_\sigma(x) \phi_\sigma''(x) dx = \frac{\hbar^2}{8m_e\sigma^2} \quad (1.8.5)$$

Thus the minimum energy is obtained from:

$$0 = E'(\sigma) = -\frac{\hbar^2}{4m_2\sigma'^3} + 3k\sigma^{*3} \rightarrow \sigma^* = \left( \frac{\hbar^2}{12km} \right)^{1/6} \quad (1.8.6)$$

And so:

$$E^* = \frac{\hbar^2}{8m \left(\frac{\hbar^2}{12km}\right)^{1/3}} + \frac{3}{4}k \left(\frac{\hbar^2}{12km}\right)^{2/3} = \frac{3\hbar^2}{16m} \frac{1}{\sigma^{*2}} \quad (1.8.7)$$

### H. Dilation relations

Consider the normalized  $\psi(x)$ . From it, we can define a family of normalized wavefunctions:

$$\psi_\gamma(x) = \gamma^{1/2} \psi(\gamma x) \quad (1.9.1)$$

We can check that each member is indeed normalized:

$$\langle \psi_\gamma | \psi_\gamma \rangle = \int |\psi_\gamma(x)|^2 dx = \gamma \int |\psi(\gamma x)|^2 dx = \int |\psi(y)|^2 dy \quad (1.9.2)$$

This operation of "stretching" the argument of a function is called dilation.

Dilation affects functionals. Consider the kinetic energy functional:

$$T[\psi] \equiv \langle \psi | \hat{T} | \psi \rangle = \int \psi(x) \left( -\frac{1}{2} \frac{d^2}{dx^2} \right) \psi(x) dx \quad (1.9.3)$$

#### Exercise I-6

Prove the dilation relation of  $T[\psi_\gamma]$  and  $T[\psi]$ :

$$T[\psi_\gamma] = \gamma^2 T[\psi] \quad (1.9.4)$$

**Solution:** We show that this is a simple relation:

$$\begin{aligned} T[\psi_\gamma] &\equiv \int \psi_\gamma(x) \left( -\frac{1}{2} \frac{d^2}{dx^2} \right) \psi_\gamma(x) dx \\ &= \gamma \int \psi(\gamma x) \left( -\frac{1}{2} \frac{d^2}{dx^2} \right) \psi(\gamma x) dx \\ &= \gamma^2 \int \psi(y) \left( -\frac{1}{2} \frac{d^2}{dy^2} \right) \psi(y) dy = \gamma^2 T[\psi] \end{aligned} \quad (1.9.5)$$

The potential between all particles is called "homogeneous of order  $n$ " if:

$$V(\gamma x_1, \dots, \gamma x_N) = \gamma^n V(x_1, \dots, x_N) \quad (1.9.6)$$

Examples are  $n = 2$  is for Harmonic potential wells and harmonic interactions and  $n = -1$  is for the Coulomb potential wells and Coulombic interactions.

**Exercise I-7**

Prove the dilation relation for homogeneous potentials of order  $n$ :

$$V[\psi_\gamma] = \gamma^{-n}V[\psi] \quad (1.9.7)$$

**Solution:** For such a potential,

$$\begin{aligned} V[\psi_\gamma] &= \int \gamma^N |\psi_\gamma(\gamma x_1, \dots, \gamma x_N)|^2 V(x_1, \dots, x_N) dx_1 \dots dx_N = \\ &= \int |\psi_\gamma(y_1, \dots, y_N)|^2 V\left(\frac{y_1}{\gamma}, \dots, \frac{y_N}{\gamma}\right) dy_1 \dots dy_N \\ &= \gamma^{-n} \int |\psi_\gamma(y_1, \dots, y_N)|^2 V(y_1, \dots, y_N) dy_1 \dots dy_N \\ &= \gamma^{-n}V[\psi_1]. \end{aligned} \quad (1.9.8)$$

We combine the results from Eqs. (1.9.4) and (1.9.7) and obtain an interesting property of the total energy for systems with homogeneous interactions:

$$E[\psi_\gamma] = T[\psi_\gamma] + V[\psi_\gamma] = \gamma^2 T[\psi] + \gamma^{-n} V[\psi] \quad (1.9.9)$$

For a molecule, the interaction between the particles (electrons and nuclei) is the Coulomb potential  $V(\mathbf{r}; \mathbf{R}) \equiv V(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}; \mathbf{R}_1, \dots, \mathbf{R}_{N_N})$  which is homogeneous of order  $n = -1$  one finds the energy of a molecule obeys:

$$E[\psi_\gamma] = T[\psi_\gamma] + V[\psi_\gamma] = \gamma^2 T[\psi] + \gamma V[\psi] \text{ (Coulomb)} \quad (1.9.10)$$

**i. The concept of the virial in classical mechanics**

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

$$virial = \langle \mathbf{q} \cdot \mathbf{F} \rangle = \frac{1}{\tau} \int_{t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} \mathbf{q}(t') \cdot \mathbf{F}(t') dt'$$

It can be shown, that for bound systems:[5]

$$virial = \langle \mathbf{q} \cdot \mathbf{F} \rangle = -2\langle T \rangle$$

For conservative systems the force is a gradient of the potential,  $\mathbf{F} = -\nabla v(\mathbf{q})$ .

The virial relates to dilation of the coordinates through:

$$\frac{d}{d\gamma} v_\gamma(\mathbf{q}) = \mathbf{q} \cdot \nabla v(\gamma\mathbf{q}) = \mathbf{q} \cdot \nabla v_\gamma(\mathbf{q}). \quad (1.9.11)$$

For homogeneous potentials we have:  $\gamma \frac{d}{d\gamma} v_\gamma(\mathbf{q}) = n\gamma^n v(\mathbf{q}) = n v_\gamma(\mathbf{q})$ , thus:

$$n v_\gamma(\mathbf{q}) = \gamma \mathbf{q} \cdot \nabla v_\gamma(\mathbf{q}) \quad (v \text{ homogeneous order } n) \quad (1.9.12)$$

In particular, for  $\gamma = 1$ :

$$n v(\mathbf{q}) = \mathbf{q} \cdot \nabla v(\mathbf{q}) \quad (v \text{ homogeneous order } n) \quad (1.9.13)$$

We will especially be interested in Coulomb systems, where  $n = -1$ , then:

$$v(\mathbf{q}) + \mathbf{q} \cdot \nabla v(\mathbf{q}) = 0 \quad (v \text{ homogeneous order } n = -1) \quad (1.9.14)$$

#### Exercise I-8

Eq. (1.9.13) is known as Euler's equation, after its discoverer. In Thermodynamics it is extremely useful. Thermodynamics stipulates that if you know the energy of a system  $U(S, V, \vec{N})$  as a function of its macroscopic parameters  $S, V, \vec{N}$  then you have complete thermodynamics knowledge. The first and second laws of thermodynamics state that:

$$dU = TdS - pdV + \vec{\mu} \cdot d\vec{N}, \quad (1.9.15)$$

where  $T, p$  and  $\vec{\mu}$  are respectively the temperature, the pressure and the chemical potentials. A second stipulation is that  $U$  is a homogeneous function of order 1.

From this, show that for all thermodynamical systems

1.  $U = TS - PV + \vec{\mu} \cdot \vec{N}$

$$2. \left(\frac{\partial p}{\partial T}\right)_{\bar{\mu}} = \frac{s}{v}.$$

For homogeneous potentials the virial theorem becomes:

$$virial = -n\langle v \rangle = -2\langle T \rangle \quad (1.9.16)$$

For Coulomb systems  $n = -1$ :

$$\langle v \rangle = -2\langle T \rangle \quad (1.9.17)$$

From which:

$$E = \langle v \rangle + \langle T \rangle = -\langle T \rangle = \frac{\langle v \rangle}{2} \quad (1.9.18)$$

We now show that this relation holds also in quantum mechanics.

## ii. The Virial Theorem in quantum mechanics

Now, if  $\psi$  is the *ground-state* energy then  $E[\psi_{\gamma=1}]$  obtains the minimum and therefore:

$$\frac{d}{d\gamma} E[\psi_{\gamma}] = 0 \quad (1.9.19)$$

Plugging this into (1.9.10), one obtains:

$$0 = 2\gamma T[\psi] - n\gamma^{-n-1}V[\psi] \quad (1.9.20)$$

Since we *know* that the optimal value of  $\gamma$  is 1, then (dropping the  $[\psi]$  symbol and remembering that all following quantities are ground state expectation value:

$$\begin{aligned} 2T &= nV \\ E = T + V &= \left(1 + \frac{2}{n}\right)T = \left(\frac{n}{2} + 1\right)V \end{aligned} \quad (1.9.21)$$

These relations show that the *virial theorem* (Eq. (1.9.18)) holds in quantum mechanics provided  $\psi$  is the full molecular eigenstate. For molecules  $n = -1$ : and one has:

$$2T = -V \quad (\text{Coulomb})$$

$$E = T + V = -T = \frac{1}{2}V \quad (1.9.22)$$

A subtle 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 homogenous, it cannot tolerate addition of a constant ( $x^2$  is homogeneous of order 2 but  $x^2 + a$  is not).

### iii. Slater's theory of the chemical bond using the virial theorem

What happens when  $\psi$  is the electronic eigenstate in the Born Oppenheimer approximation? If we look only at the electronic wave function we do not expect Eq. (1.9.21) to be valid. Indeed, using  $\hat{H}_e = \hat{T} + V(\mathbf{r}; \mathbf{R})$  where  $\hat{T} = \left(-\frac{1}{2}\nabla_r^2\right)$  and  $V(\mathbf{r}; \mathbf{R}) = \frac{-Z}{|\mathbf{r}-\mathbf{R}|}$  (with obvious summation on nuclei and electrons)

$$\hat{H}_e[\mathbf{R}]\psi[\mathbf{R}](\mathbf{r}) = E[\mathbf{R}]\psi[\mathbf{R}](\mathbf{r}) \quad (1.9.23)$$

Upon dilation  $\psi_\gamma[\mathbf{R}](\mathbf{r}) = \gamma^{3N_e/2}\psi[\mathbf{R}](\gamma\mathbf{r})$  and note that  $\mathbf{R}$  is *not* dilated. Then we have for the KE and potential:

$$\begin{aligned} T_\gamma[\mathbf{R}] &\equiv \langle \psi_\gamma | \hat{T} | \psi_\gamma \rangle = \int \psi_\gamma[\mathbf{R}](\mathbf{r}) \left(-\frac{1}{2}\nabla_r^2\right) \psi_\gamma[\mathbf{R}](\mathbf{r}) d^{3N_e r} \\ &= \gamma^{3N_e} \int \psi[\mathbf{R}](\gamma\mathbf{r}) \left(-\frac{1}{2}\nabla_r^2\right) \psi[\mathbf{R}](\gamma\mathbf{r}) d^{3N_e r} \\ &= \gamma^2 \int \psi[\mathbf{R}](\mathbf{s}) \left(-\frac{1}{2}\nabla_s^2\right) \psi[\mathbf{R}](\mathbf{s}) d^{3N_e s} = \gamma^2 T[\mathbf{R}] \end{aligned} \quad (1.9.24)$$

And we define:



$$\begin{aligned}
W_\gamma[\mathbf{R}, \mathbf{R}'] &\equiv \langle \psi_\gamma[\mathbf{R}] | \hat{V}[\mathbf{R}'] | \psi_\gamma[\mathbf{R}] \rangle = \int \psi_\gamma(\mathbf{r}, \mathbf{R})^2 V(\mathbf{r}; \mathbf{R}') d^{3N_e} r \\
&= \gamma^{3N_e} \int \psi(\gamma \mathbf{r}; \mathbf{R})^2 V(\mathbf{r}, \mathbf{R}') d^{3N_e} r \\
&= \int \psi(\mathbf{s}, \mathbf{R})^2 V(\gamma^{-1} \mathbf{s}; \mathbf{R}') d^{3N_e} s \\
&= \gamma \int \psi(\mathbf{s}, \mathbf{R})^2 V(\mathbf{s}; \gamma \mathbf{R}') d^{3N_e} s = \gamma W[\mathbf{R}, \gamma \mathbf{R}']
\end{aligned} \tag{1.9.25}$$

We have defined  $W$ , a two  $\mathbf{R}$  quantity. Of course, the physical interaction potential is  $V[\mathbf{R}] = W[\mathbf{R}, \mathbf{R}]$ . The reason we defined  $W$  this way is that we can now write:

$$E_\gamma[\mathbf{R}] = \gamma^2 T_e[\mathbf{R}] + \gamma W[\mathbf{R}, \gamma \mathbf{R}] \tag{1.9.26}$$

Note that we inserted the subscript  $e$  to the KE since now it is only the KE of the electrons.

One still has the variational principle, i.e.  $dE_\gamma[\mathbf{R}]/d\gamma|_{\gamma^*=1} = 0$  (Eq. (1.9.19)):

$$0 = 2\gamma T_e[\mathbf{R}] + W[\mathbf{R}, \gamma \mathbf{R}] + \gamma(\mathbf{R} \cdot \nabla_{\mathbf{R}'} W)[\mathbf{R}, \mathbf{R}']_{\mathbf{R}'=\gamma \mathbf{R}} \tag{1.9.27}$$

Here we used the fact that  $\frac{d}{d\gamma} f(\gamma \mathbf{R}) = [\mathbf{R} \cdot \nabla f](\gamma \mathbf{R})$ . Putting in the solution  $\gamma^* = 1$  we find:

$$0 = 2T_e[\mathbf{R}] + V[\mathbf{R}] + \mathbf{R} \cdot \nabla_{\mathbf{R}'} W[\mathbf{R}, \mathbf{R}']_{\mathbf{R}=\mathbf{R}'} \tag{1.9.28}$$

From the Hellmann-Feynman theorem (see section XXX)  $\nabla_{\mathbf{R}'} (W[\mathbf{R}, \mathbf{R}'])_{\mathbf{R}'=\mathbf{R}} = \nabla_{\mathbf{R}} E(\mathbf{R})$ . Adding to the electronic energy  $E$  the nuclear potential energy  $E_N = \frac{1}{2} \sum_{I \neq I'} \frac{Z_I Z_{I'}}{|\mathbf{R}_I - \mathbf{R}_{I'}|}$  gives the BO energy  $E_{BO} = E + E_N$  and since the nuclear energy is homogeneous of order -1, we can use Eq. (1.9.14) and:

$$0 = E_N + \sum_J \mathbf{R}_J \cdot \nabla_{\mathbf{R}_J} E_N, \tag{1.9.29}$$

Plugging all this into eq. (1.9.28) we find Slater's formula[6]:

$$T_e = -E_{BO} - \mathbf{R} \cdot \nabla_{\mathbf{R}} E_{BO} \quad (1.9.30)$$

From this, using  $E_{BO}(\mathbf{R}) = T_e + V_{tot}$  (where  $V_{tot} = V + E_N$  is the total PE of the molecule) we find also:

$$V_{tot} = 2E_{BO} + \mathbf{R} \cdot \nabla_{\mathbf{R}} E_{BO} \quad (1.9.31)$$

We point out an interesting property of Eq. (1.9.30) is that we can obtain electronic properties such as kinetic energy  $T_e$  or potential energy  $V$  directly from the potential surface  $E_{BO}$ . This will be important for analyzing properties of DFT quantities. Note that at stationary points on the Born Oppenheimer PES where  $\nabla_{\mathbf{R}} E_{BO} = 0$  the usual virial relation eq. (1.9.21) holds (with neglect of nuclear kinetic energies).

Slater derived Eq. (1.9.30) in a different manner, not from variational or dilation considerations. He used it for analyzing the general nature of chemical bonding of a diatomic molecule. We follow in his footsteps here. Consider a diatomic molecule. Suppose the BO curve is give as a potential of the following form, in terms of the inter-nuclear distance  $R$ :

$$E_{BO}(R) = \frac{A}{R^m} - \frac{B}{R^n} + E_{\infty} \quad (1.9.32)$$

Where  $A$  and  $B$  are *positive* constants and typically  $m \gg n$ . For example, in the case the PE curve is given by the Lennard-Jones potential then  $m = 12$  and  $n = 6$ . This potential describes long-range attraction and short-range repulsion. It has a minimum at:

$$R_{Min} = \left( \frac{Bn}{Am} \right)^{\frac{1}{m+n}} \quad (1.9.33)$$

One then has:

$$T_e = \frac{B(1-n)}{R^n} - \frac{A(1-m)}{R^m} - E_\infty \quad (1.9.34)$$

When  $R \rightarrow \infty$  we have:  $T_e \approx -\frac{B(n-1)}{R^n} - E_\infty$ . Let us consider what happens to the KE as the distance  $R$  is *reduced*. Clearly it initially descends, however, at some distance  $R_T$  it reaches a minimum, and below this distance it begins to increase. Since:

$$\frac{d}{dR} T_e = -\frac{Bn(1-n)}{R^{n+1}} + \frac{Am(1-m)}{R^{m+1}} \quad (1.9.35)$$

We find the minimum kinetic energy is obtained at a distance somewhat larger than  $R_{Min}$ :

$$R_T = \left( \frac{Am(m-1)}{Bn(n-1)} \right)^{\frac{1}{m-n}} = \left( \frac{m-1}{n-1} \right)^{\frac{1}{m-n}} R_{Min} > R_{Min} \quad (1.9.36)$$

For example, for  $m = 12$ ,  $n = 6$  we find:  $R_T = 1.14R_{Min}$ . The sum of electronic and nuclear potential energy is:

$$V_{tot} = \frac{A(2-m)}{R^m} - \frac{B(2-n)}{R^n} + 2E_\infty \quad (1.9.37)$$

At large  $R$ ,  $V + E_N$  actually increases as the atoms approach. The potential energy does not rise indefinitely. At some inter-nuclear distance, larger than the distance at which the BO potential obtains its minimum, it reaches a maximum and then starts decreasing. We have:

$$\frac{Am(2-m)}{R^{m+1}} = \frac{Bn(2-n)}{R^{n+1}} \quad (1.9.38)$$

$$R_V = \left( \frac{Am(m-2)}{Bn(n-2)} \right)^{\frac{1}{m-n}} = \left( \frac{m-2}{n-2} \right)^{\frac{1}{m-n}} R_{Min} > R_{Min}$$

For the Lennard Jones potential we have:  $R_V = 1.16R_{Min}$ . For  $R < R_T$  we find that the kinetic energy grows sharply as  $R$  decreases, while the potential

energy decreases. This sharp behavior can be interpreted as a condensation of the electron cloud around each nuclei as such a condensation causes increase in kinetic energy  $T$  and decrease in electronic potential energy  $V$ . The total BO energy continues to drop as long as  $R > R_{Min}$  and then starts rising.

As an application of this theory consider the  $X \dots Y^+$  system where  $X$  is a neutral atom and  $Y^+$  a distant atomic ion (can also be a molecular ion). The energy at large distance  $R$  is that of Eq. (1.9.32) with  $A = 0$  and  $B = \frac{\alpha(X)}{2}$  where  $\alpha(X)$  is the polarizability of  $X$  and  $n = 4$ . This form of the potential surface results from a simple exercise in electrostatics. Quantum mechanics is needed only for calculating the value of the polarizability of  $X$ ,  $\alpha(X)$ . Slater's analysis states that:

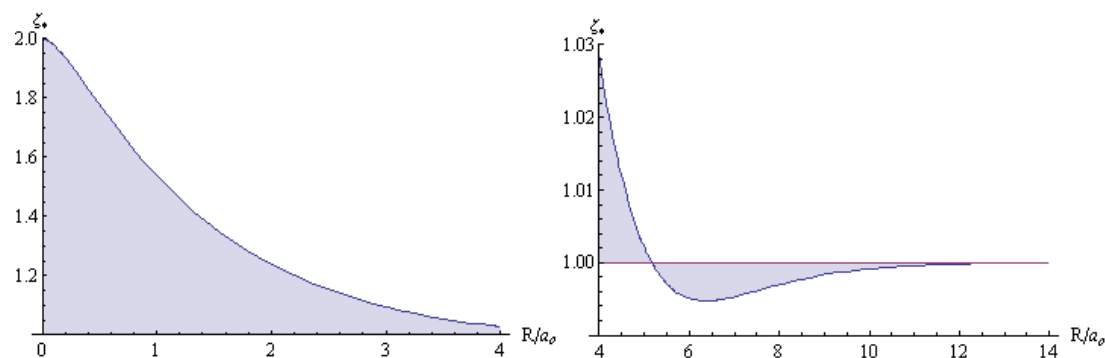
$$T_e = -\frac{3\alpha(X)}{2R^4} - E_\infty$$

$$V_{tot} = \frac{\alpha(X)}{R^4} + 2E_\infty$$
(1.9.39)

So the force due to the total potential energy is *repulsive* while that due to the KE is *attractive*. The KE attraction outweighs the PE repulsion by a factor of three halves. Slater's analysis shows that the interaction here is fundamentally different from that of electrostatics, where the force attributed entirely to the Coulombic attraction.

Another application is to the simplest chemical bond, that of the  $H_2^+$  molecule. The text-book analysis of this molecule assumes that the molecular orbital is a "bonding" linear combination of the two atomic orbitals (LCAO), both having the 1s shape:  $\psi_{1s}(r) = e^{-\frac{r}{a_0}}$ , but each centered on its corresponding nucleus. Quantitatively this picture is very inaccurate. While it does show binding, its predicted a bond length is almost 1.4 Å (while the true bond length is 1.06Å)

and its estimate of atomization energy is 1.7 eV while the experimental value is close to 2.7eV. There are no electron correlation effects here, so the only source of inaccuracy is the limited efficacy of the small basis.



**Figure I-3: The optimal dilation parameter  $\zeta_*$  as a function of the inter-nuclear distance  $R$  in  $H_2^+$ , where the wave function is the sum of 2 1s-type orbitals (localized on each of the two nuclei).**

One can add a dilation factor to the wave function  $\psi_{1s}^\zeta(r) = e^{-\frac{\zeta r}{a_0}}$  and perform a variational determination of  $\zeta$ . Then, one can obtain, for each value of the inter-nuclear distance the optimal  $\zeta$ , denoted  $\zeta_*(R)$ . At  $R \rightarrow \infty$  the dilation parameter should be equal to 1, as then the 1s orbital is exact. At  $R \rightarrow 0$  the molecule becomes the He ion and  $\zeta_*$  should approach the exact value of 2. See Figure I-3. But the optimal value  $\zeta_*$  is not monotonic! At large  $R$  it becomes smaller than 1, indicating that the electron cloud *expands*, the kinetic energy drops and the potential energy rises. Thus long-range attraction is driven by the kinetic energy. At  $R = 5.2a_0$  the value of  $\zeta_*$  is again 1 and then grows monotonically as  $R$  is reduced. Now the orbital is *contracting*, kinetic energy is growing and potential energy is decreasing. In this range, chemical binding is driven by the potential energies. At the distance minimizing the BO energy the value of  $\zeta$  is:  $\zeta_* \approx 1.23$ , i.e. the orbitals are then considerably contracted (relative to the free H atom). With such contracted orbitals, the calculated bond length of the LCAO calculation is  $1.06\text{\AA}$ . The bond energy, computed by

subtracting the calculated  $H_2^+$  energy from that of the hydrogen atom gives the bond energy of 2.35eV. The first to report this type of calculation is in ref [7], a clear exposition of the calculation is given in refs [8]. These improved theoretical estimates, with a two-function basis, achieved with just a single dilation parameter, show that orbital shapes are crucial for a good description of chemical bonding.

This result contradicts our intuition regarding the behavior of electrons in a covalent bond. It seems that a major source of the chemical bond energy in  $H_2^+$  is not due to the electrostatic benefit resulting from the placement of electrons in the *middle region* separating the two nuclei. Rather, it is due to the *contraction* of the electron wave function near each of the nuclei induced by change in Coulomb potential near that nucleus due to the presence of the other nuclei. Around one nucleus of  $H_2^+$  the nuclear Coulombic potential becomes deeper and distorted towards the second nucleus. As a result, a contraction of the electron density around the nucleus is obtained, and the charge placed between the atoms is not in the “midway” region of the bond: it is between them alright but very close to each nucleus. Each nuclei seems to share the electron not by placing it *midway* between them, but rather by having the electrons *much closer* to each nucleus.

Combining this with the above result, we find a very general conclusion:

Rule 1 of chemical attraction: As atoms approach from afar, (for  $R > R_T$ ) the kinetic energy decreases and the potential energy increases and thus it is the *lowering of kinetic energy* which is responsible for the distant attraction between molecules.

Rule 2 of chemical bonding: For distances near but larger than the bonding distance the attraction is more subtle. The electronic kinetic energy rises

sharply and the potential energy drops even more sharply. Thus, a major source of energy is the contraction of orbitals around their respective nuclei, inspired by the deepening of potential wells due to the presence of close-by nuclei. The electronic PE is thus the major glue when atoms are close to each other: It has to offset the nuclear repulsion *and* the repulsion due to kinetic energy of electrons.

#### iv. Other Dilation Facts

Consider the Schrödinger equation:

$$-\frac{\hbar^2}{2m}\psi''(x) + v(x)\psi(x) = E\psi(x) \quad (1.9.40)$$

We now dilate: we assume  $\psi_\lambda(x) = \psi(\lambda x)$  is an eigenstate and find the corresponding Hamiltonian. Note that  $\psi_\lambda''(x) = \lambda^2\psi''(\lambda x)$  thus:

$$-\frac{\hbar^2}{2m\lambda^2}\psi_\lambda''(x) + v(\lambda x)\psi_\lambda(x) = E\psi_\lambda(x) \quad (1.9.41)$$

We see that the required Hamiltonian involves a dilated potential and a *scaled* mass by a factor  $\lambda^2$ . The energy is left intact. Alternatively, we can multiply through by the square of the dilation factor  $\lambda^2$  and obtain:

$$-\frac{\hbar^2}{2m}\psi_\lambda''(x) + \lambda^2v(\lambda x)\psi_\lambda(x) = \lambda^2E\psi_\lambda(x) \quad (1.9.42)$$

Now the potential is dilated *and* scaled:  $v(x) \rightarrow \lambda^2v(\lambda x)$  when then the wave function is dilated by  $\psi(x) \rightarrow \psi(\lambda x)$  and the total energy is also scaled by  $E \rightarrow \lambda^2E$ . This is general for basically any Schrödinger equations and holds for any number of spatial dimensions.

Now consider homogeneous potentials in 3D:  $v(\lambda\mathbf{r}) = \lambda^n v(\mathbf{r})$ . In this case the potential transform is indeed a multiplicative scaling:  $v(\mathbf{r}) \rightarrow \lambda^{2+n}v(\mathbf{r})$ . In

particular, in Coulomb systems, when  $n = -1$ , we find that the potential scaling is simply by a factor  $\lambda$ . We can turn around the discussion and state:

If the Coulomb coupling constant is scaled by a factor  $\lambda$  then the total energy is scaled by a factor  $\lambda^2$  and the wave function is dilated by a factor  $\lambda$ . In particular, the density is dilated and scaled as  $n(\mathbf{r}) \rightarrow \lambda^3 n(\lambda \mathbf{r})$ .

According to the virial theorem, the potential energy and the total energy are related by  $\frac{V}{2} = -T = E$ , thus the expectation value of the potential and kinetic energies is scaled by a factor  $\lambda^2$  as well.

#### Exercise I-9

An application of the above discussion allows an interesting exact conclusion concerning the ground-state energy per particle of the homogeneous electron gas of density  $n$  (see section XXX for details). Denote this quantity,  $\epsilon^{HEG}(n, e^2)$  and the Coulomb coupling constant  $e^2/4\pi\epsilon_0$ . Show that:

$$3n \frac{\partial \epsilon^{HEG}}{\partial n} + e^2 \frac{\partial \epsilon^{HEG}}{\partial e^2} = 2\epsilon^{HEG} \quad (1.9.43)$$

Hint: Start from  $\epsilon^{HEG}(\lambda^3 n, \lambda e^2) = \lambda^2 \epsilon^{HEG}(n, e^2)$ . Then take the derivative with respect to  $\lambda$ . Then set  $\lambda = 1$ .

## I. Introduction to Functionals

### i. Definition of a functional and some examples

A formula that accepts a function and maps it to a *number* is called a *functional*. We denote the number a functional  $F$  maps to a function  $n(\mathbf{r})$  by this by the symbol  $F[n]$ . Take for example the simplest functional:  $F[n] = 0$ . This functional maps any function  $n(\mathbf{r})$  to the number zero. A more interesting functional is:



$$F_{\mathbf{r}_0}[n] = n(\mathbf{r}_0) \quad (1.10.1)$$

This functional maps to each function  $n(\mathbf{r})$  its value at a certain point  $\mathbf{r}_0$ .

Next, consider the functional

$$I_V[n] = \frac{1}{V} \int_V n(\mathbf{r}) d^3r, \quad (1.10.2)$$

mapping each function its average value in a given volume  $V$ .

Another familiar example is the kinetic energy functional for a particle of mass  $\mu$  in some wave function  $\psi(\mathbf{r})$ :

$$T[\psi] = -\frac{\hbar^2}{2\mu} \int_V \psi^*(\mathbf{r}) \nabla^2 \psi(\mathbf{r}) d^3r, \quad (1.10.3)$$

Or the energy for a given potential  $V(\mathbf{r})$ :

$$E[\psi] = T[\psi] + V[\psi], \quad (1.10.4)$$

Where:

$$V[\psi] = \int V(\mathbf{r}) |\psi(\mathbf{r})|^2 d^3r, \quad (1.10.5)$$

## ii. Linear functionals

A functional is called linear if for any two functions  $n(\mathbf{r})$  and  $m(\mathbf{r})$  one has:  $F[n(\mathbf{r}) + m(\mathbf{r})] = F[n(\mathbf{r})] + F[m(\mathbf{r})]$  and if  $F[\alpha n(\mathbf{r})] = \alpha F[n(\mathbf{r})]$  for any number  $\alpha$ . The functionals above  $F_{\mathbf{r}_0}[n]$  and  $I_V[n]$  are examples of such functionals, please check. Any linear functional can be written in the form:

$$F[n] = \int f_L(\mathbf{r}) n(\mathbf{r}) d^3r. \quad (1.10.6)$$

The function  $f_L(\mathbf{r})$  is called the linear response kernel of  $F$ . For example, the linear response kernel of  $F_{\mathbf{r}_0}$  is  $f_L(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)$ . And that of  $I_V$  is  $f_L(\mathbf{r}) = \frac{1}{V}$ .

### iii. Functional Derivatives

Let us now introduce the concept of a "functional derivative". Consider the functional  $F_{r_0}[n]$  of Eq. (1.10.1). A functional derivative, denoted  $\frac{\delta F}{\delta n(\bar{\mathbf{r}})}$  measures how  $F[n]$  changes when we make a small localized change of  $n(\mathbf{r})$  at  $\bar{\mathbf{r}}$ .

Let us give a more detailed definition. Suppose we change the argument function in  $F[n]$  by making an *arbitrarily* small change in  $(\mathbf{r}) \rightarrow n(\mathbf{r}) + \epsilon\Phi(\mathbf{r})$ . We use an arbitrary function  $\Phi(\mathbf{r})$  multiplied by an infinitesimal number  $\epsilon$  (that is, a number  $\epsilon$  we intend to make vanishingly small at a later stage). In principle, for any finite  $\epsilon$  the change  $\frac{\delta F}{\delta n(\bar{\mathbf{r}})}$  in  $F$  has a very complicated dependence on  $\epsilon$ . But, since  $\epsilon$  is assumed small, we can Taylor expand with respect to  $\epsilon\Phi$  and then "throw out" all terms beyond the term linear in  $\epsilon\Phi$ . This makes sense since we assume that  $\epsilon$  is as small as we want. Thus, the linear part of  $\delta F$  is a functional of  $\epsilon\Phi$  and it is a linear functional. We denote the linear response kernel of this functional by the symbol  $\frac{\delta F}{\delta n(\bar{\mathbf{r}})}$ , called "the functional derivative". Therefore we can write:

$$\delta F \equiv F[n + \epsilon\Phi] - F[n] = \int \frac{\delta F}{\delta n(\bar{\mathbf{r}})} \epsilon\Phi(\bar{\mathbf{r}}) d^3\bar{\mathbf{r}} + O(\epsilon^2) \quad (1.10.7)$$

Note, that this the linear response kernel  $\frac{\delta F}{\delta n(\bar{\mathbf{r}})}$  is independent of  $\Phi(\mathbf{r})$  and  $\epsilon$ . One technique to obtain the kernel is taking  $\Phi(\mathbf{r}) = \delta_{\bar{\mathbf{r}}}(\mathbf{r}) = \delta(\mathbf{r} - \bar{\mathbf{r}})$ . To make this notion precise, we define:

$$\frac{\delta F}{\delta n(\bar{\mathbf{r}})} \equiv \lim_{\epsilon \rightarrow 0} \frac{F[n + \epsilon\delta_{\bar{\mathbf{r}}}] - F[n]}{\epsilon} \quad (1.10.8)$$

Where  $\delta_{\bar{\mathbf{r}}} = \delta(\mathbf{r} - \bar{\mathbf{r}})$  is a delta function, expressing the fact that we make a localized change in  $n(\mathbf{r})$  at  $\bar{\mathbf{r}}$ .

### Examples

Apply definition (1.10.8) to  $F_{r_0}[n]$ , and  $I_V[n]$ . Find the functional derivatives and check that Eq. (1.10.7) holds.

Solution:

$$\begin{aligned}\frac{\delta F_{r_0}}{\delta n(\bar{\mathbf{r}})} &\equiv \lim_{\epsilon \rightarrow 0} \frac{F_{r_0}[n + \epsilon \delta_{\bar{\mathbf{r}}}] - F_{r_0}[n]}{\epsilon} \equiv \lim_{\epsilon \rightarrow 0} \frac{(n(\mathbf{r}_0) + \epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r}_0)) - n(\mathbf{r}_0)}{\epsilon} \\ &= \delta_{\bar{\mathbf{r}}}(\mathbf{r}_0) = \delta(\mathbf{r}_0 - \bar{\mathbf{r}})\end{aligned}\quad (1.10.9)$$

The functional derivative comes out  $\delta_{\bar{\mathbf{r}}}(\mathbf{r}) = \delta(\mathbf{r} - \bar{\mathbf{r}})$ , meaning that any change in  $n(\mathbf{r})$  made at  $\bar{\mathbf{r}}$  will not affect  $F_{r_0}[n]$ , unless it is made at  $\bar{\mathbf{r}} = \mathbf{r}_0$ . Indeed, if we make an arbitrary small change  $n \rightarrow n + \epsilon \Phi$ , we have from (1.10.7):

$$\begin{aligned}\delta F_{r_0} &\equiv F_{r_0}[n + \epsilon \delta_{\bar{\mathbf{r}}}] - F_{r_0}[n] = (n(\mathbf{r}_0) + \epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r}_0)) - n(\mathbf{r}_0) = \epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r}_0) \\ &= \epsilon \delta(\mathbf{r}_0 - \bar{\mathbf{r}})\end{aligned}\quad (1.10.10)$$

And this is seen to be indeed the case directly from the definition of  $F_{r_0}$  in Eq. (1.10.1).

As for  $I_V[n] = \frac{1}{V} \int_V n(\mathbf{r}) d^3r$ :

$$\begin{aligned}\frac{\delta I_V}{\delta n(\bar{\mathbf{r}})} &\equiv \lim_{\epsilon \rightarrow 0} \frac{I_V[n + \epsilon \delta_{\bar{\mathbf{r}}}] - I_V[n]}{\epsilon} \\ &\equiv \lim_{\epsilon \rightarrow 0} \frac{V^{-1} \int_V (n(\mathbf{r}) + \epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r})) d^3r - V^{-1} \int_V n(\mathbf{r}) d^3r}{\epsilon} \\ &= V^{-1} \int \delta_{\bar{\mathbf{r}}}(\mathbf{r}) d^3r = V^{-1}\end{aligned}\quad (1.10.11)$$

### Another Example

An important functional, we will use is the "Hartree energy", the classical repulsion energy associated with a charge distribution  $n(\mathbf{r})$ :

$$E_H[n] = \frac{1}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (1.10.12)$$

Prove that the functional derivative of the Hartree Energy, called the Hartree potential is the electrostatic potential of that density of electrons:

$$\begin{aligned}
 v_H(\bar{\mathbf{r}}) &\equiv \frac{\delta E_H}{\delta n(\bar{\mathbf{r}})} \\
 &= \lim_{\epsilon \rightarrow 0} \frac{\frac{1}{2} \iint \left( \frac{(n(\mathbf{r}) + \epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r}))(n(\mathbf{r}') + \epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} - \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right) d^3r d^3r'}{\epsilon} \\
 &= \lim_{\epsilon \rightarrow 0} \frac{\frac{1}{2} \iint \frac{(\epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r})n(\mathbf{r}') + n(\mathbf{r})\epsilon \delta_{\bar{\mathbf{r}}}(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r'}{\epsilon} \\
 &= \int \frac{n(\mathbf{r}')}{|\bar{\mathbf{r}} - \mathbf{r}'|} d^3r'
 \end{aligned} \tag{1.10.13}$$

Thus, we find as requested that  $v_H(\mathbf{r})$  is indeed the electrostatic potential of the density  $n(\mathbf{r})$ .

Two shortcuts: when “functional integrating”, one can often use regular differentiation rules. All one needs to remember is the following shortcuts:

$$\frac{\delta n(x)}{\delta n(x')} \rightarrow \delta(x - x') \quad \frac{\delta n'(x)}{\delta n(x')} \rightarrow \delta'(x - x') \tag{1.10.14}$$

And use them within “chain rules”:  $\frac{\delta f(n(x))}{\delta n(x')} \rightarrow f'(n(x))\delta(x - x')$  (where  $f'(n) \equiv \frac{d}{dn}f(n)$ ). An example of the use of this rule is the previous exercise:

$$\frac{\delta I_V}{\delta n(\bar{\mathbf{r}})} \equiv V^{-1} \int \frac{\delta n(\mathbf{r})}{\delta n(\bar{\mathbf{r}})} d^3r = V^{-1} \int \delta(\mathbf{r} - \bar{\mathbf{r}}) d^3r = V^{-1} \tag{1.10.15}$$

Another example is the following functional, related to the kinetic energy:

$$J[f] = \int_{-\infty}^{\infty} f'(x)^2 dx \tag{1.10.16}$$

The functional derivative is:

$$\begin{aligned}
\frac{\delta J[f]}{\delta f(\bar{x})} &= \int_{-\infty}^{\infty} \frac{\delta(f'(x)^2)}{\delta f(\bar{x})} dx = \int_{-\infty}^{\infty} 2f'(x) \frac{\delta f'(x)}{\delta f(\bar{x})} dx \\
&= \int_{-\infty}^{\infty} 2f'(x) \delta'(x - \bar{x}) dx = -2f''(\bar{x})
\end{aligned}
\tag{1.10.17}$$

Here we used the identity:  $\int g(x) \delta'(x - \bar{x}) dx = -g'(\bar{x})$  and the shortcut  $\frac{\delta f'(x)}{\delta f(\bar{x})} = \delta'(x - \bar{x})$ .

#### iv. Invertible function functionals and their functional derivatives

Consider a “continuous set of functionals” which is “indexed” by a continuous variable  $\mathbf{r}$  in some domain  $D$ . More precisely,  $v[n](\mathbf{r})$  assigns a functional to each  $\mathbf{r}$  in  $D$ . Let us suppose further that  $n$  is a function of  $\mathbf{r}$  defined on the same domain  $D$ . An example is the relation between the electrostatic potential and the charge density:

$$v_H[n](\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \tag{1.10.18}$$

For any charge distribution  $n(\mathbf{r}')$  the integral forms a the electrostatic potential  $v_H(\mathbf{r})$ . The potential at a give point  $\mathbf{r}$  depends in principle on the charge density at all points in space. This is a very non-local dependence.

What about the functional derivative? If one makes a small change in the function  $n$  at the point  $\mathbf{r}'$ ,  $\delta n(\mathbf{r}')$ , how will that affect the functional at  $\mathbf{r}$ ? We write this as a definition of the functional derivative:

$$\delta v[n](\mathbf{r}) = \int \frac{\delta v[n](\mathbf{r})}{\delta n(\mathbf{r}')} \delta n(\mathbf{r}') d^3r' \tag{1.10.19}$$

In the example above, it is immediately noticeable that:

$$\frac{\delta v_H[n](\mathbf{r})}{\delta n(\mathbf{r}')} = \frac{1}{|\mathbf{r} - \mathbf{r}'|} \tag{1.10.20}$$

An interesting concept can arise when one knows in advance that the relation  $v[n]$  is unique, i.e.

$$\{v[n](\mathbf{r}) = v[n'](\mathbf{r}) \text{ for all } \mathbf{r}\} \Leftrightarrow \{n(\mathbf{r}) = n'(\mathbf{r}) \text{ for all } \mathbf{r}\} \quad (1.10.21)$$

In this case one can also consider  $n$  to be a functional of  $v$ . Indeed:

$$\delta n[v](\mathbf{r}) = \int \frac{\delta n[v](\mathbf{r})}{\delta v(\mathbf{r}')} \delta v(\mathbf{r}') d^3 r' \quad (1.10.22)$$

Once can combine this equation with (1.10.20) and see that:

$$\delta n(\mathbf{r}) = \int \frac{\delta n[v](\mathbf{r})}{\delta v(\mathbf{r}')} \int \frac{\delta v[n](\mathbf{r}')}{\delta n(\mathbf{r}'')} \delta n(\mathbf{r}'') d^3 r'' d^3 r' \quad (1.10.23)$$

This shows that:

$$\int \frac{\delta n(\mathbf{r})}{\delta v(\mathbf{r}')} \frac{\delta v(\mathbf{r}')}{\delta n(\mathbf{r}'')} d^3 r' = \delta(\mathbf{r} - \mathbf{r}'') \quad (1.10.24)$$

Thus, there is an inverse relation between the functional derivatives. This is similar to what happens in usual derivatives with inverse functions. When a function  $g(x)$  is invertible we can speak of  $x(g)$ . The change in  $g$  due to a change in  $x$  is:  $dg(x) = \frac{dg}{dx} dx$ . Similarly:  $dx = \frac{dx}{dg} dg$ . Therefore:  $\left(\frac{dg}{dx}\right)_x = \left(\left(\frac{dx}{dg}\right)^{-1}\right)_{g(x)}$ .

## v. Convex functionals

A function  $f(x)$  is said to be convex if for any pair of abscisas  $x_1$  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  $(x_1, f(x_1))$  and  $(x_2, f(x_2))$ . A more formal definition is that for any  $0 < \lambda < 1$  we have:

$$f(\lambda x_2 + (1 - \lambda)x_1) \leq \lambda f(x_2) + (1 - \lambda)f(x_1) \quad (1.10.25)$$

A useful interpretation of the above definition for convexity is in terms of averages. We can view  $\lambda x_1 + (1 - \lambda)x_2$  as a weighted average of the two points, with positive weights (such an average is usually called a convex average). In this sense we have convexity obey:

$$f(\langle x \rangle) \leq \langle f(x) \rangle \quad (1.10.26)$$

This relation is much more general than just 2 points. We can easily show that for convex functions it holds for any averaging procedure  $\langle x \rangle = \sum_i c_i x_i$  and  $\langle f(x) \rangle = \sum_i c_i f(x_i)$  with  $c_i \geq 0$  and  $\sum_i c_i = 1$ . Eq. (1.10.26) is sometimes called Jensen's inequality.

One of the important implications 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. If a function is known to be convex and to have a minimum then any "descent" algorithm is surely to find the (global) minimum. This is useful.

We can Taylor expand  $f(\lambda x_2 + (1 - \lambda)x_1) = f(\lambda(x_2 - x_1) + x_1)$  with respect to  $\lambda$ , assuming now  $\lambda \ll 1$ . Then:

$$\begin{aligned} \lambda f(x_2) + (1 - \lambda)f(x_1) &\geq f(\lambda(x_2 - x_1) + x_1) \\ &= f(x_1) + \nabla f(x_1) \cdot \lambda(x_2 - x_1) + O(\lambda^2) \end{aligned} \quad (1.10.27)$$

After rearrangement and division by  $\lambda$  we set  $\lambda$  to zero and obtain:

$$f(x_2) - f(x_1) \geq \nabla f(x_1) \cdot (x_2 - x_1) \quad (1.10.28)$$

This relation, if obeyed for all pairs of points, is equivalent to the convexity of the function. By exchanging  $x_1$  and  $x_2$  we obtain also:

$$\nabla f(x_2) \cdot (x_2 - x_1) \geq f(x_2) - f(x_1) \geq \nabla f(x_1) \cdot (x_2 - x_1) \quad (1.10.29)$$

From (1.10.28) we further have, for  $\lambda \rightarrow 0$ :

$$\begin{aligned}
f(x_1) + \lambda \nabla f(x_1) \cdot u &\leq f(x_1 + \lambda u) \\
&= f(x_1) + \lambda \nabla f(x_1) \cdot \delta x_1 + \lambda^2 \frac{1}{2} \delta x_1 \cdot \nabla \nabla f(x_1) \cdot \delta x_1 \quad (1.10.30) \\
&\quad + O(\lambda^3 |u|^3)
\end{aligned}$$

From this, after cancelling and dividing by  $\lambda^2$  we find  $\frac{1}{2} u \cdot \nabla \nabla f(x_1) \cdot u \geq 0$  for arbitrary  $u$  which means the Hessian  $\nabla \nabla f(x_1)$  is positive definite: another characteristic of convex functions

A similar definition of convexity can be applied to functionals

$$F[\lambda n_1 + (1 - \lambda)n_2] \leq \lambda F[n_1] + (1 - \lambda)F[n_2] \quad (1.10.31)$$

And they too have the equivalent differential property, which will be useful below:

$$F[n_1] - F[n_2] \geq \int \left( \frac{\delta F}{\delta n(\mathbf{r})} \right)_{n_2} (n_1(\mathbf{r}) - n_2(\mathbf{r})) d^3 r \quad (1.10.32)$$

An example for a convex functional useful for density functional theory is the so-called Hartree energy functional:

$$E_H[n] = \frac{1}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r' \quad (1.10.33)$$

The second functional derivative of this function is  $|\mathbf{r} - \mathbf{r}'|^{-1}$  which is a manifestly positive matrix (this is easily seen by Fourier transforming the function  $v(\mathbf{r}) = \frac{1}{r}$  with result  $\tilde{v}(\mathbf{q}) = \frac{4\pi}{q^2}$  which is positive). Another example is the von-Weizsacker kinetic energy functional:

$$E_{vW}[n] = -\frac{1}{2} \int \sqrt{n(\mathbf{r})} \nabla^2 \sqrt{n(\mathbf{r})} d^3 r \quad (1.10.34)$$

The physical meaning of this functional is that it gives  $N$  times the kinetic energy of a single-particle with ground state wave function  $\psi(\mathbf{r}) = \sqrt{\frac{n(\mathbf{r})}{N}}$



where  $N = \int n(\mathbf{r})d^3r$ . Again, by showing the Hessian is positive definite we can straightforwardly show this functional is convex First, rewrite it as:

$$E_{vW}[n] = \frac{1}{2} \int \left( \nabla \sqrt{n(\mathbf{r})} \right)^2 d^3r = \frac{1}{8} \int \frac{(\nabla n)^2}{n} d^3r \quad (1.10.35)$$

We drop the explicit  $\mathbf{r}$  in the last expression, for brevity. Then, make an arbitray but small perturbation  $\epsilon\Phi(\mathbf{r})$ :

$$E_{vW}[n + \epsilon\Phi] = \frac{1}{8} \int \frac{(\nabla n + \epsilon\nabla\Phi)^2}{n + \epsilon\Phi} d^3r \quad (1.10.36)$$

And expand the denominator to second order in  $\epsilon$ :  $(n + \epsilon\Phi)^{-1} \rightarrow \frac{1}{n} \left( 1 - \epsilon \frac{\Phi}{n} + \epsilon^2 \left( \frac{\Phi}{n} \right)^2 + \dots \right)$ . Expanding the numerator and performing the multiplication order by order one can then read off the second order contribution and after some manipulation bring it to the form:

$$\epsilon^2 \delta^2 E_{vW}[n + \epsilon\Phi] = \frac{\epsilon^2}{8} \int \frac{\Phi^2}{n} \left( \frac{\nabla n}{n} - \frac{\nabla\Phi}{\Phi} \right)^2 d^3r \quad (1.10.37)$$

Clearly, this second order term is absolutely non-negative for all perturbations and at all physical densities. This shows that the underlying functional derivative Hessian is positive definite and thus the functional is convex.

A similar condition exists for the second functional derivative. The Jensen inequality holds here as well:

$$F[\langle n \rangle] \leq \langle F[n] \rangle \quad (1.10.38)$$

This relation, in combination with the fact that  $f(x) = e^x$  is convex is useful for developing mean field approaches in Statistical mechanics. It was used by Gibbs and later, in a quantum context by Peierls.[9]

***J. Minimizing functionals is an important part of developing a DFT***

Equation Section (Next) **Minimization of functions**

We start with the problem of minimization of a simple 1D function. Given a function  $f(x)$ , 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:

$$f'(x^*) = 0 \quad (1.11.1)$$

Let us expand in a Taylor's series the function around the point  $x^*$ . Clearly, we have:

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 + \dots \quad (1.11.2)$$

However, taking Eq. (1.11.1) into account,

$$f(x) = f(x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 + \dots \quad (1.11.3)$$

When  $x$  is extremely close to  $x^*$  we may neglect high order terms and then we see that Eq. (1.11.3) is the equation of a parabola. In order that  $f(x)$  be a minimum, we must have an ascent of  $f(x)$  when we move away from  $x^*$ , no matter which direction we take (left or right). For small displacements  $x^* \rightarrow x$ , we see from Eq. (1.11.3) that the change in  $f(x)$  is  $\frac{1}{2}f''(x)(x - x^*)^2$ . Since  $(x - x^*)^2$  is always positive, no matter the direction we move, we must demand that  $f''(x^*) > 0$  as well. Thus, our necessary conditions for a minimum are:

$$f'(x^*) = 0 \quad \text{and} \quad f''(x^*) > 0 \quad (1.11.4)$$

Now, let us consider the case of functions of two variables,  $f(\mathbf{r})$  where  $\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix} \equiv (x \ y)^T$ . Notice that we use the "transpose" symbol  $T$  to turn a column vector into a row vector. Let us Taylor expand  $f(\mathbf{r})$  around a point  $\mathbf{r}^*$ :

$$\begin{aligned} f(\mathbf{r}) &= f(\mathbf{r}^*) + \nabla f(\mathbf{r}^*)^T (\mathbf{r} - \mathbf{r}^*) + \frac{1}{2} (\mathbf{r} - \mathbf{r}^*)^T \nabla \nabla f(\mathbf{r}^*) (\mathbf{r} - \mathbf{r}^*) + \dots \\ &= f(\mathbf{r}^*) + \mathbf{g}^{*T} (\mathbf{r} - \mathbf{r}^*) + \frac{1}{2} (\mathbf{r} - \mathbf{r}^*)^T H^* (\mathbf{r} - \mathbf{r}^*) + \dots \end{aligned} \quad (1.11.5)$$

Note that we use the notations:

$$\begin{aligned} \mathbf{g}^* &= \mathbf{g}(\mathbf{r}^*) = \nabla f(\mathbf{r}^*) = \begin{pmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{pmatrix} \\ H^* &= H(\mathbf{r}^*) = \begin{pmatrix} \frac{\delta^2 f}{\delta x^2} & \frac{\delta^2 f}{\delta y \delta x} \\ \frac{\delta^2 f}{\delta x \delta y} & \frac{\delta^2 f}{\delta y^2} \end{pmatrix}_{\mathbf{r}^*} \end{aligned} \quad (1.11.6)$$

The vector  $\mathbf{g}(\mathbf{r})$  is called the gradient and the matrix  $H(\mathbf{r})$  the Hessian at  $\mathbf{r}$ . Note that the Hessian is a symmetric matrix, since the order of differentiation is immaterial. When  $\mathbf{r}^*$  is a minimum, moving infinitesimally away from it in any direction will not change the function. Why? We can show this leads to a contradiction. Let  $\mathbf{d}$  be an arbitrary direction. If the function decreases when moving from  $\mathbf{r}^*$  in that direction, no matter how small the step size, then this contradicts that  $\mathbf{r}^*$  is a minimum. Now, if  $\mathbf{d}$  is an ascent direction, then  $-\mathbf{d}$  will be a descent direction - again a contradiction to  $\mathbf{r}^*$  being a minimum. Thus we conclude, that  $f$  cannot change (to first order) or in other words, the gradient  $\mathbf{g}^*$  of  $f(\mathbf{r})$  at  $\mathbf{r}^*$  must be zero. Then Eq. (1.11.5) becomes:

$$f(\mathbf{r}) = f(\mathbf{r}^*) + \frac{1}{2} (\mathbf{r} - \mathbf{r}^*)^T H^* (\mathbf{r} - \mathbf{r}^*) + \dots \quad (1.11.7)$$

Continuing, in order for this function to have a minimum at  $\mathbf{r}^*$  the second term on the right hand side must always be *positive*. Thus Hessian matrix must be such that for *any* vector  $\neq \mathbf{0}$   $v^T \mathbf{H} v > 0$ . A matrix having this property is called "positive definite" (PD). When a PD matrix is symmetric, then its eigenvalues must be strictly positive. We can summarize the necessary conditions for a minimum at  $\mathbf{r}^*$ :

$$\mathbf{g}^* = \nabla f(\mathbf{r}^*) = \mathbf{0} \quad \text{and} \quad \mathbf{H}^* \text{ is Pos. Def.} \quad (1.11.8)$$

As an example, let us take the function  $f(x) = x^2 + y^2$ . This function is always non-negative, so its minimum is at  $x = y = 0$ , where it obtains the value 0. The gradient is indeed zero at  $(x, y) = (0, 0)$ :

$$\begin{aligned} \left(\frac{\partial f}{\partial x}\right)_{(0,0)} &= (2x)_{(0,0)} = 0 \\ \left(\frac{\partial f}{\partial y}\right)_{(0,0)} &= (2y)_{(0,0)} = 0 \end{aligned} \quad (1.11.9)$$

And the Hessian is  $2I_2$  where  $I_2$  is the  $2 \times 2$  identity matrix, thus it is trivially PD.

## ii. Constrained minimization: Lagrange multipliers

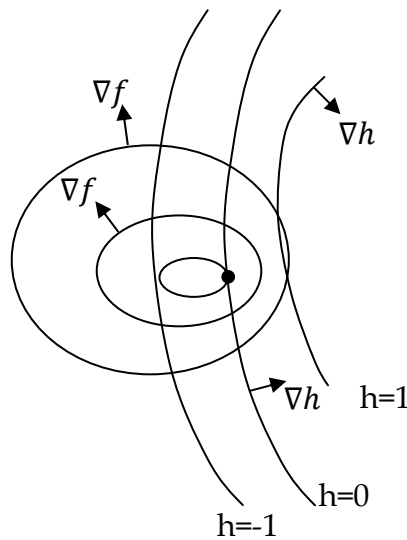
Next we discuss constrained minimization. Suppose we want to minimize  $f(x, y)$ , under a constraint that  $y = g(x)$ . This is "easy", since it can be transformed into an unconstrained minimization. All we have to do is minimize  $F(x) = f(x, g(x))$ . This will give:

$$F'(x) = \left(\frac{\partial f}{\partial x}\right)_{(x,g(x))} + \left(\frac{\partial f}{\partial y}\right)_{(x,g(x))} g'(x) \quad (1.11.10)$$

$$F''(x) = \left(\frac{\partial^2 f}{\partial x^2}\right)_{(x,g(x))} + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)_{(x,g(x))} g'(x) + \left(\frac{\partial^2 f}{\partial y^2}\right)_{(x,g(x))} g''(x) + \left(\frac{\partial^2 f}{\partial y^2}\right)_{(x,g(x))} g'(x)^2$$

And we can then search for  $x^*$  such that  $F'(x^*) = 0$  and  $F''(x^*) > 0$ .

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  $h(x, y) = 0$ . Suppose the point depicted in the following figure is this minimum.



This means that moving slightly along the isopleth of  $h = 0$  from  $\mathbf{r}^*$  will not change  $f$ : if it decreases  $f$  this is not a minimum and if it increases  $f$  we will move in the opposite direction and  $f$  will decrease.

Thus at  $\mathbf{r}^*$  the gradient of  $f$ ,  $\nabla f(\mathbf{r}^*)$ , must be normal to the line  $h = 0$ . Furthermore, *by definition* the gradient of  $h$  is normal to its isopleth. 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  $\lambda^*$ :

$$\begin{aligned}\nabla f(\mathbf{r}^*) &= \lambda^* \nabla h(\mathbf{r}^*) \\ h(\mathbf{r}^*) &= 0\end{aligned}\tag{1.11.11}$$

Lagrange noticed that this relation and wrote a *Lagrangian* for the constrained problem:

$$L(\mathbf{r}, \lambda) = f(\mathbf{r}) - \lambda h(\mathbf{r})\tag{1.11.12}$$

The new variable  $\lambda$  is called a "Lagrange multiplier". The necessary conditions for a minimum in Eq. (1.11.11) translate into the condition of finding a stationary point of  $L$

$$\nabla L(\mathbf{r}^*, \lambda^*) = 0\tag{1.11.13}$$

Where  $\nabla \equiv (\partial_x, \partial_y, \partial_z, \partial_\lambda)$ . The derivative with respect to  $\lambda$  gives us back the constraint. As an example, consider minimizing  $f(x, y) = x + y$  under the constraint  $h(x, y) = x^2 + y^2 - 1$  (i.e. find the point on the unit circle for which  $x + y$  is minimal). We have:

$$L(\mathbf{r}, \lambda) = (x + y) - \lambda(x^2 + y^2 - 1)\tag{1.11.14}$$

The necessary condition is:

$$\begin{aligned}\nabla L &= \left(1 - 2\lambda^* x^*, 1 - 2\lambda^* y^*, -(x^{*2} + y^{*2} - 1)\right) = (0, 0, 0) \\ \Rightarrow x^* &= y^* = \frac{1}{2\lambda^*}, \quad 2\left(\frac{1}{2\lambda^*}\right)^2 = 1\end{aligned}\tag{1.11.15}$$

The solution is thus:

$$\lambda^* = \pm \frac{1}{\sqrt{2}} \quad x^* = y^* = \pm \frac{1}{\sqrt{2}}\tag{1.11.16}$$

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

The method of Lagrange multipliers is often more convenient to work with than direct replacement. The problem is thus that of finding a minimum of  $f(\mathbf{r}_1, \dots, \mathbf{r}_N)$  under the  $M < N$  constraints  $\mathbf{h}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathbf{c}$ , or  $h_i(\mathbf{r}_1, \dots, \mathbf{r}_N) = c_i$   $i = 1, \dots, M$  is:

$$\begin{aligned} \text{find } \mathbf{r}^* \text{ such that: } \mathbf{h}(\mathbf{r}^*) = \mathbf{c} \text{ for which : } f(\mathbf{r}^*) \leq f(\mathbf{r}) \\ \text{for any } \mathbf{r} \text{ such that: } \mathbf{h}(\mathbf{r}) = \mathbf{c} \end{aligned} \quad (1.11.17)$$

To facilitate such a search, we formulate the Lagrangian function:

$$L(\mathbf{r}, \boldsymbol{\lambda}; \mathbf{c}) = f(\mathbf{r}) - \sum_{i=1}^M \lambda_i (h_i(\mathbf{r}) - c_i) \quad (1.11.18)$$

Our plan is to find the position of  $\mathbf{r}$  which minimizes  $L$  for any choice of  $\boldsymbol{\lambda}$  and then change  $\mathbf{c}$  until  $\mathbf{h}(\mathbf{r}) = \mathbf{c}$ . At  $\boldsymbol{\lambda}^*$  we have  $L$  assuming a minimum at a point  $\mathbf{r}^*$ . A necessary condition for the constrained minimum to be achieved at the point  $\mathbf{r}^*$  and with the Lagrange multipliers  $\boldsymbol{\lambda}^*$  is:

$$\begin{aligned} \nabla_{\mathbf{r}} L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) &= 0 \\ \nabla_{\boldsymbol{\lambda}} L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) &= \mathbf{h}(\mathbf{r}^*) - \mathbf{c} = 0 \end{aligned} \quad (1.11.19)$$

Note that  $\boldsymbol{\lambda}^*$  is not necessarily a minimizer of  $L$ . In fact the opposite is true:  $\boldsymbol{\lambda}^*$  is a maximizer of  $L$  and the search for constrained minimizations is a search for saddle points (see ref. [10] for a method to solve such a problem on the large scale).

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

$$\delta L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) = L(\mathbf{r}^* + \delta \mathbf{r}^*, \boldsymbol{\lambda}^* + \delta \boldsymbol{\lambda}; \mathbf{c} + \delta \mathbf{c}) - L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) \quad (1.11.20)$$

Now, because of Eq. (1.11.19) we find:

$$\begin{aligned}\delta L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) &= L(\mathbf{r}^* + \delta \mathbf{r}^*, \boldsymbol{\lambda}^* + \delta \boldsymbol{\lambda}^*; \mathbf{c} + \delta \mathbf{c}) - L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) \\ &= L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c} + \delta \mathbf{c}) - L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c})\end{aligned}\quad (1.11.21)$$

Thus:

$$\frac{d}{dc_i} L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) = \frac{\delta}{\delta c_i} L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) = \lambda_i^* \quad (1.11.22)$$

Since  $L(\mathbf{r}^*, \boldsymbol{\lambda}^*; \mathbf{c}) = f(\mathbf{r}^*; \mathbf{c})$  we find:

$$\frac{d}{dc_i} f(\mathbf{r}^*; \mathbf{c}) = \lambda_i^* \quad \text{or} \quad \nabla_{\mathbf{c}} f(\mathbf{r}^*; \mathbf{c}) = \boldsymbol{\lambda}^* \quad (1.11.23)$$

This equation reveals the “meaning” of the Lagrange multipliers  $\lambda_i^*$  at the optimal point: they are equal to the rate at which the optimal value of the minimized function  $f$  changes when  $c_i$ , the value of the  $i$  constraint, is changed. This is an important result which we use below whenever we want to give physical significance to Lagrange multipliers.

### iii. Minimization of functionals

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

$$\frac{\delta I}{\delta f(\mathbf{r})} = 0 \quad (1.11.24)$$

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 of this particle:

$$S[x] = \int_{t_0}^{t_f} \left[ \frac{1}{2} m \dot{x}(t)^2 + v(x(t)) \right] dt \quad (1.11.25)$$

For any trajectory  $x(t)$  between times  $t_0$  and  $t_f$ ,  $L[x]$  returns a number. Lagrange showed that finding the trajectory that makes  $L[x]$  stationary



(although, not necessarily minimal), under the condition that  $x(t_0) = x_0$  and  $x(t_f) = x_f$  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:

$$\begin{aligned} \frac{\delta}{\delta x(t)} \int_{t_0}^{t_f} \dot{x}(t')^2 dt' &= 2 \int_{t_0}^{t_f} \dot{x}(t') \frac{\delta \dot{x}(t')}{\delta x(t)} dt' = 2 \int_{t_0}^{t_f} \dot{x}(t') \delta(t - t') dt' \\ &= -2 \int_{t_0}^{t_f} \ddot{x}(t') \delta(t - t') dt' = -2\ddot{x}(t) \end{aligned} \quad (1.11.26)$$

The condition for stationarity of the action under changes in the trajectory which leave the edges intact give:

$$0 = \frac{\delta S}{\delta x(t)} = -m\ddot{x}(t) - v'(x(t)) \quad (1.11.27)$$

From which we obtain Newton's equation  $m\ddot{x}(t) = -v'(x(t))$ . This equation must be solved subject to the given constraints at the endpoints, i.e.  $x(t_0) = x_0$  and  $x(t_f) = x_f$ .

Minimizing functionals with constraints can again be done using Lagrange multipliers. Then one defines:

$$L[f, \lambda] = I[f] + \lambda h[f] \quad (1.11.28)$$

And the necessary conditions are:

$$\left. \frac{\delta L}{\delta f(\mathbf{r})} \right|_{f^*, \lambda^*} = 0 \quad \left. \frac{\delta L}{\delta \lambda} \right|_{f^*, \lambda^*} = h[f^*] = 0 \quad (1.11.29)$$

For example, let us solve the following problem. Find the shape in a plane of a closed contour encircling maximal area under the constraint of a given circumference  $l_0$ . We need to find the contour given by  $r(\theta)$  with  $r(\theta + 2\pi) = r(\theta)$  which gives zero variation to

$$L[r, \lambda] = S[r] + \lambda(L[r] - l_0) \quad (1.11.30)$$

Where  $S$  is the area and  $L$  is the circumference. We limit ourselves to curves that can be uniquely defined. A point on the curve, is defined by the angle  $\theta$  and the distance from the origin  $r(\theta)$ . The vector to a point on the curve and its derivative is:

$$\begin{aligned} \mathbf{r}(\theta) &= r(\theta)(\cos \theta, \sin \theta) \\ \mathbf{r}'(\theta) &= r'(\theta)(\cos \theta, \sin \theta) + r(\theta)(-\sin \theta, \cos \theta). \end{aligned} \quad (1.11.31)$$

The area of a small arc from  $\theta$  to  $\theta + d\theta$  is  $dS = \frac{1}{2}|\mathbf{r}(\theta) \times \mathbf{r}(\theta + d\theta)| = \frac{1}{2}|\mathbf{r}(\theta) \times \mathbf{r}'(\theta)|d\theta$ . Notice that  $|\mathbf{r} \times \mathbf{r}'| = r^2$  and so  $dS = \frac{1}{2}r(\theta)^2 d\theta$ .

The square circumference is:  $dL^2 = (\mathbf{r}(\theta + d\theta) - \mathbf{r}(\theta))^2 = (\mathbf{r}'(\theta))^2 d\theta^2$ . Thus  $dL = |\mathbf{r}'(\theta)|d\theta = \sqrt{r(\theta)^2 + r'(\theta)^2}d\theta$ . Thus, the area is a functional is:

$$\begin{aligned} S[r] &= \int_0^{2\pi} \frac{1}{2} r(\theta)^2 d\theta \\ L[r] &= \int_0^{2\pi} \sqrt{r(\theta)^2 + r'(\theta)^2} d\theta \end{aligned} \quad (1.11.32)$$

The functional derivative of  $S$  is easy:

$$\frac{\delta S[r]}{\delta r(\theta)} = r(\theta) \quad (1.11.33)$$

The functional derivative of  $L$  is computed by adding a "small" function  $\epsilon(\theta)$ :

$$\begin{aligned} \delta l &= \int_0^{2\pi} \sqrt{(r + \epsilon)^2 + (r' + \epsilon')^2} d\theta' - l \\ &= \int_0^{2\pi} \sqrt{r^2 + 2r\epsilon + r'^2 + 2r'\epsilon'} d\theta' - l \\ &= \int_0^{2\pi} \sqrt{r^2 + r'^2} \sqrt{1 + \frac{2r\epsilon + 2r'\epsilon'}{r^2 + r'^2}} d\theta' - l \end{aligned} \quad (1.11.34)$$

Then, we can use  $\sqrt{1 + \eta} = 1 + \frac{1}{2}\eta + O(\eta^2)$  so:

$$\delta l = \int_0^{2\pi} \frac{r\epsilon + r'\epsilon'}{\sqrt{r^2 + r'^2}} d\theta' \quad (1.11.35)$$

Integrating by parts the second term (remembering that the end terms drop because  $\theta = 2\pi$  is the same point as  $\theta = 0$ ):

$$\delta l = \int_0^{2\pi} \frac{r\epsilon}{\sqrt{r^2 + r'^2}} d\theta' - \int_0^{2\pi} \left( \frac{r}{\sqrt{r^2 + r'^2}} \right)' \epsilon d\theta' \quad (1.11.36)$$

We thus get:

$$\frac{\delta l}{\delta r(\theta)} = \frac{r}{\sqrt{r^2 + r'^2}} - \left( \frac{r}{\sqrt{r^2 + r'^2}} \right)' \quad (1.11.37)$$

We define the Lagrangian:

$$L[r] = S[r] - \lambda(l[r] - l_0) \quad (1.11.38)$$

And get the maximum from:

$$0 = \frac{\delta L}{\delta r(\theta)} = r - \lambda \left( \frac{r}{\sqrt{r^2 + r'^2}} - \left( \frac{r}{\sqrt{r^2 + r'^2}} \right)' \right) \quad (1.11.39)$$

Clearly, a solution is  $r(\theta) = \lambda$ , a constant, since then all derivatives are zero.

The curve is then circle of radius  $\lambda$  and given the circumference, we have:

$$\lambda = l_0/2\pi.$$

### EXERCISES FOR CHAPTER 1

- 1) Consider a system of two particles in a harmonic well. Each particle repels the other. The potential energy is  $V(x_1, x_2) = \frac{1}{2}kx_1^2 + \frac{1}{2}kx_2^2 - \frac{1}{2}q(x_1 - x_2)^2$ . We assume that the system is bound (so  $0 < 2q < k$ ). (a) Write down the exact Hamiltonian for this system, assuming the mass are  $m_1$  and  $m_2$ . (b) Make a transformation separating the Hamiltonian into two uncoupled particles: the center of mass  $X = \frac{m_1x_1+m_2x_2}{m_1+m_2}$  and the reduced mass  $x = x_2 - x_1$ . Write the Hamiltonian of each part and determine the mass of each particle. What are the masses when  $m_1 \gg m_2$ ? (c) Write down the energy levels of the system and the eigenfunctions. (d) Now make a Born-Oppenheimer approximation. Assuming  $m_1 \gg m_2$ , determine the Born

Oppenheimer potential surface (obtained when  $x_1$  is considered a parameter).

Solution: The Hamiltonian is:

$$H = -\frac{\hbar^2}{2m_1}\partial_1^2 - \frac{\hbar^2}{2m_1}\partial_2^2 + \frac{1}{2}kx_1^2 + \frac{1}{2}kx_2^2 - \frac{1}{2}q(x_1 - x_2)^2.$$

In the new coordinates:

$$X = \frac{(m_1x_1 + m_2x_2)}{M}, \quad x = x_2 - x_1, \quad x_1 = X - \frac{m_2}{M}x, \quad x_2 = X + \frac{m_1}{M}x.$$

If  $f(x_1, x_2)$  is any function than we can define the "same" scalar function as

$$F(x, X) = f\left(X - \frac{m_2}{M}x, X + \frac{m_1}{M}x\right).$$

- 2) Calculate the correlation in the case of a singlet, where both electrons are in the ground state orbital and when one is in the ground state orbital and the other in the first excited orbital.

## II. My first density functional:

### Thomas-Fermi Theory

#### *A. Basic concepts in the electron gas and the Thomas-Fermi Theory*

Thomas and Fermi assumed that the energy of an atom or a molecule can be written as a functional of the 1-particle density as follows:

$$E_{TF}[n] = T_{TF}[n] + \int v_{ext}(\mathbf{r})n(\mathbf{r})d^3r + \frac{1}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (2.1.1)$$

(Note, for use via the Born-Oppenheimer approximation, to this energy we need to add the nuclear-nuclear repulsion energy.) They then assumed that the density that characterizes the ground-state minimizes this functional under the constraint:

$$\int n(\mathbf{r})d^3r = N_e \quad (2.1.2)$$

The first question, beyond the rigor of this approach is, what is the kinetic energy functional? In order to take into account the Fermi nature and the

quantum nature of the electrons, this functional must include both these considerations. The Thomas Fermi solution is to assume:

$$T[n] = \int t_S(n(\mathbf{r}))n(\mathbf{r})d^3r \quad (2.1.3)$$

What shall we take for  $t_S(n)$ ? Consider first a simple case: a homogeneous gas of density  $n$  (i.e.  $n(\mathbf{r})$  is independent of  $\mathbf{r}$ ). Furthermore, let us assume that the electrons are non-interacting. This is a simple enough system to enable the analytic calculation of the kinetic energy functional. From the form of (2.1.3) we see that the total kinetic energy is the sum of contributions of various infinitesimal cells in space. Each cell contains  $n(\mathbf{r})d^3r$  electrons and so, if we interpret  $t(n)$  as the kinetic energy per electron of a homogeneous gas of non-interacting electrons then this sum yields exactly the total kinetic energy for this homogeneous gas. The Thomas-Fermi approximation then uses this same  $t(n)$  also for the inhomogeneous interacting case.

Let us now compute  $t(n)$ . Consider a *homogeneous gas of  $N$  uncharged electrons. They are non-interacting.* These electrons are put in a cubic cell of length  $L$ . The electron density is everywhere the same  $n = \frac{N_e}{V} = \frac{N_e}{L^3}$ .

We assume the wave functions are periodic in the box. According to Fourier's theorem, we can write any periodic wave function as a linear combination of plane-waves, as follows:

$$\psi(\mathbf{r}) = \sum_{\mathbf{n}} \frac{a_{\mathbf{n}} e^{i\frac{2\pi}{L}\mathbf{n}\cdot\mathbf{r}}}{\sqrt{V}} \quad (2.1.4)$$

Where:

$$\mathbf{n} = (l_x, l_y, l_z). \quad (2.1.5)$$

and  $l_{x/y/z}$  are integers. Fourier's theorem is based on the orthonormality of the plane waves

$$\langle \psi_{\mathbf{k}} | \psi_{\mathbf{k}'} \rangle = \frac{1}{V} \iiint_V e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}} d^3 r = \delta_{\mathbf{k}, \mathbf{k}'} \quad (2.1.6)$$

Where we defined

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{e^{i \mathbf{k} \cdot \mathbf{r}}}{\sqrt{V}} \quad \mathbf{k} = \frac{2\pi}{L} \mathbf{n} \quad (2.1.7)$$

We imagine 3-dimensional  $\mathbf{k}$ -space divided into an array of small compartments, indexed by a set of integers  $\mathbf{n} = (l_x, l_y, l_z)$  or by the vector  $\mathbf{k}$ . Each compartment is of  $\mathbf{k}$ -length  $\Delta k = \frac{2\pi}{L}$  and its  $\mathbf{k}$ -volume is  $\Delta k^3 = \frac{(2\pi)^3}{V}$ . For large  $\mathbf{r}$ -space boxes the  $\mathbf{k}$ -space compartment is extremely small since  $\Delta k^3$  is proportional to the inverse box volume. Since we are interested eventually in the limit  $L \rightarrow \infty$ , we may assume approximate sums of any function  $f(\mathbf{k})$  over the discrete values of  $\mathbf{k} = \frac{2\pi}{L} \mathbf{n}$  by integrals:

$$\sum_{\mathbf{k}} f(\mathbf{k}) \rightarrow \frac{V}{(2\pi)^3} \int f(\mathbf{k}) d^3 k \quad (V \rightarrow \infty) \quad (2.1.8)$$

Let's show that plane-waves are eigenstates of kinetic energy operator  $\hat{T}_1$ :

$$\hat{T}_1 \psi_{\mathbf{k}}(\mathbf{r}) = -\frac{\hbar^2}{2\mu} \frac{1}{\sqrt{V}} \nabla^2 e^{i\mathbf{k} \cdot \mathbf{r}} = \frac{\hbar^2 \mathbf{k}^2}{2m_e} \psi_{\mathbf{k}}(\mathbf{r}) \quad (2.1.9)$$

Now, consider the wavefunction of the  $N_e$  non-interacting electrons in their ground-state. Since they are non-interacting, this wave-function is a product of single-electron wave-functions:

$$\Psi = \psi_{\mathbf{k}_1}(\mathbf{r}_1) \bar{\psi}_{\mathbf{k}_1}(\mathbf{r}_2) \cdots \psi_{\mathbf{k}_{N_e/2}}(\mathbf{r}_{N_e-1}) \bar{\psi}_{\mathbf{k}_{N_e/2}}(\mathbf{r}_{N_e}) \quad (2.1.10)$$

Here  $\psi_{\mathbf{k}}(\mathbf{r})$  is the state of a spin-up electron with wave vector  $\mathbf{k}$ , while  $\bar{\psi}_{\mathbf{k}}(\mathbf{r})$  is the state of a spin down electron with wave vector  $\mathbf{k}$ . Anticipating the

antisymmetry, we build this wave function by placing 2 electrons in the same spatial orbital (once with spin up and the other with spin down). Since non-interacting electrons have only one type of energy, i.e. kinetic energy:

$\hat{H} = \sum_{n=1}^{N_e} \left( -\frac{\hbar^2}{2m_e} \nabla_n^2 \right)$ , we can easily show that (2.1.10) is an eigenstate of the Hamiltonian:

$$\begin{aligned} \hat{H}\Psi &= \sum_{n=1}^{N_e} \left( -\frac{\hbar^2}{2m_e} \nabla_n^2 \right) \psi_{\mathbf{k}_1}(\mathbf{r}_1) \bar{\psi}_{\mathbf{k}_1}(\mathbf{r}_2) \cdots \psi_{\mathbf{k}_{N_e/2}}(\mathbf{r}_{N_e-1}) \bar{\psi}_{\mathbf{k}_{N_e/2}}(\mathbf{r}_{N_e}) \\ &= \sum_{n=1}^{N_e/2} 2 \left( \frac{\hbar^2 \mathbf{k}_n^2}{2m_e} \right) \psi_{\mathbf{k}_1}(\mathbf{r}_1) \bar{\psi}_{\mathbf{k}_1}(\mathbf{r}_2) \cdots \psi_{\mathbf{k}_{N_e/2}}(\mathbf{r}_{N_e-1}) \bar{\psi}_{\mathbf{k}_{N_e/2}}(\mathbf{r}_{N_e}) \\ &= \sum_{n=1}^{N_e/2} 2 \left( \frac{\hbar^2 \mathbf{k}_n^2}{2m_e} \right) \Psi \end{aligned} \quad (2.1.11)$$

One sees that the energy is just the sum of kinetic energy  $\sum_{n=1}^{N_e/2} 2 \left( \frac{\hbar^2 \mathbf{k}_n^2}{2m_e} \right)$  in each spin-orbital of the product wave function. Let us now antisymmetrize this product wave function. We do this by adding all products resulting from even permutations of the electrons and subtracting all odd permutation products. One convenient way to represent such a sum is using a determinant, called a Slater wave function:

$$\Psi = \frac{1}{\sqrt{N!}} \det \begin{bmatrix} \psi_{\mathbf{k}_1}(\mathbf{r}_1) & \cdots & \bar{\psi}_{\mathbf{k}_{N_e}}(\mathbf{r}_1) \\ \vdots & \ddots & \vdots \\ \psi_{\mathbf{k}_1}(\mathbf{r}_{N_e/2}) & \cdots & \bar{\psi}_{\mathbf{k}_{N_e}}(\mathbf{r}_{N_e/2}) \end{bmatrix}. \quad (2.1.12)$$

For this wave function to be minimal energy must fill 2 electrons per level starting from the lowest kinetic energy and going up until electrons are exhausted. Denote the highest filled level by  $\mathbf{k}_F$ . Then:

$$N_{filled} = \frac{V}{(2\pi)^3} \iiint \theta(k_F - k) d^3k. \quad (2.1.13)$$

Where  $\theta(x)$  is 0 if  $x$  is negative and 1 otherwise. This is called the Heaviside function. We now perform the integral using spherical coordinates:

$$N_{filled} = \frac{V}{(2\pi)^3} \int_0^{k_F} 4\pi k^2 dk = \frac{V}{(2\pi)^3} \frac{4\pi}{3} k_F^3 = \frac{V}{2\pi^2} \frac{k_F^3}{3}. \quad (2.1.14)$$

The number of filled orbitals is the product of the real-space volume  $V$  and the k-space occupied state volume, divided by  $(2\pi)^3$ . Since  $N_e = 2N_{filled}$  and the density is  $n = \frac{N_e}{V}$  we have:

$$n = \frac{2N_{filled}}{V} = \frac{k_F^3}{3\pi^2}. \quad (2.1.15)$$

The electron density determined directly the highest filled momentum state:

$$T_S[n] = \frac{2V}{(2\pi)^3} \int_0^{k_F} \left( \frac{\hbar^2 k^2}{2m_e} \right) 4\pi k^2 dk = \frac{2V}{(2\pi)^3} 4\pi \frac{\hbar^2}{2m_e} \frac{k_F^5}{5} = \frac{2V}{2\pi^2} \frac{\hbar^2}{2m_e} \frac{k_F^5}{5}. \quad (2.1.16)$$

The energy per particle is:

$$t_S(n) = \frac{T_S}{N} = \frac{1}{\pi^2 n} \frac{\hbar^2}{2m_e} \frac{k_F^5}{5} = \frac{3}{5} \frac{\hbar^2 k_F^2}{2m_e} = \frac{3}{5} \frac{\hbar^2}{2m_e} (3\pi^2 n)^{2/3} = C_{TF} n^{2/3}, \quad (2.1.17)$$

where:

$$C_{TF} = \frac{3}{5} \frac{\hbar^2}{2m_e} (3\pi^2)^{2/3} = \frac{3}{10} (3\pi^2)^{2/3} \text{au} = 2.871 \text{ au}. \quad (2.1.18)$$

Plugging into Eq. (2.1.3), the Thomas-Fermi kinetic energy functional is obtained to be used in Eq. (2.1.1):

$$T_S[n] = C_{TF} \int n(\mathbf{r})^{5/3} d^3r. \quad (2.1.19)$$

### Exercise: The Thomas Fermi functional for the hydrogen atom.

- Repeat the calculation above but now for a “spin-polarized 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.



- b. Since the electron in a hydrogen-like atom is spin-polarized, use the Thomas-Fermi KE functional derived in (a) and compare its estimation of the kinetic energy of the electron in a hydrogen-like atom to the exact value. Using the exact kinetic energy in the hydrogen atom (you can find it using the virial theorem), assess the quality of the result as a function of the nucleus charge  $Z$ .

### ***B.Minimization of the Thomas-Fermi energy***

Now, according to the TF theory, the true ground-state electron density is the one that *minimizes*  $E_{TF}[n]$ . But the electron density must also account for the required number of electrons of  $N_e$ , so there is a constraint for the minimization:

$$\int n(\mathbf{r})d^3r = N_e \quad (2.1.20)$$

Thus, we must build a Lagrangian to be minimized as:

$$L[n, \lambda] = E_{TF}[n] - \mu \left[ \int n(\mathbf{r})d^3r - N_e \right] \quad (2.1.21)$$

Minimizing it gives the Thomas-Fermi equation:

$$0 = \frac{\delta L}{\delta n(\mathbf{r})} = \frac{\delta E_{TF}}{\delta n(\mathbf{r})} - \mu \quad (2.1.22)$$

We see that the Lagrange constant  $\mu$  is the chemical potential, since it is equal to the change in energy when we perturb the density and this change is everywhere constant. We must now compute the functional derivatives:

$$\begin{aligned} \frac{\delta}{\delta n(\mathbf{r})} \int C_{TF}n(\mathbf{r}')^{5/3}d^3r' &= \frac{5}{3}C_{TF}n(\mathbf{r})^{2/3} \\ \frac{\delta}{\delta n(\mathbf{r})} \left( \int v(\mathbf{r}')n(\mathbf{r}')d^3r' \right) &= v(\mathbf{r}) \end{aligned} \quad (2.1.23)$$

$$\frac{\delta}{\delta n(\mathbf{r})} \frac{1}{2} \iint \frac{n(\mathbf{r}'')n(\mathbf{r}')}{|\mathbf{r}'' - \mathbf{r}'|} d^3r'' d^3r' = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r'$$

Plugging into Eq.(2.1.22), one obtains the Thomas-Fermi equation for an atom:

$$\mu = \frac{5}{3} C_{TF} n(r)^{2/3} + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \quad (2.1.24)$$

This is an integral equation for  $n(\mathbf{r})$ . It is called the integral Thomas-Fermi equation. The potential  $v(\mathbf{r})$  is due to the positive charge, hence we can write:

$v(\mathbf{r}) = - \int \frac{n_+(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r'$ . Now we can define a *total potential*

$$\phi(\mathbf{r}) = \mu + \int \frac{n_+(\mathbf{r}') - n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r', \quad (2.1.25)$$

as the sum of the total electrostatic potential and the chemical potential. Since  $\nabla^2 \frac{1}{r} = 4\pi\delta(\mathbf{r})$ , we have:

$$-\nabla^2 \phi(\mathbf{r}) = 4\pi(n_+(\mathbf{r}) - n(\mathbf{r})), \quad (2.1.26)$$

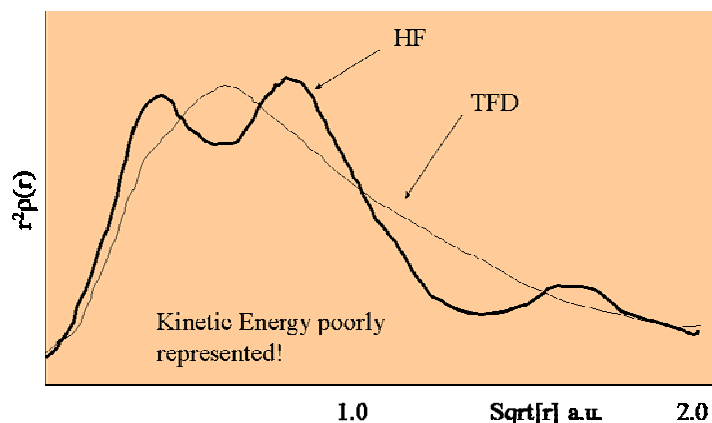
leading to an equation for the total potential:

$$\frac{1}{4\pi} \nabla^2 \phi = \left[ \frac{3}{5C_{TF}} \phi(\mathbf{r}) \right]^{3/2} - n_+(\mathbf{r}) \quad (2.1.27)$$

This is called the "differential **Thomas Fermi equation**". The constant  $\mu$  is buried in  $\phi$  but it did not disappear: it must be chosen so that Eq. (2.1.20) is obeyed. Furthermore, it is clear that the potential  $\phi(\mathbf{r})$  is manifestly non-negative in TF theory (otherwise we could not take its square root).

As we showed (see Eq. (1.11.23)),  $\mu$  can be shown to act as a chemical potential, i.e.  $-\frac{\partial E_{TF}(N)}{\partial N} = \mu(N)$ . We will not solve this equation for atoms or

molecules. We just comment that it gives a smoothed value for the atomic density, not showing the shell structure. We can see this in the following figure, where we plot the radial density as computed by a relatively accurate theory, such as Hartree-Fock and the TFD (Thomas-Fermi-Dirac) theory.



There is a question of how does the minimal energy of the Thomas Fermi functional compare with the accurate quantum mechanical energy. This question has been examined. It was found that for atoms with  $Z \rightarrow \infty$  we have:

$$\lim_{Z \rightarrow \infty} \frac{E_{TF}(N = \lambda Z)}{E_{exact}(N = \lambda Z)} = 1 \quad (2.1.28)$$

For  $0 < \lambda < 1$  (i.e. the number of electrons is smaller than that of the protons and  $\frac{N}{Z}$  is held while  $Z \rightarrow \infty$ ). Note that the Thomas Fermi energy for an atom has the property that:

$$E_{atom}^{TF}(\lambda, Z) = Z^{7/3} E_{atom}^{TF}(\lambda, 1) \quad (2.1.29)$$

### ***C. Thomas-Fermi does not account for molecules***

Consider an existing system, and now add a bit of positive charge  $\delta q = \int \delta n_+(\mathbf{r}) d^3r$ , where  $\delta n_+(\mathbf{r}) \geq 0$  everywhere. Also, increase the electronic charge  $\delta q$  so as to preserve neutrality. The electron distribution is changed by  $\delta n(\mathbf{r})$  (so that  $\delta q = \int \delta n(\mathbf{r}) d^3r$ ). The change in the total energy  $\delta E[n]$  is composed of the change in electronic energy,  $\mu \delta q$ , and the corresponding change in the electrostatic positive charge energy:

$$\begin{aligned}\delta E[n] &= \delta E_{TF}[n] + \delta E_{NN}[n] \\ &= \mu\delta q + \iint \frac{(n_+(\mathbf{r}) - n(\mathbf{r}))\delta n_+(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r'\end{aligned}\quad (2.1.30)$$

Using the definition of the total potential in Eq. (2.1.25) we find:

$$\delta E[n] = \int \phi(\mathbf{r})\delta n_+(\mathbf{r})d^3r \quad (2.1.31)$$

Since *everywhere we add positive charge*, i.e.  $\delta n_+(\mathbf{r}) \geq 0$  and since  $\phi(\mathbf{r})$  is non-negative everywhere, this process *increases* the total energy! In TF theory addition of infinitesimal positive charge followed by addition of a compensating electronic charge causes an *increase* in total energy of the system. In a more elaborate treatment Teller showed[11] that the total energy of any diatomic molecule is higher than the sum of energy of its constituent atoms, i.e. that TF theory cannot account for stable molecules! He concluded that Thomas Fermi theory is not very useful for chemistry.

#### ***D. Thomas-Fermi Screening***

When a point impurity  $Ze$  is inserted into an electronic system, it pulls ( $Z$  positive) or repels ( $Z$  negative) electrons towards it. This has an effect that the impurity is partially screened by opposite charge and so it has a smaller effect on distant charges. Let us study this phenomenon in the electron gas, using Thomas-Fermi theory. The homogeneous gas of electrons is a model for ideal metals, so the screening effect we address here is relevant for many metallic systems. Macroscopically, the “free” metal electrons completely screen the charged impurity. However microscopically, perfect screening is not possible because electrons have kinetic energy – even at zero temperature – and a short ranged electric field develops around the impurity. Thomas Fermi theory

takes kinetic energy effects into account and can be used to estimate the form of the local electric field, specifically its size or length scale.

Let us study an unperturbed homogeneous electron gas using Thomas-Fermi theory. Such a “gas” has no structure and it is characterized by only one parameter: its density  $n_0$ . In order to neutralize it and support the electron homogeneity, we add positive smeared homogeneous charge density  $+en_0$ . All the Coulomb energies (e-e, e-N and N-N) cancel exactly so the only energy left is the electronic kinetic energy:

$$E_{TF}[n_0] = \int C_{TF} n_0(\mathbf{r})^{5/3} d^3r \quad (2.2.1)$$

The constraint minimization of this functional yields the following condition, relating the density to the chemical potential:

$$\mu = \frac{\delta E_{TF}}{\delta n_0(\mathbf{r})} = \frac{5}{3} C_{TF} n_0(\mathbf{r})^{2/3} \quad (2.2.2)$$

Comparing with Eq. (2.1.15), and using Eq. (2.1.18) we find for the chemical potential:

$$\mu = \frac{\delta E_{TF}}{\delta n_0(\mathbf{r})} = \frac{5}{3} C_{TF} \left( \frac{k_F^3}{3\pi^2} \right)^{2/3} = \frac{\hbar^2 k_F^2}{2m_e} \quad (2.2.3)$$

Thus we see that indeed the electron density is constant and the chemical potential is equal to the kinetic energy corresponding to the maximal occupied momentum  $\hbar k_F$ .

Now we introduce a positive charge  $Ze$ . The density of electrons is changed:

$$n(\mathbf{r}) = n_0 + n_1(\mathbf{r}) \quad (2.2.4)$$

It is physically clear that  $n_1(\mathbf{r})$  is localized around the impurity (assumed at the origin). We therefore have for the total energy of the system in terms of  $n_1$ :

$$E_{TF}[n_1] = \int C_{TF}(n_0 + n_1(\mathbf{r}))^{5/3} d^3r - \int \frac{Ze^2 n_1(\mathbf{r})}{r} d^3r + \frac{e^2}{2} \iint \frac{n_1(\mathbf{r})n_1(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (2.2.5)$$

The corresponding TF equation comes from minimizing:

$$\mu = \frac{\delta E_{TF}}{\delta n_1(\mathbf{r})} = \frac{5}{3} C_{TF}(n_0 + n_1(\mathbf{r}))^{2/3} - \frac{Ze^2}{r} + e^2 \int \frac{n_1(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (2.2.6)$$

We write:  $\phi(\mathbf{r}) = e \int \frac{n_1(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r'$  and so:

$$\mu = \frac{\delta E_{TF}}{\delta n_1(\mathbf{r})} = \frac{5}{3} C_{TF}(n_0 + n_1(\mathbf{r}))^{2/3} - \frac{Ze^2}{r} + e\phi(\mathbf{r}) \quad (2.2.7)$$

Upon linearizing, assuming  $n_1 \ll n_0$ :

$$\frac{5}{3} C_{TF} n_0^{2/3} \left(1 + \frac{2n_1}{3n_0}\right) - \frac{Ze^2}{r} + e\phi(\mathbf{r}) = \mu \quad (2.2.8)$$

We can write:  $\frac{5}{3} C_{TF} n_0^{2/3} = \mu_0$  and so:

$$\frac{10}{9} C_{TF} n_0^{-1/3} n_1 - \frac{Ze^2}{r} + e\phi(\mathbf{r}) = \mu - \mu_0 \quad (2.2.9)$$

Finally since  $\nabla^2 \phi = -4\pi e n_1$  we have:

$$-\frac{10}{9} \frac{1}{4\pi e^2} n_0^{-1/3} C_{TF} \nabla^2 \phi - \frac{Ze}{r} + \phi(\mathbf{r}) = \frac{\mu - \mu_0}{e} \quad (2.2.10)$$

We have from Eq. (2.1.15)  $k_F = (3\pi^2 n_0)^{1/3}$  and we use the definition of the Bohr radius  $a_0 = -\frac{\hbar^2}{m_e e^2}$ , defining the Thomas Fermi screening parameter  $k_{TF}$ :

$$\frac{10}{9} \frac{1}{4\pi e^2} n_0^{-1/3} C_{TF} = \frac{\pi a_0}{4k_F} \equiv \frac{1}{k_{TF}^2} \quad (2.2.11)$$

With this we have the equation:

$$\nabla^2 \phi = k_{TF}^2 \left( \phi(\mathbf{r}) - \frac{Ze}{r} - \frac{\mu - \mu_0}{e} \right) \quad (2.2.12)$$

Passing to spherical coordinates we find:

$$\frac{1}{r}(r\phi(\mathbf{r}))'' = k_{TF}^2 \left( \phi(\mathbf{r}) - \frac{Ze}{r} - \Delta\mu \right). \quad (2.2.13)$$

Defining  $\chi = r\phi$  we find:

$$\chi'' = k_{TF}^2(\chi - Ze - \Delta\mu r) \quad (2.2.14)$$

The homogeneous equation is  $\chi'' = k_{TF}^2\chi$  which has the solution  $\chi_H = Ae^{-k_{TF}r} + Be^{k_{TF}r}$ . Clearly, for a localized potential solution we must take  $B = 0$ . To this we need add *any* solution of the inhomogeneous equation which clearly is  $\chi_{IH} = Ze + \Delta\mu r$ . Thus:

$$\chi = Ae^{-k_{TF}r} + Ze + \Delta\mu r \quad (2.2.15)$$

This leads to:

$$\phi = \frac{Ae^{-k_{TF}r}}{r} + \frac{Ze}{r} + \Delta\mu \quad (2.2.16)$$

In the limit that  $r \rightarrow 0$  we must have  $r\phi(r) \rightarrow 0$  since the electronic charge  $n_1$  has no cusps. Thus  $A = -Ze$ . The total electrostatic potential is

$$\phi_{tot}(r) = \phi(r) - \frac{Ze}{r} = \frac{-Ze}{r}e^{-k_{TF}r} + \Delta\mu \quad (2.2.17)$$

Aside from the constant  $\Delta\mu$ , far from the impurity the surface integral of  $\nabla\phi_{tot}$  evaluates to zero and by Gauss's theorem a large sphere around the impurity includes zero charge in it, meaning that the total amount of electronic charge pulled into the sphere is exactly equal to that of the impurity ( $Z$ ).

It is interesting that the screening length is proportional to  $k_F^{-1/2}$  or to  $n_0^{-1/6}$ . The higher the density the smaller the length, i.e. the more efficient is the screening, however, the dependence on  $n_0$  is mild because of the small exponent. It is also interesting to note that  $k_{TF}$  is independent of  $Z$ . However, this latter results holds only in so far as our linearization is valid.

impurities the non-linear equation will give a different result and the screening will depend on  $Z$ .

### ***E. 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. 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 the kinetic energy is:  $T = \int \psi(\mathbf{r}) \left( -\frac{\hbar^2}{2\mu_e} \nabla^2 \right) \psi(\mathbf{r}) d^3r$ . For wave functions that decay to zero at  $r \rightarrow \infty$ , one can integrate by parts and obtain  $T = \frac{\hbar^2}{2\mu_e} \int (\nabla\psi(\mathbf{r}))^2 d^3r$ , stressing the absolute positivity of kinetic energy (it cannot be zero). Finally, if  $\psi(\mathbf{r})$  is a non-degenerate ground-state it can be written as  $\psi(\mathbf{r}) = \sqrt{n(\mathbf{r})}$  and so we obtain the kinetic energy functional of von Weizsäcker:

$$T_{vW}[n] = \frac{\hbar^2}{2\mu_e} \int (\nabla\sqrt{n(\mathbf{r})})^2 d^3r \quad (2.2.18)$$

Which can be written as follows, using local wave vector:

$$\mathbf{k}(\mathbf{r}) \equiv \frac{1}{2} \frac{\nabla n(\mathbf{r})}{n(\mathbf{r})} \quad (2.2.19)$$

So:

$$T_{vW}[n] = \int \frac{\hbar^2 \mathbf{k}(\mathbf{r})^2}{2\mu_e} n(\mathbf{r}) d^3r \quad (2.2.20)$$

This functional is now used for any density, even a many electron one. The variation is:

$$\delta T_{vW} = \frac{\hbar^2}{8\mu_e} \int \left( \frac{(\nabla(n(\mathbf{r}') + \epsilon\Phi(\mathbf{r}')))^2}{n(\mathbf{r}') + \epsilon\Phi(\mathbf{r}')} - \frac{(\nabla(n(\mathbf{r}')))^2}{n(\mathbf{r}')} \right) d^3r' \quad (2.2.21)$$



Working this out to linear terms in  $\epsilon$ , using:  $(n(\mathbf{r}') + \epsilon\Phi(\mathbf{r}'))^{-1} \approx n(\mathbf{r}')^{-1} (1 - \epsilon\Phi(\mathbf{r}')/n(\mathbf{r}'))$  we obtain:

$$\delta T_{vW} = \frac{\hbar^2}{8\mu_e} \epsilon \int \left( \frac{2\nabla n(\mathbf{r}') \cdot \nabla \Phi(\mathbf{r}')}{n(\mathbf{r}')} - \left( \frac{\nabla(n(\mathbf{r}'))}{n(\mathbf{r}')} \right)^2 \frac{\Phi(\mathbf{r}')}{n(\mathbf{r}')^2} \right) d^3r' \quad (2.2.22)$$

Which after integration by parts of the first term finally gives:

$$\delta T_{vW} = \frac{\hbar^2}{8\mu_e} \epsilon \int \left( -2\nabla \cdot \left( \frac{\nabla n(\mathbf{r}')}{n(\mathbf{r}')} \right) - \left( \frac{\nabla(n(\mathbf{r}'))}{n(\mathbf{r}')} \right)^2 \right) \Phi(\mathbf{r}') d^3r' \quad (2.2.23)$$

Thus the von-Weizsäcker potential is:

$$v_{vW}(\mathbf{r}) = -\frac{\hbar^2}{8\mu_e} \left( 2\nabla \cdot \left( \frac{\nabla n(\mathbf{r})}{n(\mathbf{r})} \right) + \left( \frac{\nabla(n(\mathbf{r}))}{n(\mathbf{r})} \right)^2 \right) \quad (2.2.24)$$

Which can be written more compactly as:

$$v_{vW}(\mathbf{r}) = -\frac{\hbar^2}{2\mu_e} [\nabla \cdot \mathbf{k}(\mathbf{r}) + k(\mathbf{r})^2] \quad (2.2.25)$$

**Exercise:** For 1-electron system, discuss the claims: 1) The wave vector  $\mathbf{k}(\mathbf{r})$  is the gradient of the log of the wavefunction:  $\mathbf{k}(\mathbf{r}) = \nabla \log \psi(\mathbf{r})$  (2) the von Weizsäcker potential is the potential for which  $n(\mathbf{r})$  is the ground state density.

### III. Many-electron wave functions

#### A. The electron spin

Zeeman has shown that a small magnetic field causes the splitting of energy levels in atoms. Each atomic level is split into a doublet. The amount of splitting is proportional to the field. At zero field these doublets are degenerate. The conclusion is that the electron has an intrinsic magnetic moment which can take two values. The states of the internal magnetic

moment of the electron are assumed to be proportional to an internal angular momentum called “spin”. The spin of an electron can be assumed to have two values  $\pm \frac{\hbar}{2}$ . This is an additional “degree of freedom”. It is not continuous, but is nevertheless it is degree of freedom. We denote a spin-orbital  $\psi_n(\mathbf{x}) = \psi_n(\mathbf{r}, s)$  where  $\mathbf{r}$  is a point in 3D space and  $s$  is a “spin variable”, which allows us to perform an inner product of spin as explained now. There are two possible spin functions for an electron,  $\alpha(s)$  denotes spin up and  $\beta(s)$  spin down. These two states are complete and orthonormal:

$$\int \alpha(s)^* \alpha(s) ds = \langle \alpha | \alpha \rangle = 1$$

$$\int \beta(s)^* \alpha(s) ds = \langle \beta | \alpha \rangle = 0$$

$$\int \alpha(s)^* \beta(s) ds = \langle \alpha | \beta \rangle = 0$$

$$\int \beta(s)^* \beta(s) ds = \langle \beta | \beta \rangle = 1$$

The variable  $s$  is just a mnemonic. With new notation, we have:

$$\langle \psi | \phi \rangle = \int \psi(\mathbf{x})^* \phi(\mathbf{x}) d\mathbf{x} = \int \int \psi(\mathbf{r}, s)^* \phi(\mathbf{r}, s) ds d^3r \quad (3.1.1)$$

### ***B. The Pauli principle***

The electronic wavefunctions are functions of  $N_e$  electronic coordinates and spins  $\psi(\mathbf{x}_1, \dots, \mathbf{x}_{N_e})$ . Here  $\mathbf{x}_j \equiv (\mathbf{r}_j, s_j)$ . The Pauli principle states that this wavefunction must be antisymmetric with respect to interchange of two electrons:

$$\psi(\dots \mathbf{x}_i, \dots, \mathbf{x}_j, \dots) \equiv -\psi(\dots \mathbf{x}_j, \dots, \mathbf{x}_i, \dots) \quad (3.2.1)$$

This is a “boundary condition” we impose while solving for any electronic wave function.

### ***C. The Excited states of the Helium atom***

How should we represent the, in an approximate form, the low lying excited states of the Helium atom. He<sup>+</sup> has two low lying orbitals 1s and 2s (the 2p orbitals are degenerate with the 2s, but we will not consider them because in the Helium atom they are of much higher energy. We can form a 2-electron wavefunction by:  $\psi(\mathbf{x}_1, \mathbf{x}_2) = \psi_{1s}(r_1)\psi_{1s}(r_2)[\alpha(s_1)\beta(s_2) - \alpha(s_2)\beta(s_1)]$ .

The excited states will involve excitation of an electron to the 2s orbital. We can then write:

$$\begin{aligned}
 \psi_2(x_1, x_2) &= [\psi_{1s}(r_1)\psi_{2s}(r_2) - \psi_{1s}(r_2)\psi_{2s}(r_1)]\alpha(s_1)\beta(s_2) \\
 \psi_3(x_1, x_2) &= [\psi_{1s}(r_1)\psi_{2s}(r_2) - \psi_{1s}(r_2)\psi_{2s}(r_1)][\alpha(s_1)\beta(s_2) \\
 &\quad + \alpha(s_2)\beta(s_1)] \\
 \psi_4(x_1, x_2) &= [\psi_{1s}(r_1)\psi_{2s}(r_2) - \psi_{1s}(r_2)\psi_{2s}(r_1)][\beta(s_1)\beta(s_2)] \\
 \psi_5(x_1, x_2) &= [\psi_{1s}(r_1)\psi_{2s}(r_2) + \psi_{1s}(r_2)\psi_{2s}(r_1)][\alpha(s_1)\beta(s_2) \\
 &\quad - \alpha(s_2)\beta(s_1)]
 \end{aligned}
 \tag{3.3.1}$$

The first 3 states form a triplet the total spin is 1. The last is again a singlet (like the ground state).

### ***D. The Slater wave function is the basic anti-symmetric function describing N electrons in N orbitals***

The previous example is difficult to generalize. In order to develop a way to easily represent antisymmetric functions of all types, we consider the following 2-electron function, composed of 2 1-electron spin-orbitals:

$$\begin{aligned}\Psi(x_1, x_2) &= \frac{1}{\sqrt{2}} [\phi_1(x_1)\phi_2(x_2) - \phi_1(x_2)\phi_2(x_1)] \\ &= \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) \\ \phi_2(x_1) & \phi_2(x_2) \end{vmatrix}\end{aligned}\quad (3.4.1)$$

If we choose the orbitals to be orthonormal,  $\langle \phi_i | \phi_j \rangle = \delta_{ij}$  then:

$$\begin{aligned}\langle \Psi | \Psi \rangle &= \iint \Psi(x_1, x_2)^* \Psi(x_1, x_2) dx_1 dx_2 \\ &= \frac{1}{2} \iint [\phi_1(x_1)\phi_2(x_2) - \phi_1(x_2)\phi_2(x_1)]^2 dx_1 dx_2 \\ &= \frac{1}{2} \iint [\phi_1(x_1)^2 \phi_2(x_2)^2 + \phi_1(x_2)^2 \phi_2(x_1)^2 \\ &\quad + 2\phi_1(x_2)\phi_2(x_1)\phi_1(x_2)\phi_2(x_1)] dx_1 dx_2 = 1\end{aligned}\quad (3.4.2)$$

***E. Without loss of generality, we may assume the orbitals of a Slater wave function are orthogonal***

But what happens if the orbitals are not orthogonal? Suppose that the orbitals were not orthonormal:

$$\langle \phi_i | \phi_j \rangle = S_{ij} \quad (3.5.1)$$

It is then possible to “orthonormalize” them. i.e define two linear combinations which are orthonormal. Define:

$$\xi_n = \sum_m \phi_m A_{mn} \quad (3.5.2)$$

and demand:  $\langle \xi_i | \xi_j \rangle = \delta_{ij}$ . Then:

$$\delta_{ij} \langle \xi_i | \xi_j \rangle = \sum_{m,n} A_{mi}^* \langle \phi_i | \phi_j \rangle A_{nj} = (A^\dagger S A)_{ij} \quad (3.5.3)$$

Thus:  $A^\dagger S A = I$  or  $AA^\dagger = S^{-1}$ . Note also that  $\det A = (\det S)^{-1/2}$ . There are many solutions to this equation. (For example, if  $A_0$  is a solution then so is  $A_0 U$  where  $U$  is any unitary matrix.) Each solution will give us a different set

of orthonormal orbitals. The Slater wave function made out from these new orbitals is:

$$\begin{aligned}\det\{\xi_i(x_j)\} &= \det\{(A^T)_{ni}\phi_i(x_j)\} = \det A \times \det\{\phi_i(x_j)\} \\ &= \frac{1}{\sqrt{\det S}} \det\{\phi_i(x_j)\}\end{aligned}\quad (3.5.4)$$

Thus, the new wavefunction is the same as the old one, up to multiplication by constant! Yet, it is always more convenient to work with normalized orbitals, so we can assume the orbitals are orthonormal without any loss of generality. This development also shows that given any set of  $N$  orbitals from which the Slater wave function has been constructed, we can take  $N$  linear combinations of the orbitals to obtain new orbitals that give the same Slater wave function up to a constant factor.

***F. Any antisymmetric function can be expanded as a sum of basic Slater (determinantal) functions***

For orthonormal orbitals, the normalization is easy to compute. We write explicitly the determinant as:

$$\det[\phi_1 \dots \phi_N](x_1, \dots, x_N) = \sum_{i_1 \dots i_N} (-)^{P_{i_1 \dots i_N}} \prod_{k=1}^N \phi_{i_k}(x_k) \quad (3.6.1)$$

Where  $(i_1 \dots i_N)$  is a permutation of the numbers  $1 \dots N$  (there are  $N!$  such permutations). Each permutation can be obtained from a series of pair swapping operations. For example: (132) is obtained from (123) by switching the pair of numbers in position 2 and 3. We write this as:  $(132) = S_{23}(123)$ . (2413) is obtained from (1234) by three operations:

$$S_{23}S_{34}S_{12}(1234) = S_{23}S_{34}(2134) = S_{23}(2143) = (2413) \quad (3.6.2)$$

If the number of switches is odd the permutation is odd and  $P_{i_1 \dots i_N} = 1$ ; if the number of switches is even, the permutation is even and  $P_{i_1 \dots i_N} = 0$ . The normalization of a determinantal wave function composed of orthonormal orbitals is:

$$\begin{aligned}
 & \int |\det[\phi_1 \dots \phi_N](\mathbf{x}_1, \dots, \mathbf{x}_N)|^2 dx_1 \dots dx_N \\
 &= \sum_{i_1 \dots i_N j_1 \dots j_N} (-)^{P_{i_1 \dots i_N}} (-)^{P_{j_1 \dots j_N}} \int \prod_{k=1}^N \phi_{i_k}(\mathbf{x}_k) \prod_{q=1}^N \phi_{j_q}(\mathbf{x}_q) dx_1 \dots dx_N \quad (3.6.3) \\
 &= \sum_{i_1 \dots i_N j_1 \dots j_N} (-)^{P_{i_1 \dots i_N}} (-)^{P_{j_1 \dots j_N}} \prod_{k=1}^N \langle \phi_{i_k} | \phi_{j_k} \rangle
 \end{aligned}$$

Because of orthonormality the orbital integral  $\langle \phi_{i_k} | \phi_{j_k} \rangle$  is zero unless  $i_k = j_k$ . These integrals appear in products so the product is non-zero only if  $i_k = j_k$  for *all*  $k = 1 \dots N$ . The only conclusion is, that the two permutations *must* be identical and:

$$\int |\det[\xi_1 \dots \xi_N](\mathbf{x}_1, \dots, \mathbf{x}_N)|^2 dx_1 \dots dx_N = \sum_{i_1 \dots i_N j_1 \dots j_N} 1 = N! \quad (3.6.4)$$

We conclude that the normalization factor of a determinantal wave function of orthonormal orbitals is  $\frac{1}{\sqrt{N!}}$  and write:

$$\frac{1}{\sqrt{N!}} \det[\xi_{i_1} \dots \xi_{i_N}] \equiv |i_1 \dots i_N\rangle \quad (3.6.5)$$

Given a set of  $M > N$  orthonormal single-electron spin-orbitals  $\phi_n(\mathbf{x}) = \phi_n(\mathbf{r}, s)$   $n = 1, 2, \dots, M$ , we can consider the space of all linear combinations of all  $N$ -particle determinants that can be made. There are  $\binom{M}{N} = \frac{M!}{N!(M-N)!}$  ways to select determinants so this is the dimension of the space. The dimension grows factorially with  $M$ . A typical antisymmetric wave function can be approached by linear combinations of these determinants:

$$\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \sum_{\{i_1, \dots, i_N\}} C_{i_1, \dots, i_N} \det[\phi_{i_1}(\mathbf{x}_1) \dots \phi_{i_N}(\mathbf{x}_N)] \quad (3.6.6)$$

The sum is over all selections of  $N$  integers, where each selection is ordered so that  $i_1 < i_2 < \dots$ . If the orbitals are orthogonal, the constants  $C_{i_1, \dots, i_N}$  are obtained from:

$$C_{i_1, \dots, i_N} = \frac{1}{\sqrt{N!}} \int \det[\phi_{i_1}(\mathbf{x}_1) \dots \phi_{i_N}(\mathbf{x}_N)] \Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) d\mathbf{x}_1 \dots d\mathbf{x}_N \quad (3.6.7)$$

### ***G. Determinant expectation values***

In this section we discuss the calculation of expectation values of many-electron operators for  $N$  electrons within a given Slater wave function  $\Psi_S(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det[\phi_1(\mathbf{x}_1) \dots \phi_N(\mathbf{x}_N)]$ . We assume the orbitals  $\phi_i(\mathbf{x})$  are orthonormal:  $\langle \phi_i | \phi_j \rangle = \delta_{ij}$ .

#### **i. One-body operators**

Consider an operator  $\hat{o}_1(\mathbf{x})$  which operates on an electron with spin-coordinates  $\mathbf{x}$ . For  $N$  electrons we define the sum of  $\hat{o}_1$  for each electron

$$\hat{O} = \sum_{n=1}^N \hat{o}_1(\mathbf{x}_n), \quad (3.6.8)$$

Examples: when electrons are in a potential well  $v(\mathbf{r})$ , the total potential energy operator is  $\hat{V} = \sum_{n=1}^N v(\mathbf{r}_n)$ ; the total kinetic energy is:  $\hat{T} = \sum_{n=1}^N \hat{T}_1(n) = \sum_{n=1}^N \left( -\frac{\hbar^2}{2\mu_e} \frac{\partial^2}{\partial x_n^2} \right)$ .

When the system of  $N$  electrons is in a given Slater wave function  $\Psi_S(\mathbf{x}_1, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \det[\phi_1(\mathbf{x}_1) \dots \phi_N(\mathbf{x}_N)]$ , then using the notation of (3.6.3), we have:

$$\begin{aligned}
& \langle \Psi_S | \hat{O} | \Psi_S \rangle \\
&= \frac{1}{N!} \int \det[\phi_1(\mathbf{x}_1) \dots \phi_N(\mathbf{x}_N)] \hat{O} \det[\phi_1(\mathbf{x}_1) \dots \phi_N(\mathbf{x}_N)] d\mathbf{x}_1 \dots d\mathbf{x}_N \\
&= \frac{1}{N!} \sum_{n=1}^N \sum_{\substack{i_1 \dots i_N \\ j_1 \dots j_N}} (-)^{P_i + P_j} \prod_{\substack{k=1 \\ k \neq n}}^N \langle \phi_{i_k} | \phi_{j_k} \rangle \langle \phi_{i_n} | \hat{O}_1 | \phi_{j_n} \rangle
\end{aligned} \tag{3.6.9}$$

Once again, a massive cancellation of terms happens in the first integral. Inspection shows that both permutations,  $i$  and  $j$ , must be equal otherwise there is always an orbital integral for which the integrals  $\langle \xi_{i_k} | \xi_{j_k} \rangle$  is zero. When the permutations are identical we have:

$$\langle \Psi_S | \hat{O} | \Psi_S \rangle = \frac{1}{N!} \sum_{n=1}^N \sum_{i_1 \dots i_N} \langle \phi_{i_n} | \hat{O}_1 | \phi_{j_n} \rangle = \sum_{n=1}^N \langle \phi_j | \hat{O}_1 | \phi_n \rangle \tag{3.6.10}$$

As an example, let us take the electron density  $\hat{n}(\mathbf{r}) = \sum_{n=1}^N \delta(\mathbf{r} - \mathbf{r}_n)$ . Thus:

$$n(\mathbf{r}) = \sum_{n=1}^N |\phi_n(\mathbf{r}, s)|^2 \tag{3.6.11}$$

**Conclusion:** The matrix element of a 1-particle operator is the sum of its single-electron matrix elements.

## ii. Two-body operators

Consider an operator  $\hat{o}_{12}(\mathbf{x}_1, \mathbf{x}_2)$  which operates on two electrons with spin-coordinates  $\mathbf{x}_1$  and  $\mathbf{x}_2$ . For  $N$  electrons we define the sum of  $\hat{o}_{12}$  on all pairs of electrons

$$\hat{O} = \sum_{n < m = 1}^N \hat{o}_{12}(\mathbf{x}_n, \mathbf{x}_m) = \frac{1}{2} \sum_{n \neq m = 1}^N \hat{o}_{12}(\mathbf{x}_n, \mathbf{x}_m), \tag{3.6.12}$$

Examples: The 2-electron interaction potential is  $u_{12}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\kappa e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}$ . The total interaction energy operator is:  $\hat{U} = \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N u_{12}(\mathbf{r}_i, \mathbf{r}_j)$ .



We compute the expectation value

$$\begin{aligned}
& \langle \Psi_S | \hat{U} | \Psi_S \rangle \\
&= \frac{1}{N!} \int \det[\phi_1(\mathbf{x}_1) \dots \phi_N(\mathbf{x}_N)] \hat{U} \det[\phi_1(\mathbf{x}_1) \dots \phi_N(\mathbf{x}_N)] dx_1 \dots dx_N \\
&= \frac{1}{N!} \frac{1}{2} \sum_{\substack{n,m=1 \\ n \neq m}}^N \sum_{\substack{i_1 \dots i_N \\ j_1 \dots j_N}} (-)^{P_i+P_j} \prod_{\substack{k \neq n \\ k \neq m}}^N \langle \phi_{i_k} | \phi_{j_k} \rangle \langle \phi_{i_m} \phi_{j_m} | u_{12} | \phi_{i_n} \phi_{j_n} \rangle
\end{aligned} \tag{3.6.13}$$

Where we used the notation:

$$\langle \phi_i \phi_j | u_{12} | \phi_{i'} \phi_{j'} \rangle \equiv \int \phi_i(\mathbf{x}) \phi_j(\mathbf{x}') u_{12}(\mathbf{r}, \mathbf{r}') \phi_{i'}(\mathbf{x}) \phi_{j'}(\mathbf{x}') dx dx' \tag{3.6.14}$$

The following symmetry properties hold from the above definition:

$$\begin{aligned}
\langle \phi_i \phi_j | u_{12} | \phi_{i'} \phi_{j'} \rangle &= \langle \phi_{i'} \phi_{j'} | u_{12} | \phi_i \phi_j \rangle = \langle \phi_i \phi_{j'} | u_{12} | \phi_{i'} \phi_j \rangle \\
&= \langle \phi_{i'} \phi_{j'} | u_{12} | \phi_i \phi_j \rangle = \text{etc}
\end{aligned} \tag{3.6.15}$$

For a pair of permutations to contribute to the integral in Eq. (3.6.13), the permutations must either be identical or involve the permutation of a single pair of orbitals. Thus:

$$\langle \Psi_S | \hat{U} | \Psi_S \rangle = \frac{1}{2} \sum_{\substack{n,m=1 \\ n \neq m}}^N [\langle \phi_n \phi_m | u_{12} | \phi_n \phi_m \rangle - \langle \phi_n \phi_m | u_{12} | \phi_m \phi_n \rangle] \tag{3.6.16}$$

## IV. The Hartree-Fock Theory

### A. The Hartree-Fock Energy and Equations

The variational principle says that the lowest expectation value of the electronic Hamiltonian attained by the ground-state. This Hamiltonian, in the non-relativistic approximation, for  $N$  electrons is given by:

$$\hat{H} = \hat{h} + \hat{U} = \sum_{n=1}^N \hat{h}_n + \frac{e^2}{2} \sum_{n \neq m=1}^N \frac{1}{r_{nm}} \quad (4.1.1)$$

Where the “one body” operator is

$$\hat{h} = \sum_{m=1}^N \hat{h}_m = \sum_{m=1}^N -\frac{\hbar^2}{2m_e} \nabla_m^2 + v(\mathbf{r}_m) \quad (4.1.2)$$

And the two-body Coulomb repulsion operator is:

$$\hat{U} = \frac{e^2}{2} \sum_{n \neq m=1}^N \frac{1}{|\mathbf{r}_n - \mathbf{r}_m|} \quad (4.1.3)$$

Given a family of wave functions we can choose the “best” of them by finding that which minimizes the expectation value of  $\hat{H}$ . For the Slater wave functions  $\Psi_S = |\phi_1 \dots \phi_N\rangle$ , the energy to be minimized is:

$$E[\Psi_S] = \langle \Psi_S | \hat{H} | \Psi_S \rangle \equiv E_{HF}[\phi_1 \dots \phi_N] \quad (4.1.4)$$

Because we want the orbitals to be orthonormal, we write a Lagrangian:

$$L[\Psi_S] = L_{HF}[\phi_1 \dots \phi_N] = E_{HF}[\phi_1 \dots \phi_N] - \sum_{i,j=1}^N \mu_{ij} [\langle \phi_i | \phi_j \rangle - \delta_{ij}] \quad (4.1.5)$$

The minimum is attained by:

$$0 = \frac{\delta L_{HF}[\phi_1 \dots \phi_N]}{\delta \phi_i(x)} = \frac{\delta E_{HF}[\phi_1 \dots \phi_N]}{\delta \phi_i(x)} - \sum_{j=1}^N (\mu_{ij} + \mu_{ji}) \phi_j(x) \quad (4.1.6)$$

Let us try a solution with

$$\mu_{ij} = \frac{1}{2} \epsilon_i \delta_{ij} \quad (4.1.7)$$

i.e.:

$$\frac{\delta E_{HF}[\phi_1 \dots \phi_N]}{\delta \phi_i(x)} = \epsilon_i \phi_i(x) \quad (4.1.8)$$

If we can solve this equation, and if the solutions are naturally orthogonal, then we have obtained the necessary conditions for a minimum. Now, we only need to estimate the left hand side of this equation. From the previous work, we know:

$$E_{\text{HF}}[\phi_1 \dots \phi_N] = \langle \Psi_S | \hat{h} | \Psi_S \rangle + \langle \Psi_S | \hat{U} | \Psi_S \rangle \quad (4.1.9)$$

where

$$\langle \Psi_S | \hat{h} | \Psi_S \rangle \equiv h[\phi_1 \dots \phi_N] = \sum_{m=1}^N \langle \phi_m | \hat{h}_1 | \phi_m \rangle \quad (4.1.10)$$

Is the "one body energy" defined the as the sum of kinetic energy and "external" energy (i.e. energy due to the frozen nuclei):

$$\hat{h}_1 = -\frac{\hbar^2}{2\mu_e} \nabla^2 + v(\mathbf{r}) \quad (4.1.11)$$

Furthermore, the 2-electron energy can be written as:

$$\begin{aligned} \langle \Psi_S | \hat{U} | \Psi_S \rangle &= \frac{e^2}{2} \sum_{\substack{n,m=1 \\ n \neq m}}^N \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_n \phi_m \right\rangle - \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_m \phi_n \right\rangle \\ &\equiv J[\phi_1 \dots \phi_N] + K[\phi_1 \dots \phi_N] \end{aligned} \quad (4.1.12)$$

Where the direct or Hartree energy is:

$$J[\phi_1 \dots \phi_N] = \frac{e^2}{2} \sum_{n,m=1}^N \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_n \phi_m \right\rangle, \quad (4.1.13)$$

and the energy exchange is

$$K[\phi_1 \dots \phi_N] = -\frac{e^2}{2} \sum_{n,m=1}^N \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_m \phi_n \right\rangle. \quad (4.1.14)$$

The direct energy is numerically equal to the Hartree energy  $E_H[n]$  which is a functional of the Slater wave function electron density  $n(\mathbf{r}) = \langle \Psi_S | \hat{n}(\mathbf{r}) | \Psi_S \rangle = \sum_{m=1}^{N_e} |\phi_m(\mathbf{r}, s)|^2$ :

$$E_H[n] = \frac{e^2}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (4.1.15)$$

The exchange energy is numerically equal to the “exchange energy”  $E_X[\rho]$  which is a functional of the *density matrix* defined by

$$\rho(\mathbf{x}, \mathbf{x}') = \sum_{m=1}^N \phi_m(\mathbf{x}) \phi_m(\mathbf{x}') \quad (4.1.16)$$

As:

$$E_X[\rho] = -\frac{e^2}{2} \iint \frac{|\rho(\mathbf{x}, \mathbf{x}')|^2}{|\mathbf{r} - \mathbf{r}'|} dx dx' \quad (4.1.17)$$

Notice that:  $\int \rho(\mathbf{x}, \mathbf{x}) ds = n(\mathbf{r})$ . The density matrix is *idempotent*:

$$\int \rho(\mathbf{x}, \mathbf{x}'') \rho(\mathbf{x}'', \mathbf{x}') d^3x'' = \rho(\mathbf{x}, \mathbf{x}') \quad (4.1.18)$$

This result shows that the DM is a projection operator, projecting onto the space of orbitals which defines the Slater wave function.

With the direct and exchange energies we also define their functional derivatives:

$$\begin{aligned}
\frac{\delta J}{\delta \phi_i(\mathbf{x})} &= \frac{1}{2} \sum_{n,m} \frac{\delta}{\delta \phi_i(\mathbf{x})} \left\langle \phi_n \phi_m \left| \frac{e^2}{r_{12}} \right| \phi_n \phi_m \right\rangle \\
&= \frac{e^2}{2} \sum_{n,m} \frac{\delta}{\delta \phi_i(\mathbf{x})} \iint \frac{\phi_n(\mathbf{x}_1)^2 \phi_m(\mathbf{x}_2)^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \frac{e^2}{2} \sum_{n,m} \left( 2\delta_{in} \phi_n(\mathbf{x}) \int \frac{\phi_m(\mathbf{x}_2)^2}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{x}_2 \right. \\
&\quad \left. + 2\delta_{im} \phi_m(\mathbf{x}) \int \frac{\phi_n(\mathbf{x}_1)^2}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{x}_1 \right) \\
&= 2e^2 \left( \int \frac{n(\mathbf{r}_2)}{|\mathbf{r} - \mathbf{r}_2|} d^3 r_2 \right) \phi_i(\mathbf{x}) \equiv 2v_H(\mathbf{r}) \phi_i(\mathbf{x}) \\
&\equiv 2\hat{v}_H \phi_i(\mathbf{x})
\end{aligned} \tag{4.1.19}$$

Where in the last line we defined the Hartree potential:

$$v_H(\mathbf{r}) = e^2 \sum_{m=1}^N \int \frac{|\phi_m(\mathbf{x})|^2}{|\mathbf{r} - \mathbf{r}'|} d^3 r' = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \tag{4.1.20}$$

Then the direct energy can also be written as a functional of the orbitals:

$$J[\phi_1 \dots \phi_N] = \frac{1}{2} \sum_{m=1}^N \langle \phi_m | \hat{v}_H | \phi_m \rangle = \frac{1}{2} \int v_H(\mathbf{r}) n(\mathbf{r}) d^3 r = \frac{1}{2} \langle \Psi_S | \hat{V}_H | \Psi_S \rangle \tag{4.1.21}$$

where  $\hat{V}_H = \sum_{n=1}^N v_H(\mathbf{r}_n)$  is the total Hartree potential. A similar treatment exists for the exchange energy functional derivative:

$$\begin{aligned}
\frac{\delta K}{\delta \phi_i(\mathbf{x})} &= -\frac{e^2}{2} \sum_{n,m} \frac{\delta}{\delta \phi_i(\mathbf{x})} \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_m \phi_n \right\rangle \\
&= -\frac{e^2}{2} \sum_{n,m} \left( 2\delta_{in} \phi_m(\mathbf{x}) \int \frac{\phi_m(\mathbf{x}_2) \phi_n(\mathbf{x}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{x}_2 \right. \\
&\quad \left. + 2\delta_{im} \phi_n(\mathbf{x}) \int \frac{\phi_n(\mathbf{x}_2) \phi_m(\mathbf{x}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{x}_2 \right) = \\
&= -2e^2 \sum_n \int \frac{\phi_n(\mathbf{x}_2) \phi_i(\mathbf{x}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{r}_2 \phi_n(\mathbf{x}) = 2\hat{K} \phi_i(\mathbf{x})
\end{aligned} \tag{4.1.22}$$

Where the last equality is based on a definition of a one-particle exchange operator:

$$\begin{aligned}\hat{K}_1\psi(\mathbf{x}) &\equiv -e^2 \sum_{m=1}^N \phi_m(\mathbf{x}) \int \frac{\phi_m(\mathbf{x}')\psi(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}' \\ &= -e^2 \int \frac{\rho(\mathbf{x},\mathbf{x}')\psi(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}'\end{aligned}\quad (4.1.23)$$

Then the exchange energy is written as a functional of the orbitals:

$$K[\phi_1 \dots \phi_N] = \frac{1}{2} \sum_{n=1}^N \langle \phi_n | \hat{K}_1 | \phi_n \rangle = \frac{1}{2} \langle \Psi_S | \hat{Q}_H | \Psi_S \rangle \quad (4.1.24)$$

where  $\hat{Q}_H = \sum_{n=1}^N \hat{K}_n$  is the total exchange operator. The other functional derivatives needed are:

$$\frac{\delta}{\delta\phi_i(\mathbf{x})} \int \phi_n(\mathbf{x}') \nabla^2 \phi_n(\mathbf{x}') d^3x' = 2\nabla^2 \phi_n(\mathbf{x}) \delta_{in} \quad (4.1.25)$$

and

$$\frac{\delta}{\delta\phi_i(\mathbf{x})} \int \phi_n(\mathbf{x}')^2 v(\mathbf{r}') d\mathbf{x}' = 2v(\mathbf{r}) \phi_n(\mathbf{x}) \delta_{in} \quad (4.1.26)$$

Thus:

$$\frac{\delta}{\delta\phi_i(\mathbf{x})} \int \phi_n(\mathbf{x}') \hat{h}(\mathbf{r}') \phi_n(\mathbf{x}') d\mathbf{x}' = 2\hat{h}\phi_n(\mathbf{x}) \delta_{in} \quad (4.1.27)$$

Plugging all these terms into Eq. (4.1.7), we obtain the **Hartree-Fock equations**:

$$\hat{F}\phi_i(\mathbf{x}) = \epsilon_i\phi_i(\mathbf{x}) \quad (4.1.28)$$

where:

$$\hat{F} = \hat{h} + v_H(\mathbf{r}) + \hat{K}_1 \quad (4.1.29)$$

Equations (4.1.28) seem very much like 1-electron eigenvalue equations of the Schrödinger equation, except that instead of a regular Hamiltonian, we have a Fock operator including the non-local exchange. We showed in the exercise that  $\hat{K}$  is Hermitean and thus so is  $\hat{F}$ , i.e.  $\langle \phi | \hat{F} | \psi \rangle = \langle \psi | \hat{F} | \phi \rangle^*$ . We can thus choose the orbital solutions of the eigenvalue equation (4.1.28) orthonormal. This shows that the choice Eq. (4.1.7) is indeed acceptable.

Because  $v_H(\mathbf{r})$  and  $\hat{K}$  themselves depend on  $\phi_m$ , the Hartree-Fock equations are fundamentally different from the Schrödinger Equation: they are *nonlinear equations*.

Now that the sum of orbital energies is:

$$\begin{aligned} \sum_{m=1}^N \epsilon_m &= \sum_{m=1}^N \langle \phi_m | \hat{F} | \phi_m \rangle = \sum_{m=1}^N [\langle \phi_m | \hat{h}_1 | \phi_m \rangle + \langle \phi_m | \hat{v}_H + \hat{K}_1 | \phi_m \rangle] \\ &= h[\phi_1, \dots, \phi_N] + 2J[\phi_1, \dots, \phi_N] + 2K[\phi_1, \dots, \phi_N] \end{aligned} \quad (4.1.30)$$

This shows that the orbital sum is not equal to the energy of the wave function, since it involves double counting of the direct and exchange energies. The HF energy is thus:

$$E_{HF}[\phi_1, \dots, \phi_N] = \sum_{m=1}^N \epsilon_m - (J[\phi_1, \dots, \phi_N] + K[\phi_1, \dots, \phi_N]) \quad (4.1.31)$$

### ***B. Restricted closed-shell Hartree-Fock***

For molecules with even number  $N$  of electrons in a spin-singlet state, we can impose the following structure on the Slater wave function. We can assume that the  $2N$  spin-orbitals come in pairs:  $\phi_{2j-1}(\mathbf{x}) = \psi_j(\mathbf{r})\alpha(\omega)$  and  $\phi_{2j}(\mathbf{x}) = \psi_j(\mathbf{r})\beta(\omega)$ . Thus each orbital  $\psi_j$ ,  $j = 1, \dots, N/2$  is “doubly occupied” by electrons of both spins. By imposing this constraint we obtain the “restricted” Hartree-Fock ground state. It will sometime be of higher energy than the fully

unrestricted case. However, the wave function has a well-defined spin which may be advantageous in some applications.

The Restricted Hartree-Fock (RHF) energy in the closed shell case remain essentially the same except for counting business. We can formulate all expressions using only the spatial orbitals. Indeed, the RHF energy is given by:

$$E_{RHF}[\psi_1, \dots, \psi_{N/2}] = 2 \sum_{m=1}^{N/2} \langle \psi_m | \hat{h} | \psi_m \rangle + 4J[\psi_1, \dots, \psi_{N/2}] + 2K[\psi_1, \dots, \psi_{N/2}] \quad (4.2.1)$$

Where  $J[\psi_1, \dots, \psi_{N/2}]$  and  $K[\psi_1, \dots, \psi_{N/2}]$  are the orbital functionals defined in Eqs. (4.1.13) and (4.1.14) respectively. The reason for multiplying the one body part by two is evident: each orbitals is double occupied so has double contribution. The direct part is multiplied by 4 since the density is multiplied by two and the direct part depends on the density multiplied by itself. Finally, the exchange part is multiplied by 2 and not 4 since only  $\alpha - \alpha$  and  $\beta - \beta$  contribute, while  $\alpha - \beta$  and  $\beta - \alpha$  do not (so only half the contribution of direct).

The RHF equations are then:

$$\left( -\frac{\hbar^2}{2\mu_e} \nabla^2 + v(\mathbf{r}) + v_H(\mathbf{r}) \right) \psi_m(\mathbf{r}) + \hat{K}_1 \psi_m(\mathbf{r}) = \epsilon_m \psi_m(\mathbf{r}) \quad (4.2.2)$$

Where:

$$v_H(\mathbf{r}) = 2e^2 \sum_{m=1}^{N/2} \int \frac{|\psi_m(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (RHF) \quad (4.2.3)$$

and



$$\hat{K}_1\psi(\mathbf{r}) \equiv -e^2 \sum_{m=1}^{N/2} \psi_m(\mathbf{r}) \int \frac{\psi_m(\mathbf{r}')\psi(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \quad (RHF) \quad (4.2.4)$$

Note that the Hartree interaction is between each electron and all other electrons regardless of their spin while the exchange interaction involves each electrons with all other electrons *of the same spin*.

Example: The H<sub>2</sub> molecule

We apply the RHF theory for the H<sub>2</sub>, having a pair of electrons. The 2-electron wave function includes just one spatial orbital populated by spin-paired electrons:

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = \begin{vmatrix} \psi(\mathbf{r}_1) & \psi(\mathbf{r}_2) \\ \bar{\psi}(\mathbf{r}_1) & \bar{\psi}(\mathbf{r}_2) \end{vmatrix} = \psi(\mathbf{r}_1)\psi(\mathbf{r}_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)] \quad (4.2.5)$$

Since there is only one orbital the exchange K and direct J involve just one and the same integral. Thus the RHF energy is in this case:  $E_{RHF}[\psi] = 2\langle\psi_m|\hat{h}|\psi_m\rangle + 2J[\psi]$  and the RHF equation is

$$\left( -\frac{\hbar^2}{2\mu_e}\nabla^2 + v(\mathbf{r}) + \frac{1}{2}v_H(\mathbf{r}) \right) \psi(\mathbf{r}) = \epsilon\psi(\mathbf{r}) \quad (4.2.6)$$

where  $v(\mathbf{r}) = -\frac{e^2}{|\mathbf{r}-\mathbf{R}_A|} - \frac{e^2}{|\mathbf{r}-\mathbf{R}_B|}$  and  $v_H(\mathbf{r}) = 2e^2 \int \frac{|\psi(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d^3r'$ . The effect of exchange here is to annihilate the Coulomb repulsion of the  $\alpha$  ( $\beta$ ) electron with itself, leaving only the interaction of the  $\alpha - \beta$  electrons.

This RHF approach works nicely for the case that the distance between the nuclei  $|\mathbf{R}_A - \mathbf{R}_B|$  is close to the typical bond length of H<sub>2</sub> (which is close to  $1.4a_0$ ). The energy of the molecule at this configuration can be calculated numerically and results in  $E_{RHF} = -1.134E_h$ . Compared to the energy of 2 H atoms ( $-1E_h$ ) this results indicates that the atomization energy of H<sub>2</sub> is  $0.134E_h = 3.65\text{eV}$ . The atomization energy based on experimental results is

about 4.75eV. This shows that the RHF approximation does not give high quality atomization energies, since a deficit of 1 eV is very substantial in Chemical energy terms.

A more severe problem arises when we place the two nuclei far from each other. We expect the resulting energy and wave function to resemble that of two separated H atoms. I.e the exact wave-function should approach something like  $\Psi = |\psi_{1s}(\mathbf{r}_1 - \mathbf{R}_A)\bar{\psi}_{1s}(\mathbf{r}_2 - \mathbf{R}_B)| - |\psi_{1s}(\mathbf{r}_2 - \mathbf{R}_A)\bar{\psi}_{1s}(\mathbf{r}_1 - \mathbf{R}_B)|$ . This form however is not supported by the RHF ansatz of Eq. (4.2.5). Indeed, if we think of the solution of the RHF equation as being approximately given by  $\psi \approx \psi_A(\mathbf{r}) + \psi_B(\mathbf{r})$ , where  $\psi_X(\mathbf{r}) = \psi(\mathbf{r} - \mathbf{R}_X)$ ,  $X = A, B$ , then the RHF wave function is:

$$\begin{aligned} \Phi(\mathbf{x}_1, \mathbf{x}_2) &\approx \left| (\psi_A(\mathbf{r}_1) + \psi_B(\mathbf{r}_1)) (\bar{\psi}_A(\mathbf{r}_2) + \bar{\psi}_B(\mathbf{r}_2)) \right| \\ &= \left[ |\psi_A(\mathbf{r}_1)\bar{\psi}_A(\mathbf{r}_2)| + |\psi_B(\mathbf{r}_1)\bar{\psi}_B(\mathbf{r}_2)| \right] \\ &\quad + \left[ |\psi_A(\mathbf{r}_1)\bar{\psi}_B(\mathbf{r}_2)| + |\psi_B(\mathbf{r}_1)\bar{\psi}_A(\mathbf{r}_2)| \right] \end{aligned} \quad (4.2.7)$$

The first term is an ionic term, where both electrons are on the same atom (either A or B) while the second term places one electron on each atom – “neutral” term. The problem of RHF theory is the ionic term. It may be important when the atoms are close but it should go to zero when they are far.

### ***C. Atomic Orbitals and Gaussian Basis sets***

Where do we get "good" basis functions? What is "good"?

We want a small basis that can still describe the electrons. On natural source are the atomic orbitals of the atoms. These are of the form resembling exponentials time polynomials. Thus, one choice is:

$$\chi_{lm}^A(\mathbf{r}) = \xi_{lm}^{\zeta_A}(\mathbf{r} - \mathbf{R}_A) \quad (4.2.8)$$

Where:

$$\xi_{lm}^{\zeta}(\mathbf{r}) = r^l e^{-\zeta r} Y_{lm}(\theta, \phi) \quad (4.2.9)$$

Where  $Y_{lm}$  are the spherical harmonics. One can also take appropriate combinations of these functions to make them all real. These functions have a desired analytical property: their  $l$ 's derivative exhibits a cusp of the correct order and structure at  $\mathbf{r} = \mathbf{R}_A$ . There exist analytical formula for doing the overlap and "one-body" integrals. But there are no convenient formulae for the 2-body integrals, although some progress was made in recent years (see articles by Handy).

A more convenient, although less natural choice (no cusp). Is the use of Gaussian functions, for example:

$$\xi_{lm}^{\zeta}(\mathbf{r}) = r^l \left( \sum_{\mu} a_{\mu,l} e^{-\beta_{\mu} r^2} \right) Y_{lm}(\theta, \phi) \quad (4.2.10)$$

Where  $a_{l,\mu}$  are called "contraction coefficients. These are chosen so that  $(\sum_{\mu} a_{\mu,l} e^{-\beta_{\mu} r^2})$  resembles  $e^{-\zeta r}$ . With Gaussian functions very effective and rapid algorithms were published allowing extremely fast 2-electron integrals.

#### ***D. Variational-Algebraic approach Hartree-Fock***

We have seen that the Hartree-Fock equations can be derived by searching for that the most general Slater wave function that minimizes the Hartree-Fock functional. However implementing a solution to such equations is usually very difficult, if not impossible in practice. A more practical approach, that keeps the spirit of the Hartree\_Fock approach was developed by Roothan and Hall. In this approach we find the optimal Slater wave function of orbitals constrained to lie in a finite dimensional vector space spanned by *basis*

functions, usually called atomic orbitals (although, up to a point, we need not assume this)  $\chi_\sigma(\mathbf{x})$ ,  $\sigma = 1, \dots, M$ . Thus, a set of  $N$  molecular orbitals  $\phi_n(\mathbf{x})$  ( $n = 1, \dots, N$ ) in a determinant  $\Psi_S[C]$  of this form must all be of the following form:

$$\phi_n(\mathbf{x}) = \sum_{\sigma=1}^M \chi_\sigma(\mathbf{x}) C_{\sigma n} \quad (4.3.1)$$

The  $C_{\sigma n}$  coefficients form an  $M \times N$  matrix, called the MO coefficient matrix for the determinant. Note that for this to make sense we must demand  $M > N$ . The HF energy functional now becomes a function of these coefficients  $C_{\sigma n}$ . The constraint that the MO's are orthonormal,  $\langle \phi_n | \phi_m \rangle = \delta_{nm}$ , becomes:

$$\delta_{nm} = \left\langle \sum_{\sigma=1}^M \chi_\sigma C_{\sigma n} \left| \sum_{\sigma'=1}^M \chi_{\sigma'} C_{\sigma' m} \right. \right\rangle = (C^T S C)_{nm} \quad (4.3.2)$$

Where we use matrix algebra notation and the  $M \times M$  matrix  $S$  is defined by:

$$S_{\sigma\sigma'} = \langle \chi_\sigma | \chi_{\sigma'} \rangle \quad (4.3.3)$$

Thus, the orthonormality condition is:

$$C^T S C = I_N \quad (4.3.4)$$

Where  $I_N$  is the  $N \times N$  unit matrix. Let us now derive an expression for the expectation value of a one-body operator in a Slater wave function of these MO's, by Eq. (3.6.10):

$$\langle \Psi_S | \hat{O} | \Psi_S \rangle = \sum_{n=1}^N \langle \phi_n | \hat{O}_1 | \phi_n \rangle = \text{Tr}[C^T O C] \quad (4.3.5)$$

Where  $O$  is the  $M \times M$  matrix in the AO basis:

$$O_{\sigma\sigma'} = \langle \chi_\sigma | \hat{O}_1 | \chi_{\sigma'} \rangle \quad (4.3.6)$$

It is customary to define the  $M \times M$  density matrix:

$$P = CC^T \quad (4.3.7)$$

And with it we can write:

$$\langle \Psi_S | \hat{O} | \Psi_S \rangle = Tr[CC^T O C] = Tr[PO] \quad (4.3.8)$$

In the last step we used the fact that the trace of the product of two matrices is invariant to their order of multiplication:

$$Tr[AB] = \sum_{\sigma=1}^M (AB)_{\sigma\sigma} = \sum_{\sigma=1}^M \sum_{n=1}^N A_{\sigma n} B_{n\sigma} = \sum_{n=1}^N \sum_{\sigma=1}^M B_{n\sigma} A_{\sigma n} = Tr[BA] \quad (4.3.9)$$

Notice that the DM has the generalized idempotency property:

$$PSP = CC^T S C C^T = CC^T = P \quad (4.3.10)$$

One can see that  $P$  is a symmetric matrix. Furthermore, one can see that it is positive semi-definite, i.e. for any vector:  $v^T P v = v^T C C^T v = (C^T v)^T (C^T v) \geq 0$ .

Furthermore:

$$Tr[PS] = Tr[CC^T S] = Tr[C^T S C] = N$$

This last step is a result of Eq. (4.3.4). Finally, the 2-body operator, by Eq. (4.1.12), we need the direct and exchange. We use:

$$[kl|mn] = \sum_{\sigma, \sigma', \tau, \tau'=1}^N C_{k\sigma}^T C_{\sigma'l} [\sigma\sigma' | \tau\tau'] C_{m\tau}^T C_{\tau'n}, \quad (4.3.11)$$

Then:

$$\begin{aligned} J[C] &= \frac{e^2}{2} C_{n\sigma}^T C_{\sigma'n} [\sigma\sigma' | \tau\tau'] C_{m\tau}^T C_{\tau'm} \\ K[C] &= -\frac{e^2}{2} C_{n\sigma}^T C_{\sigma'm} [\sigma\sigma' | \tau\tau'] C_{m\tau}^T C_{\tau'n}, \end{aligned} \quad (4.3.12)$$

Where we use the convention that all repeated indices are summed over: latin indices are summed between 1 and  $N$  and greek indices between 1 and  $M$ . The summations on  $n, m$  can be done first and we obtain:

$$\begin{aligned} J[C] &= \frac{e^2}{2} P_{\sigma'\sigma} [\sigma\sigma' | \tau\tau'] P_{\tau\tau'} \\ K[C] &= -\frac{e^2}{2} P_{\sigma\tau'} [\sigma\sigma' | \tau\tau'] P_{\sigma'\tau} \end{aligned} \quad (4.3.13)$$

The last expression can be reindexed (assuming  $P$  is a symmetric matrix) as:

$$K[C] = -\frac{e^2}{2} P_{\sigma'\sigma} [\sigma\tau' | \tau\sigma'] P_{\tau\tau'} \quad (4.3.14)$$

Thus:

$$J[C] + K[C] = \frac{e^2}{2} P_{\sigma'\sigma} ([\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \tau\sigma']) P_{\tau\tau'} \quad (4.3.15)$$

Finally, defining:

$$V_{IJ} \equiv V_{\sigma\sigma', \tau\tau'} = [\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \tau\sigma'] = [\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \sigma'\tau] \quad (4.3.16)$$

Using the double indexing  $I \equiv (\sigma\sigma')$  and  $J = (\sigma'\sigma''')$ . Note that  $V_{IJ} = V_{JI}$  as can be seen from:

$$\begin{aligned} V_{IJ} = V_{\sigma\sigma', \tau\tau'} &= [\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \sigma'\tau] = [\tau\tau' | \sigma\sigma'] - [\tau\sigma' | \tau'\sigma] = V_{\tau\tau' \sigma\sigma'} \\ &= V_{JI} \end{aligned} \quad (4.3.17)$$

We may thus write:

$$J[C] + K[C] = \frac{e^2}{2} P^t V P \quad (4.3.18)$$

Where now, we consider  $P$  not as a  $M \times M$  matrix, but as a column vector of  $M^2$  elements.  $P^t$  is the corresponding row-vector. Similarly,  $V$  is not a  $M \times M \times M \times M$  tensor but as  $M^2 \times M^2$  matrix.

The HF energy is thus compactly written as:

$$E_{HF}[P] = P^t h + \frac{1}{2} P^t V P \quad (4.3.19)$$

Note, that our unknown variable is now the DM  $P$ . We want to minimize  $E_{HF}$  with respect to  $P$ , however, we need to impose 2 types of constraints. First, we need to specify the subject to the constraints:

$$\begin{aligned} Tr[PS] &= N \\ G &= P S P S - P S = 0 \end{aligned} \quad (4.3.20)$$

In order to minimize the energy we introduce the Lagrangian:

$$L[P] = E_{HF}[P] - \mu(P^t S - N) - \Lambda^t G \quad (4.3.21)$$

The number  $\mu$  and the  $M \times M$  matrix  $\Lambda$  are Lagrange multipliers. The algebraic Hartree-Fock equations are now  $\frac{\delta L}{\delta P_{\sigma\tau}} = 0$ . In order to obtain working expressions we derive:

$$\frac{\delta E_{HF}}{\delta P_I} = h_I + \frac{1}{2} (P_J V_{JI} + V_{IJ} P_J) = h_I + V_{IJ} P_J \quad (4.3.22)$$

Where again, we use the convention that when a super index appears twice we sum over it. This can be written more compactly as:

$$\frac{\delta E_{HF}}{\delta P} = h + V P \equiv F \quad (4.3.23)$$

This gradient is what we call the ‘‘Fock matrix’’  $F$ . In our present notation  $F$  is a  $M^2$  vector  $F_I$  with index  $I$ . But soon we will consider it as a matrix with two indices  $F_{\sigma\sigma'}$ . The constraints can easily be derived in a similar way, leading to the following Lagrangian gradient:

$$\frac{\delta L}{\delta P} = F - \mu S - (S P S \Lambda + \Lambda P S - S \Lambda) \quad (4.3.24)$$

It is easy to convince one’s self that  $F$  is always a symmetric matrix, for any  $P$ .

The condition for minimum is:

$$F - \mu S - (SPS\Lambda + S\Lambda PS - S\Lambda) = 0 \quad (4.3.25)$$

Multiplying by  $SP$  from the left we find:

$$SPF - \mu SPS - SPS\Lambda PS = 0 \quad (4.3.26)$$

Multiplying by  $PS$  from the right we find:

$$FPS - \mu SPS - SPS\Lambda PS = 0 \quad (4.3.27)$$

Subtracting, we obtain:

$$SPF - FPS = 0 \quad (4.3.28)$$

The set of equations that need to be solved simultaneously is:

$$\begin{aligned} SPF - FPS = 0 & & F = h + VP \\ PSP = P & & Tr[PS] = N \end{aligned} \quad (4.3.29)$$

One practical way of doing this is to go back to the matrix  $C$ . In terms of these, the equations become:

$$SCC^T F - FCC^T S = 0 \quad F = h + VCC^T \quad C^T SC = I_N \quad (4.3.30)$$

These equations can all be met if we demand that:

$$FC = SCE \quad F = h + VCC^T \quad (4.3.31)$$

Where  $E$  is a  $M \times M$  diagonal matrix. Indeed, from this equation we also have, from the symmetry of  $F$  and  $S$ :  $C^T F = EC^T S$ . Left-multiplying by  $SC$  we find  $SCC^T F = SCEC^T S$  and using the first equation in (4.3.31) on the right hand side we obtain the first equation in (4.3.30). Furthermore, multiplying the first equation in (4.3.31) by  $C^T$  we find:  $C^T FC = C^T SCE$ . On the left we replace  $C^T F$  by  $EC^T S$  and obtain:

$$[E, C^T SC] = 0$$



We find that  $C^TSC$  is commutative with a diagonal matrix. If no two elements on the diagonal of  $E$  are equal then  $C^TSC$  is diagonal. We know that the diagonal entries must be positive since  $S$  is positive definite. Furthermore, we can choose the norm of the columns of  $C$  so that all diagonal elements of  $C^TSC$  are equal to 1. In this case then  $C^TSC = I$ . When there are several elements on the diagonal of  $E$  which are exactly equal, then one can take linear combinations of the corresponding columns of the C-matrix, without disturbing their eigenstatishness. Once can then always create a situation which again allows  $C^TSC = I$ . We thus find that the procedure of finding the generalized eigenstates and eigenvalues of  $F$  is indeed a procedure for finding the minimum.

Thus Eq. (4.3.31) is the algebraic Hartree Fock equation. In actual calculations, it is very common that programs solve self consistently the algebraic HF equation. This procedure is appropriate for small to medium sized systems. But for larger system it may be beneficial to directly the minimize of the Lagrangian, using the gradient in Eq. (4.3.24). Of course, iterations are still needed because a search must be made for the Lagrange multiplier  $\Lambda$ .

### ***E. The Algebraic Density Matrix and Charge Analysis***

We have seen that the density matrix is defined by the relation  $P = CC^T$  where  $C_{an}$  is the coefficient of the AO  $\chi_\alpha(\mathbf{x})$  in the expansion of the MO  $\phi_n(\mathbf{x})$ . The relation of  $P$  the the real space density matrix defined in XXX is (we use the convention that repeated roman indices are summed from 1 to  $N$  and repeated greek indices are summed from 1 to  $M$ ):

$$\rho(\mathbf{x}, \mathbf{x}') = \phi_n(\mathbf{x})\phi_n(\mathbf{x}') = C_{an}\chi_\alpha(\mathbf{x})\chi_\beta(\mathbf{x}')C_{\beta n} = P_{\alpha\beta}\chi_\alpha(\mathbf{x})\chi_\beta(\mathbf{x}')$$

Thus we see that  $P_{\alpha\beta}$  determines  $\rho$ . Hence the common name for these two quantities. Note also that  $P_{\alpha\beta}$  determines the density  $n(\mathbf{r})$  since that is easily obtained by place  $\mathbf{x} = \mathbf{x}'$  in the DM and integrating over spin:

$$n(\mathbf{r}) = \lim_{\eta \rightarrow 0^+} P_{\alpha\beta} S_{\alpha\beta} \frac{\chi_\alpha(\mathbf{r})\chi_\beta(\mathbf{r})}{S_{\alpha\beta} + i\eta} = q_\alpha n_\alpha(\mathbf{r}) + \sum_{\alpha \neq \beta} q_{\alpha\beta} n_{\alpha\beta}(\mathbf{r})$$

Where:

$$n_\alpha(\mathbf{r}) = \frac{\chi_\alpha(\mathbf{r})^2}{S_{\alpha\alpha}}$$

Are the “atomic electron number densities” (each integrates to 1) and

$$n_{\alpha\beta}(\mathbf{r}) = \frac{\chi_\alpha(\mathbf{r})\chi_\beta(\mathbf{r})}{S_{\alpha\beta} + i\eta}$$

Is the bond electron number density (again, integrating to 1, or zero). The atomic charges are then  $q_\alpha = P_{\alpha\alpha} S_{\alpha\alpha}$  and the bond charges are  $q_{\alpha\beta} = P_{\alpha\beta} S_{\alpha\beta}$ .

This form of charge analysis is very popular and allows to obtain “intuitive” pictures for the charge distribution in the molecule. While useful to many, the user should be warned that this analysis is “basis-form” dependent. What we mean by this is that if we take different linear combinations of the *same* set of basis functions (i.e. we stay in the same Hilbert space), our charge analysis will yield totally different results. This is because when we take linear combinations:  $\chi' = \mathbf{T}\chi$  then the density matrix changes by  $P' = \mathbf{T}^T P \mathbf{T}$ . Thus in general the charges on each atom can change by this procedure.

Note that when one integrates over  $\mathbf{r}$ , on the left hand one gets  $N$ . On the right hand the first gives (assuming the basis functions are normalized  $P_{\alpha\alpha}$

## ***F. Solving the Hartree-Fock Equations***

A plausible “algorithm” for solving the Hartree-Fock equations is as follows:

1. Guess  $\phi_m(\mathbf{r})$ ,  $m = 1, \dots, N_e$ .
2. Build  $v_H(\mathbf{r})$  (Eq.(4.1.20)),  $\hat{K}$  (Eq.(4.1.23)) thus determining  $\hat{F}$  (eq. (4.1.29)).
3. Solve the eigenvalue equations (Eq. (4.1.28) to get a new set of orbitals corresponding to the lowest energy orbitals.
4. Redo from step 2 using the new orbitals, until you converge – i.e. until the orbitals change no more.

While this algorithm seems reasonable, in practice it rarely converges. There are several ways to make an algorithm “practical”.

### **i. Direct inversion in iterative space (DIIS)**

This method, devised by Pulay (P. Pulay, Chem. Phys. Lett. 73, 393 (1980)) is designed to speed up the convergence. Suppose the iterative process has produced  $M$  iterants  $\vec{v}_m$   $m = 1, \dots, M$  (Fockians, density matrices or sets of  $N_e$  orbitals). We can define residuals by:

$$\delta\vec{v}_m = \vec{v}_m - \vec{v}_{m-1} \quad (4.4.1)$$

We want to produce a new iterant by interpolation:

$$\vec{v} = \sum_{m=1}^M w_m \vec{v}_m \quad (4.4.2)$$

where  $w_m$  are the weights and they sum to unity:

$$\sum_{m=1}^M w_m = 1 \quad (4.4.3)$$

These weights are obtained by minimizing the residual, assuming linearity:

$$\delta\vec{v} = \sum_{m=1}^M w_m \delta\vec{v}_m \quad (4.4.4)$$

The function to be minimized is:

$$\begin{aligned}
 J[w] &= \frac{1}{2} \delta \vec{v}^T \delta \vec{v} - \lambda \left[ \sum_{m=1}^M w_m - 1 \right] \\
 &= \frac{1}{2} \sum_{m,m'=1}^M w_m B_{m,m'} w_{m'} - \lambda \left[ \sum_{m=1}^M w_m - 1 \right]
 \end{aligned} \tag{4.4.5}$$

Where:

$$B_{m,m'} = \delta \vec{v}_m^T \delta \vec{v}_{m'} \tag{4.4.6}$$

Differentiating with respect to  $w_k$  gives:

$$0 = \frac{\partial J}{\partial w_k} = B_{km} w_m - \lambda w_k \tag{4.4.7}$$

The solution of these equations, together with the constraints Eq. (4.4.3) gives:

$$\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1M} & -1 \\ B_{21} & B_{22} & \cdots & B_{2M} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{M1} & B_{M2} & \cdots & B_{MM} & -1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{4.4.8}$$

The solution of this equation gives the desired weights. The use of this algorithm can be done in the following way:

1. Get new  $\vec{v}$  as output from the iterative procedure. Add it to the list i.e. designate it as  $\vec{v}_M$
2. Find weights from which get interpolant  $\vec{v}' = \sum_{m=1}^M w_m \vec{v}_m$ .
3. Use  $\vec{v}'$  as input to the iterative procedure and redo from 1.

## ii. Direct Minimization

Sometimes the DIIS procedure is not effective and other methods are tried. One of the most useful methods is to use numerical minimization techniques, such as the conjugate gradients algorithm to directly minimize the energy of

the Slater wave function under the constraints. This method is especially useful when the molecule being studied is very large. Special tricks are used to formulate the minimization problem as an “unconstrained minimization” (see for example, Nunes et al, Phys. Rev. B 17611, 50 (1994)).

### ***G. Performance of the Hartree-Fock approximation***

We examine the performance of Hartree-Fock approximation on, for example Formaldehyde. There are 2 sources of error. One is in the application, since we use finite basis sets. Then there is the intrinsic error.

In the table below, we see the prediction of various properties of formaldehyde, calculated with increasing quality of basis set and compared to experiment.

Basis	R(CO)Å	R(CH)Å	A(O-C-H)	Energy(au)
sto-3g	1.2169	1.1014	122.73	-112.35435
3-21g	1.2071	1.0833	122.51	-113.22182
sto-6g	1.2163	1.0981	122.61	-113.44078
6-31g	1.2103	1.0816	121.69	-113.80837
D95	1.2170	1.0843	121.57	-113.83071
D95v*	1.1887	1.0935	121.96	-113.89173
6-311g**	1.1787	1.0949	122.09	-113.89915
6-311++g**	1.1797	1.0943	121.97	-113.90287
apvtz	1.1786	1.0927	121.94	-113.91534
experimental	1.210	1.1020	121.1	

We see Hartree-Fock converges when basis set quality increases. However the converged quantity deviates somewhat from experimental values.

This deviance exists because Hartree-Fock theory is only an approximation. What it assumes is that the electrons act as if they are independent particles (since it imposes a single determinant). The real ground-state is composed from a huge series of determinants. The "independent" particles interact with

the mean field of all other particles, while in essence each electrons has to interact with each other electron, trying as much as possible to avoid it, without paying too much in kinetic or electron-nuclear potential energy.

### ***H. Beyond Hartree-Fock***

The Hartree-Fock method is very successful, since it typically accounts for over 99% of the electronic energy of molecules. Yet it is not accurate enough for most applications in chemistry. The reason is that most quantities of chemical significance are *energy differences* - not absolute energies. When differences are considered the errors in the Hartree-Fock approach are not small.

One way to improve the situation is to approximate the groundstate wave function by a series of determinants:

$$\Psi_{gs} = \sum_{[n_1, \dots, n_{N_e}]} c_{[n_1, \dots, n_{N_e}]} |\phi_{n_1} \dots \phi_{n_{N_e}}| \quad (4.6.1)$$

where  $\phi_n$  are an infinite orthonormal set of orbitals. Such an expansion can always be made, with any such set. We can thus take the orbitals produced by the Hartree-Fock process. This has the added nicety that the first determinant is already a good approximation to the ground-state.

In this case we can classify the determinants in the following way. We divide the orbitals into two sets: one is the set of  $N_e$  HF orbitals, called the occupied orbitals, the rest of orbitals named *virtualls*. We then classify the determinants by the number of *occupieds* missing. Thus we speak of all single substitutions, double substitutions etc.

A commonly employed method is the configuration-interaction (CI) method. In a CI one takes a number of determinants  $D_\alpha$   $\alpha = 1, \dots, M$  and uses them to minimize the energy:

$$E[\vec{c}] = \sum_{m,n=1}^M c_m \langle D_m | \hat{H} | D_n \rangle c_n \quad (4.6.2)$$

under the constraint of normalization:  $\sum_{m,n=1}^M c_m \langle D_m | D_n \rangle c_n = 1$ . One common way of choosing the determinants that go into this expansion is by collecting all single, double, triple etc excitations. A determinant is singly excited if when compared to the HF determinant it has one occupied molecular orbital replaced by some virtual orbital. Virtual orbitals are excited eigenfunctions of the Fock operator. One can show that the singles alone do not allow a reduction of energy. However, singles and doubles give sometimes good results. Such a method is called singles-doubles CI (SDCI). One problem with this theory is that it is not "size consistent". For example, calculating the energy of 2 distant Helium atoms will not give the twice the energy of one Helium atom under the same order of theory.

Another approach is many-body perturbation theory, called Moller-Plesser theory. In this approach, one writes the many-body Hamiltonian as:

$$\hat{H} = \sum_{n=1}^M \hat{F}_n + \hat{W} \quad (4.6.3)$$

Where  $\hat{W} = \hat{H} - \sum_{n=1}^N \hat{F}_n$  is considered a "small" perturbation. This quantity is not small enough and high order MP theory does not converge. However, second order MP theory, called MP2, is sometimes a useful approach. It is size consistent. However, it relies heavily on the quality of the Hartree-Fock

solution: Hartree-Fock must be a good reference on which to base a perturbation theory.

All wave-function methods beyond Hartree-Fock theory become quickly very expensive as system size grows. In fact, the numerical cost of good methods typically scales as  $O(N^7)$  for the coupled cluster method, which is a size-consistent non-variational method (variational methods are derived from the variational theorem), not discussed above. Thus, every enlargement of number of electrons by a factor of two makes the calculation a factor 100 more expensive!

## **V. Advanced topics in Hartree-Fock theory**

In this chapter we will continue our study of the Hartree-Fock approximation, and look into some of the formal issues, like stability, excitations and ionization and generalizations like fractional occupations numbers. We will then discuss the homogeneous electron gas in the Hartree-Fock approximation and show that it breaks down when treating this system.

### ***A. Low-lying excitations and the stability of the Hartree-Fock ground state***

#### **i. CI-Singles and Brillouin's theorem**

For a given system, one can think of the simplest excited states as linear combinations involving low lying energy determinants. For example, all determinants where one occupied orbital in the HF determinant is replaced by an orbital which is a "virtual" eigenstate of the Fockian (we call those  $N$  eigenstates of the Fockian which are part of the HF determinant "occupied"



and those which are not we term “virtual”). In the algebraic form of HF theory, if there are  $N$  electrons and the size of the basis is  $M > N$  then there are  $N$  ways to choose the occupied orbital to be replaced and  $(M - N)$  ways to choose the virtual orbital and so there are  $N(M - N)$  such “singly excited determinant”. Together with the HF determinants we can form a  $N(M - N) + 1$  dimensional determinant space and diagonalize the exact Hamiltonian in it. This will give an approximate description of the low lying excited states of the system. In fact, there is no need to include the HF determinant itself in this scheme since Brillouin showed that for an singly excited determinant  $\Psi_a^n$  (this notation is for a determinant that is obtained from the HF determinant by replacing the occupied  $a$  orbital by the virtual  $n$  orbital):

$$\langle \Psi_a^n | \hat{H} | \Psi_0 \rangle = 0 \quad (5.1.1)$$

Thus, the HF determinant is decoupled from the singly excited determinants and one can just diagonalize the Hamiltonian in the singly excited space. This approach is often called “configuration interaction – singles” (CIS) and it is a standard method for calculating excitation energies in HF theory. In fact, one cannot expect CIS to give a good approximation for the excited states, since the “real” excited states are intricate linear combinations of determinants with multi-excited electrons. On the other hand, our approximate ground state wave function is also just a single (Hartree-Fock) determinant and it too misses all this essential mixing with doubly and triply etc excitations. One can hope however for a mutual cancellation of errors. Indeed, there are many examples where the CIS method gives quite respectable excitation energies, even when the wave functions are of questionable quality.

## ii. Hartree-Fock Stability

We now consider the question of stability of the Hartree-Fock solution. Basically the issue is this: how do we know that we have produced the truly lowest energy by occupying the  $N$  orbitals with the lowest orbital energy? Maybe if we used a different ansatz we could have produced a lower energy. In other words, how do we make sure that all singly excited determinants are higher in energy than the HF determinant.

Suppose we have set up the Hartree-Fock equations and solved them to obtain a Fockian, a set of orbitals and orbital energies, out of which  $N_e$  are "occupied". Suppose the Hartree-Fock determinantal wave function is  $\Psi_0 = |\phi_1, \dots, \phi_a, \dots, \phi_{N_e}|$ . Let us consider the determinantal wave function  $\Psi_a^n = |\phi_1, \dots, \phi_n, \dots, \phi_{N_e}|$  obtained from the Hartree-Fock function by replacing an occupied spin-orbital  $\phi_a$  by the unoccupied spin-orbital  $\phi_n$ . This can be viewed as an electron excitation process: a hole is made in  $\phi_a$  and an electron is formed in  $\phi_n$ . The excitation energy for this excited state is the difference between the expectation values:

$$\begin{aligned} \epsilon_a^n &= \langle \Psi_a^n | \hat{H} | \Psi_a^n \rangle - \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \\ &= \langle \phi_n | \hat{h}_1 | \phi_n \rangle - \langle \phi_a | \hat{h}_1 | \phi_a \rangle \\ &\quad + \sum_m f_m (1 - \delta_{ma}) (\langle nm | nm \rangle - \langle nm | mn \rangle) \\ &\quad - \sum_m f_m (\langle am | am \rangle - \langle am | ma \rangle) \end{aligned} \quad (5.1.2)$$

Here, we introduced, for convenience the orbital occupation  $f_m = 1$  for  $m \leq N_e$  and  $f_m = 0$  otherwise. Rewriting:

$$\begin{aligned}
\epsilon_a^n = & \left[ \langle \phi_n | \hat{h}_1 | \phi_n \rangle + \sum_m f_m (\langle nm | nm \rangle - \langle nm | mn \rangle) \right] \\
& - \left[ \langle \phi_a | \hat{h}_1 | \phi_a \rangle + \sum_m f_m (\langle am | am \rangle - \langle am | ma \rangle) \right] \\
& - (\langle na | na \rangle - \langle na | an \rangle)
\end{aligned} \tag{5.1.3}$$

Using the definition of orbital energies in (4.1.30) we can thus write this “electron hole” excitation energy as:

$$\epsilon_a^n = \epsilon_n - \epsilon_a - \Delta_{an} \tag{5.1.4}$$

where:

$$\Delta_{an} = \langle an | an \rangle - \langle an | na \rangle. \tag{5.1.5}$$

One way to understand  $\Delta_{an}$  is as an over-counting term. We thus see that excitation energies as calculated as the difference between excited state energies in Hartree-Fock theory are *not* simply the differences of the Hartree-Fock orbital energies. They must actually be corrected for over-counting by subtracting a quantity  $\Delta_{an}$ , which we show henceforth to be manifestly positive. One conclusion is that orbital energy differences in Hartree-Fock form an upper-bound to the excitation energies, as determined from the Hartree-Fock single excited determinants. We may think of  $-\Delta_{an}$  as the *Coulomb energy of attraction* between the excited electron and the hole it leaves behind. Indeed  $\Delta_{an}$  is composed of the electrostatic interaction of the electron  $\phi_n$  and the hole  $\phi_a$ ,  $\left\langle \phi_a \phi_n \left| \frac{e^2}{r_{12}} \right| \phi_a \phi_n \right\rangle$  corrected by a corresponding exchange  $-\left\langle \phi_a \phi_n \left| \frac{e^2}{r_{12}} \right| \phi_n \phi_a \right\rangle$  term.

To prove  $\Delta_{an}$  is positive, note that it is symmetrical:  $\Delta_{an} = \Delta_{na}$ . Thus:

$$\begin{aligned}
\Delta_{an} &= \frac{1}{2}[\Delta_{an} + \Delta_{na}] \\
&= \frac{1}{2} \left[ \left\langle \phi_a \phi_n \left| \frac{e^2}{r_{12}} \right| \phi_a \phi_n - \phi_n \phi_a \right\rangle \right. \\
&\quad \left. + \left\langle \phi_n \phi_a \left| \frac{e^2}{r_{12}} \right| \phi_n \phi_a - \phi_a \phi_n \right\rangle \right] \\
&= \frac{1}{2} \left\langle \phi_a \phi_n - \phi_n \phi_a \left| \frac{e^2}{r_{12}} \right| \phi_a \phi_n - \phi_n \phi_a \right\rangle
\end{aligned} \tag{5.1.6}$$

This shows that

$$\Delta_{an} = \frac{e^2}{2} \iint \frac{(\phi_a(1)\phi_n(2) - \phi_a(2)\phi_n(1))^2}{|\mathbf{r}_2 - \mathbf{r}_1|} d^3r_1 d^3r_2 > 0 \tag{5.1.7}$$

Stability is obtained when  $\epsilon_a^n > 0$ , or, in other words when  $\epsilon_n - \epsilon_a > \Delta_a^n$ . Since  $\Delta_a^n > 0$ , we find that a *necessary* condition for stability is:

$$\epsilon_n - \epsilon_a > 0 \tag{5.1.8}$$

Clearly this might not be sufficient, however, it is necessary. That is: for stability to be possible, the orbital energy of all occupied orbitals must be lower than that of all the unoccupied orbitals otherwise there will definitely be “singly excited determinants” with lower energy.

### ***B. Koopmans’ Theorem***

What is the physical meaning of the orbital energies  $\epsilon_m$  in Eq. (4.1.28)? This was first discussed by the Dutch-American scientist Tjalling C. Koopmans (*Physica* **1934**, 1, 104.) in his PhD thesis. (After his Ph.D. with Hans Kramers, Koopmans began a scientific career in economics. He was awarded the 1975 Nobel Prize in economics “for his contribution to the optimal allocation of resources”).

Let us consider the ionization energy of a molecule in the Hartree-Fock approximation:

$$IP_{HF} = E_{HF}[\tilde{\phi}_1, \dots, \tilde{\phi}_{N-1}] - E_{HF}[\phi_1, \dots, \phi_N] \quad (5.2.1)$$

Note that we calculated the HF energy for two systems, the system  $N$  electrons and the ionized system of  $N - 1$  electrons. HF approximation will in general give *different* orbitals to the two systems, hence the notation:  $\phi_n$  for the  $N$  electron system and  $\tilde{\phi}_n$  for the ionized system. Now Koopmans *assumed* that the two sets of orbitals are identical. Of course they are not, but it is known that often they are similar so we neglect their difference. Actually, when the system is very large and is homogeneous (repeats itself), like an infinite crystalline solid, this assumption is expected to be exact because the orbitals are spread out on the entire system and therefore removal of just one out of an infinite number cannot make a difference. For molecules this assumption is a severe approximation. Nevertheless, under this approximation we see that all the one body terms cancel except the last and a large cancelation of two body terms takes place as well. Only the terms which involve the removed orbital stay, these include direct and exchange terms:

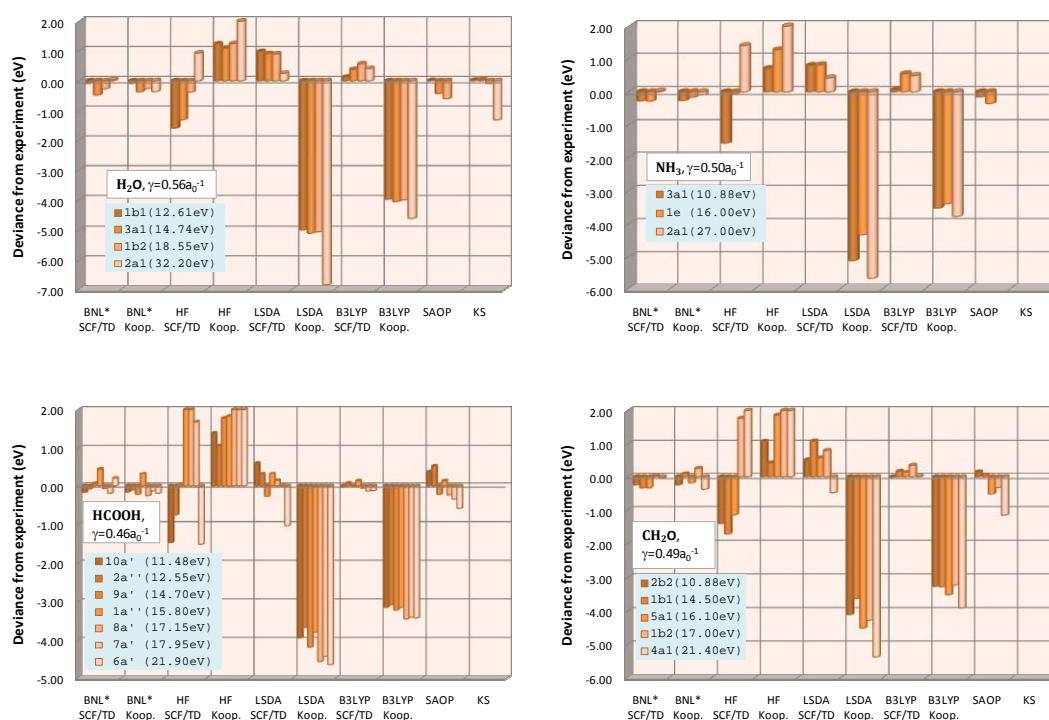
$$\begin{aligned} -IP_{HF} = & \langle \phi_{N_e} | \hat{h} | \phi_{N_e} \rangle + \frac{1}{2} \sum_{m=1}^N (\langle Nm | Nm \rangle - \langle Nm | mN \rangle) \\ & + \frac{1}{2} \sum_{m=1}^N (\langle mN | mN \rangle - \langle mN | Nm \rangle) \end{aligned} \quad (5.2.2)$$

Woring out the expression gives:

$$-IP_{HF} = \langle \phi_N | \hat{h} | \phi_N \rangle + \sum_{m=1}^N (\langle Nm | Nm \rangle - \langle Nm | mN \rangle) = \epsilon_N \quad (5.2.3)$$

The conclusion:  $\epsilon_{N_e}$  is the HF approximation to the ionization energy. Similarly,  $\epsilon_{N_e-1}$  approximates the next ionization energy etc.

The flaw in this “theorem” is the neglect of orbital relaxation. In the next section we will give a generalized formulation of Koopmans’ theorem which is exact. In the section after that we discuss the homogeneous electron gas which is a system for which orbital relaxation does not exist.



**Figure 4: Estimation of ionization energies in HF, approximate and accurate DFT methods. Two approaches are given for each method, one based on Koopmans’ approach, using the orbital energy of the neutral. One can see that Koopmans’ HF orbital energies usually overestimate the IPs by ~2 eV. Taken from ref. [12].**

Examples of the performance of Koopmans’ theorem within Hartree Fock and some DFT brands are give in the Figure I-1. In the figure the first IP is always the first left hand (brown) bar. Also shown are calculations for the IP’s using the  $\Delta$ SCF approach, where the IP is simply the difference between the HF energies of the cation and the neutral. In HF theory the Koopmans’ approach for the first IP are off by ~1 eV  $\Delta$ SCF has errors tend to be a bit larger. The

usual brands of DFT, such as LSDA, B3LYP have large Koopmans' errors but small  $\Delta$ SCF errors. The brand called BNL has small errors in both respects.

***C. Fractional occupation numbers, the HF orbital functional and the generalized Koopmans' theorem***

We now make a fundamental generalization of the Hartree-Fock energy. Instead of viewing it as an expectation value of a determinantal wave function, we view it as a new fundamental concept: *an orbital functional*. We take Eq. (4.1.9) and write it as a functional of "all" orbitals and occupation numbers:

$$\begin{aligned} \tilde{E}_{HF}[\phi_1, \phi_2, \dots; f_1, f_2, \dots] \\ = \sum_n f_n \langle \phi_n | \hat{h}_1 | \phi_n \rangle + \frac{1}{2} \sum_{nm} f_n f_m (\langle nm | nm \rangle - \langle nm | mn \rangle) \end{aligned} \quad (5.2.4)$$

Here the functional depends on an (in principle) infinite set of orthonormal orbitals and their corresponding occupation numbers. Since electrons are fermions each occupation number is limited to the unit interval, i.e.  $0 \leq f_i \leq 1$  (since you cannot have negative occupation and you cannot have more than 1 electron in a given orbital because of the Pauli principle). Let us minimize this functional for  $N$  electrons. What we mean by this is that we minimize this energy  $\tilde{E}_{HF}$  with respect to the orbitals  $\phi_n(\mathbf{r})$  and the occupation numbers  $f_n$  under the constraints that the orbitals are orthonormal and the occupation numbers are non-negative, not greater than 1 and that they sum up to  $N$ . Thus the Lagrangian for this constraint minimization is

$$\begin{aligned}
\tilde{L}_{HF}[\phi_1, \phi_2, \dots; f_1, f_2, \dots] \\
&= \tilde{E}_{HF}[\phi_1, \phi_2, \dots; f_1, f_2, \dots] - \sum_n f_n \epsilon_n [\langle \phi_n | \phi_n \rangle - 1] \\
&\quad - \chi \left( \sum_n f_n - N \right)
\end{aligned} \tag{5.2.5}$$

The derivative with respect to the orbitals give the analogous HF equations, which are similar to Eqs. (4.1.28) with an important difference: all sums on  $N$  orbitals are replaced by weighted sums:  $\sum_{n=1}^N X_n \rightarrow \sum_n f_n X_n$ . In essence, the usual HF theory is the constant occupation numbers  $f_n = \begin{cases} 1 & n \leq N \\ 0 & n > N \end{cases}$ . The new equations are:

$$\begin{aligned}
\hat{h}_1 |\phi_n\rangle + \hat{v}_H |\phi_n\rangle + \hat{K}_H |\phi_n\rangle &= \epsilon_n |\phi_n\rangle \\
\langle \mathbf{x} | \hat{v}_H | \psi \rangle &= e^2 \int \left( \sum_m \frac{f_m |\phi_m(\mathbf{x}')|^2}{|\mathbf{r}' - \mathbf{r}|} \right) dx' \psi(\mathbf{r}) \\
\langle \mathbf{x} | \hat{K} | \psi \rangle &= -e^2 \sum_n \phi_n(\mathbf{x}) \int \frac{f_m \phi_m(\mathbf{x}')^* \psi(\mathbf{r}_1)}{|\mathbf{r}' - \mathbf{r}|} dx'
\end{aligned} \tag{5.2.6}$$

Now, let's discuss the equations obtained by demanding variational behavior with respect to  $f_n$ . Since we have a constraint that  $f_n \in [0,1]$ , we must differentiate between three cases: the  $f_n = 0$ ,  $f_n = 1$  and *fractional* ( $0 < f_n < 1$ ) cases. To derive with respect to  $f_n$  is allowed only when you are not at constraint boundaries. So in the *fractional* case an arbitrary infinitesimal change in  $f_n$  is indeed meaningful and the derivative of  $\tilde{L}$  with respect to  $f_n$  is equal to zero at the variationally optimal point. This gives:

$$\chi = \frac{\delta \tilde{E}_{HF}}{\delta f_n} = \langle \phi_n | \hat{h}_1 | \phi_n \rangle + \sum_m f_m (\langle nm | nm \rangle - \langle nm | mn \rangle) = \epsilon_n \tag{5.2.7}$$

This equation shows that all fractionally occupied orbitals must be of the same energy  $-\chi$ . We can divide all orbitals to be full ( $f_n = 1$ ) or empty



( $f_n = 0$ ) and partially occupied but these latter orbitals all have the same orbital energy. We have seen that for the Hartree-Fock solution to be stable, only the lowest orbitals can be occupied. Thus the picture that emerges is: there are  $N - j$  fully occupied orbitals. Then there can be several degenerate orbitals with fractional occupation that adds up to  $j$  (assuming  $N$  is integer). The rest of the orbitals are unoccupied.

Eq. (5.2.7) is now an *exact* formulation of Koopmans' theorem. Suppose we have a system with slightly less than an integer number of electrons:  $N - \eta$  where  $\eta$  is a small fraction. Then:

$$\tilde{E}_{HF}[N] - \tilde{E}_{HF}[N - \eta] = \epsilon_H(N)\eta + O(\eta^2) \quad (5.2.8)$$

Where  $\epsilon_H(N)$  is the energy of the highest (can be partially) occupied orbital for the  $N$ -electron system. Thus we see that if we "slightly" ionize the system Koopmans' theorem holds exactly (in a sense, the change is so small that there will be no orbital relaxation). Of course, in real molecules there is no such thing as a fractional electron. But still, in terms of the *orbital HF theory* there is.

This new concept of an orbital functional has allowed us to consider an exact and generalized version of Koopmans' theorem. We will see in later chapters that orbital functionals play an important role in advanced approaches to DFT.

## ***D. Hartree-Fock for the homogeneous electron gas***

### **i. Hartree-Fock orbitals and orbital energies of HEG**

Let us now apply the Hartree-Fock theory to an important system which is a model for the valence electrons of a simple metal such as sodium. This is once again the homogeneous electron, gas of  $N$  electrons in a cubic box of volume  $V$  under periodic boundary conditions. Note that we are imposing on our

electron gas to be uniform, although it is not necessarily known that this is the lowest energy solution. In Hartree-Fock theory for example there are non-homogeneous (symmetry broken) states which lead to lower energy. We will not study these here however. The density is  $n = N/V$ . For an interacting system one must have charge neutrality, so a static positive charge of the same density as the electrons is smeared in the box. This positive charge is called "Jellium". The e-e direct term, the Jellium self energy (positive-positive interaction) and the electron-Jellium energy must all cancel each other:

$$\begin{aligned}
 J_{ee} + \int v n d^3r + j_{pp} \\
 = \frac{e^2 n^2}{2} \iint \frac{d^3r d^3r'}{|\mathbf{r} - \mathbf{r}'|} - e^2 n^2 \iint \frac{d^3r d^3r'}{|\mathbf{r} - \mathbf{r}'|} \\
 + \frac{e^2 n^2}{2} \iint \frac{d^3r d^3r'}{|\mathbf{r} - \mathbf{r}'|} = 0
 \end{aligned} \tag{5.2.9}$$

So in the Hartree-Fock approach, all that is left are the kinetic energy and the exchange energies:

$$\epsilon_{HF}(n) = t(n) + \epsilon_X(n) \tag{5.2.10}$$

The HF equations are:

$$-\frac{\hbar^2}{2m_e} \nabla^2 \phi_k(\mathbf{r}) + \hat{K} \phi_k(\mathbf{r}) = \epsilon_k \phi_k(\mathbf{r}) \tag{5.2.11}$$

The solutions must be plane waves (since all points are equivalent). Thus:

$$\epsilon_k = \frac{\hbar^2 k^2}{2m_e} + X_k^{-1} \tag{5.2.12}$$

And:

$$\hat{K} \phi_k = X_k^{-1} \phi_k \tag{5.2.13}$$

Let us calculate the exchange eigenvalues  $X_k^{-1}$ . Assume double occupancy of all orbitals. Then:

$$\begin{aligned}
\widehat{K}\phi_{\mathbf{k}}(\mathbf{r}) &= -\frac{2e^2}{V^{3/2}} \sum_{\{\mathbf{q}|\epsilon_{\mathbf{q}} < \epsilon_F\}} e^{i\mathbf{q}\cdot\mathbf{r}} \int \frac{e^{-i\mathbf{q}\cdot\mathbf{r}'} e^{i\mathbf{k}\cdot\mathbf{r}'}}{|\mathbf{r}-\mathbf{r}'|} d^3r' \\
&= -\frac{2e^2}{V^{3/2}} \sum_{\{\mathbf{q}|\epsilon_{\mathbf{q}} < \epsilon_F\}} e^{i\mathbf{q}\cdot\mathbf{r}} \int \frac{e^{i(\mathbf{k}-\mathbf{q})\cdot\mathbf{r}'}}{|\mathbf{r}-\mathbf{r}'|} d^3r' \\
&= -\frac{2e^2}{V^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\{\mathbf{q}|\epsilon_{\mathbf{q}} < \epsilon_F\}} \int \frac{e^{i(\mathbf{k}-\mathbf{q})\cdot(\mathbf{r}'-\mathbf{r})}}{|\mathbf{r}'-\mathbf{r}|} d^3r'
\end{aligned} \tag{5.2.14}$$

Here the summation symbol with  $\{\mathbf{q}|\epsilon_{\mathbf{q}} < \epsilon_F\}$  means that we sum over all momentum states  $\mathbf{q}$  with  $\epsilon_{\mathbf{q}} < \epsilon_F$ . We now discuss the evaluation of  $\int \frac{e^{i\mathbf{w}\cdot\mathbf{x}}}{|\mathbf{x}|} d^3x$ . Let's assume the box is a sphere of radius  $R$  and add a damping factor  $\eta > 0$  (which can be set to zero after the calculation is done). For a finite volume

$$\begin{aligned}
\int \frac{e^{i\mathbf{w}\cdot\mathbf{x}}}{x} e^{-\eta x} d^3x &= \int_0^R e^{-\eta x} \int_0^\pi \frac{e^{iwx \cos \theta}}{x} \sin \theta d\theta 2\pi x^2 dx \\
&= \int_0^R \frac{e^{(i\mathbf{w}-\eta)x} - e^{(-i\mathbf{w}-\eta)x}}{iwx^2} 2\pi x^2 dx \\
&= \frac{2\pi}{i\mathbf{w}} \left[ \frac{e^{(i\mathbf{w}-\eta)R} - 1}{(i\mathbf{w}-\eta)} + \frac{e^{(-i\mathbf{w}-\eta)R} - 1}{(i\mathbf{w}+\eta)} \right] \\
&= \frac{2\pi}{i\mathbf{w}} \left[ \frac{(e^{(i\mathbf{w}-\eta)R} - 1)(i\mathbf{w}+\eta) + (e^{(-i\mathbf{w}-\eta)R} - 1)(i\mathbf{w}-\eta)}{(i\mathbf{w}-\eta)(i\mathbf{w}+\eta)} \right] \\
&= \frac{4\pi}{\mathbf{w}^2 + \eta^2} \left[ (1 - e^{-\eta R} \cos(\mathbf{w}R)) - \frac{\eta}{\mathbf{w}} e^{-\eta R} \sin(\mathbf{w}R) \right]
\end{aligned}$$

Take the limit  $R \rightarrow \infty$  and *after* that the limit  $\eta \rightarrow 0$ . This will lead to the result  $\frac{4\pi}{\mathbf{w}^2}$ . For a finite  $R$ , take the  $\eta \rightarrow 0$  limit first:

$$\int \frac{e^{i\mathbf{w}\cdot\mathbf{x}}}{x} d^3x = \frac{4\pi}{\mathbf{w}^2} \begin{cases} 1 - \cos(\mathbf{w}R) & R \text{ finite} \\ 1 & R \rightarrow \infty \end{cases}$$

The result of infinite  $R$  becomes undefined in the second case. So we use the first. Thus we write:

$$\hat{K}\phi_{\mathbf{k}}(\mathbf{r}) = -\frac{2}{V^{3/2}} \sum_{\{\mathbf{q}|\epsilon_{\mathbf{q}} < \epsilon_F\}} e^{i\mathbf{k}\cdot\mathbf{r}} \frac{4\pi e^2}{(\mathbf{k}-\mathbf{q})^2} = \frac{1}{X_{\mathbf{k}}} \phi_{\mathbf{k}}(\mathbf{r}) \quad (5.2.15)$$

Where:

$$\frac{1}{X_{\mathbf{k}}} = -\frac{2}{V} \sum_{\{\mathbf{q}|\epsilon_{\mathbf{q}} < \epsilon_F\}} \frac{4\pi e^2}{(\mathbf{k}-\mathbf{q})^2} \quad (5.2.16)$$

Now, to proceed, we want to assume that the momentum states get filled up in just the same way as they did in the Thomas-Fermi approach for a non-interacting gas. This will be valid if the orbital energies  $\epsilon_k$  are ascending functions of  $k$ , which according to Eq. (5.2.12) is valid when  $\frac{1}{X_k}$  is an increasing function of  $k$  as well. We can only know that however, after we evaluate the summation in Eq. (5.2.16)... What we can do, is work in the spirit of the self consistent field approach. We will *assume* that  $\epsilon_k$  increasing functions of  $k$  then sum Eq. (5.2.16) and check, for self consistency, that  $\frac{1}{X_k}$  is an increasing function of  $k$ . Under this plan, we replace in Eq. (5.2.16) the summation over all  $\mathbf{q}$  with  $\epsilon_{\mathbf{q}} < \epsilon_F$  by a summation over all  $\mathbf{q}$  with  $q < k_F$ , which can be approximated by the an integral (Eq. (2.1.8)), we find:

$$\frac{1}{X_{\mathbf{k}}} = -\frac{2}{(2\pi)^3} \int_{k < k_F} \frac{4\pi}{(\mathbf{k}-\mathbf{q})^2} d^3q \quad (5.2.17)$$

To evaluate this integral we pass over to spherical coordinates,  $k$ ,  $\theta$  and  $\phi$ :

$$\frac{1}{X_{\mathbf{k}}} = -\frac{2}{\pi} \int_0^{k_F} \int_0^{\pi} \frac{\sin\theta d\theta}{q^2 + k^2 - 2kq\cos\theta} q^2 dq \quad (5.2.18)$$

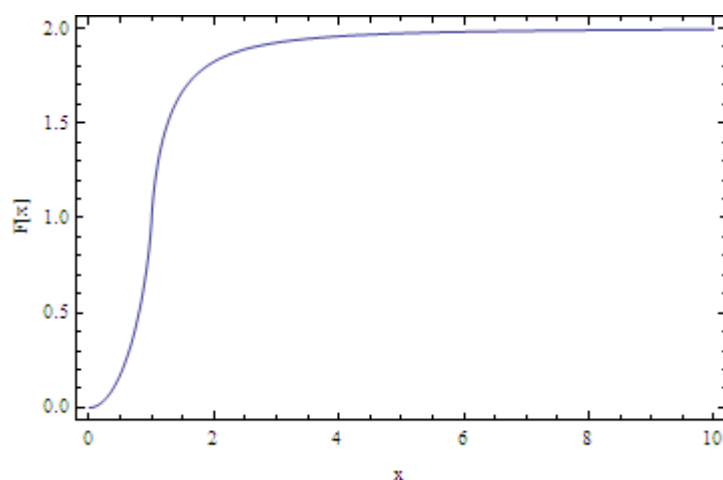
The factor of  $2\pi$  is a consequence of the integral over the angle  $\phi$  (the integrand is independent of  $\phi$ ). The integral over  $\theta$  is can be performed after change of variables  $x = \sin\theta$ , noting that  $\int_{-1}^1 \frac{dx}{q^2 + k^2 - 2kqx} = \left( \frac{\ln(q^2 + k^2 - 2kqx)}{-2kq} \right)_{-1}^1$  so:

$$\frac{1}{X_k} = -\frac{2}{k\pi} \int_0^{k_F} \ln \left| \frac{k+q}{k-q} \right| q dq \equiv -\frac{2k_F}{\pi} F\left(\frac{k_F}{k}\right) \quad (5.2.19)$$

We changed variable to  $x = \frac{q}{k}$  and gave the integral over  $x$  a new name:

$$F(x) = \frac{1}{x} \int_0^x \ln \left| \frac{1+x}{1-x} \right| x dx = 1 + \frac{1}{2} \left( x - \frac{1}{x} \right) \ln \left| \frac{x+1}{x-1} \right| \quad (5.2.20)$$

This function is shown in Figure V-5. It is monotonically increasing from zero reaching 2 when  $x \rightarrow \infty$ .

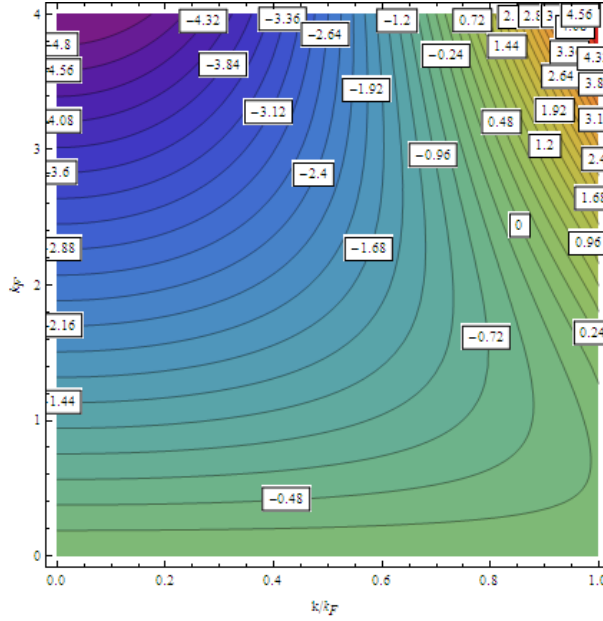


**Figure V-5: The function  $F(x)$  defined in Eq. (5.2.20)**

The HF orbital energies are now

$$\epsilon_k = \frac{\hbar^2 k^2}{2m_e} - \frac{e^2}{4\pi\epsilon_0} \frac{2k_F}{\pi} F\left(\frac{k_F}{k}\right) \quad (5.2.21)$$

Clearly, as  $F(x)$  is increasing as a function of  $x$ , it is decreasing with  $k$ , yet it comes into the orbital energy expression with a minus sign so overall  $\epsilon_k$  is an increasing function of  $k$ . A plot of  $\epsilon_k$ , in atomic units is given in Figure V-6. It is seen that indeed  $\epsilon_k$  is increasing monotonically with  $k$ .



**Figure V-6: The Hartree-Fock orbital energies  $\epsilon_k$  for various values of  $k_F$  (all in atomic units). For any given  $k_F$  the function  $\epsilon_k$  is increasing with  $k$ .**

Exercise: Determine the Hartree-Fock density matrix of the HEG.

We have shown that the Hartree-Fock orbitals are plane waves with wave vectors less than  $k_F$ . Thus:

$$\rho(\mathbf{r}, \mathbf{r}') = \frac{1}{V} \sum_{k < k_F} e^{i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}')} \quad (5.2.22)$$

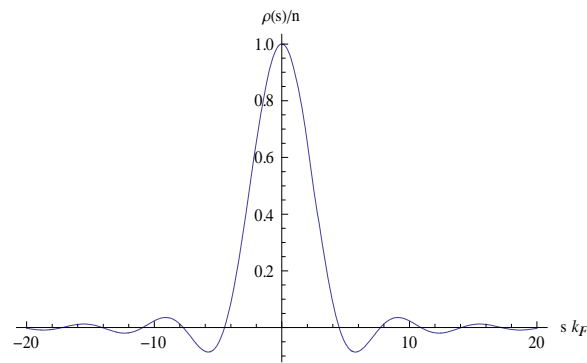
Denoting  $\mathbf{r} - \mathbf{r}' \equiv \mathbf{s}$  we find that  $\rho$  depends only on  $s$  and so:

$$\begin{aligned} \rho(s) &= 2 \frac{1}{(2\pi)^3} \int_{k < k_F} e^{i\mathbf{k} \cdot \mathbf{s}} d^3 k = \frac{2}{(2\pi)^3} \int_0^{k_F} \int_{-1}^1 e^{iksx} dx 2\pi k^2 dk \\ &= \frac{2}{(2\pi)^3} \frac{4\pi}{s^3} \int_0^{sk_F} \sin x x dx = \frac{1}{(2\pi)^3} \frac{8\pi}{s^3} (\sin x - x \cos x)_0^{sk_F} \quad (5.2.23) \\ &= \frac{k_F^3}{\pi^2} \left( \frac{\sin sk_F - sk_F \cos sk_F}{(sk_F)^3} \right) \end{aligned}$$

Since  $k_F^3 = 3\pi^2 n$  we find:

$$\rho(s) = 3n \left( \frac{\sin sk_F - sk_F \cos sk_F}{(sk_F)^3} \right) \quad (5.2.24)$$

We know that  $\rho(\mathbf{r}, \mathbf{r}) = \rho(0) = n$ . A plot of the DM is plotted here:



## ii. The density of states of the HEG

An interesting quantity, determining many of the properties of the substance is the density of single-particle states (DOS)  $D(\epsilon_F)$  at the Fermi level. One can calculate and obtain:

$$\begin{aligned}
 D(\epsilon) &= \frac{V}{(2\pi)^3} \int_0^{k_F} \delta(\epsilon - \epsilon_k) 4\pi k^2 dk \\
 &= \frac{V}{(2\pi)^3} \int_0^{\epsilon_F} \delta(\epsilon - \epsilon_k) 4\pi k(\epsilon_k)^2 \frac{dk}{d\epsilon_k} d\epsilon_k \\
 &= \frac{V}{2\pi^2} \theta(\epsilon_F - \epsilon) k(\epsilon)^2 k'(\epsilon)
 \end{aligned} \tag{5.2.25}$$

Here we need to invert the relation  $\epsilon_k$  to  $k(\epsilon)$ . Clearly, one can write this as:

$$D(\epsilon_k) = \frac{V}{2\pi^2} \theta(\epsilon_F - \epsilon_k) k^2 (\epsilon'(k))^{-1} \tag{5.2.26}$$

Thus, to obtain the DOS we need to take the inverse of the derivative  $\epsilon'_k$ . Let us look more closely into that. Note that for free electrons  $\frac{k^2}{\epsilon'(k)} = \sqrt{2\mu^3\epsilon}$  thus  $D(\epsilon) \propto \sqrt{\epsilon}$ . In particular, the DOS at the Fermi level is finite, a typical situation for metals, which is what the HEG is. Now, let us look what happens according to HF theory. First, let us look numerically-graphically at  $\epsilon_k$  for some arbitrary definite value  $k_F = 1$ . This is shown in Figure V-7 (left). It seems perfectly OK. However, when we plot the derivative  $\epsilon'_k$ , one notices a

divergent behavior at  $k \rightarrow k_F$ . Indeed, taking the derivative of Eq. (5.2.21) we obtain, using  $F'(x) = \frac{1}{x} \left( -1 + \frac{1}{2} \left( x + \frac{1}{x} \right) \ln \left| \frac{x+1}{x-1} \right| \right)$ :

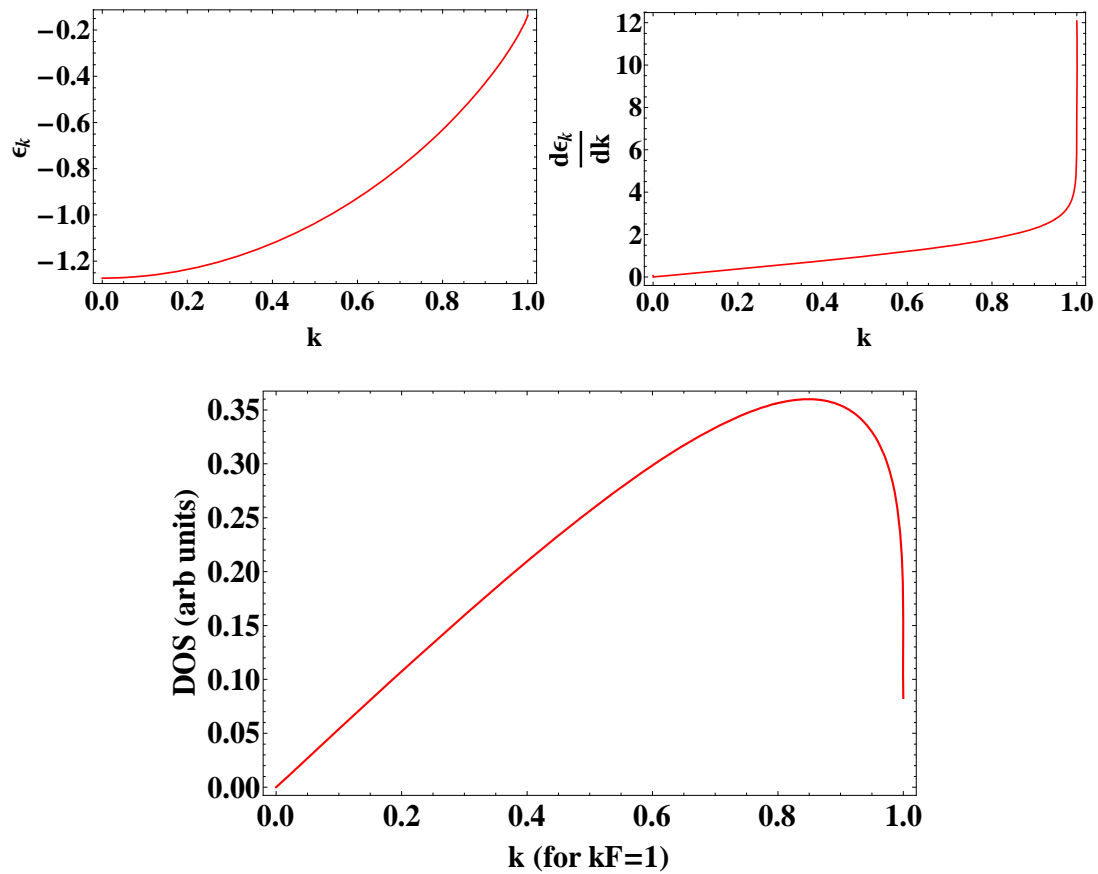


Figure V-7: The functions  $\epsilon_k$  (top left) and  $\epsilon'_k$  (top right) and the DOS of the HEG, according to Hartree-Fock theory.

$$\epsilon'_k = \frac{\hbar^2 k}{m_e} + \frac{e^2}{4\pi\epsilon_0} \frac{2}{\pi} \frac{k_F}{k} \left( -1 + \frac{1}{2} \left( \frac{k_F}{k} + \frac{k}{k_F} \right) \ln \left| \frac{k_F + k}{k_F - k} \right| \right) \quad (5.2.27)$$

Indeed, as  $k \rightarrow k_F$  there is a logarithmic divergence of this expression. Since the DOS is proportional to the inverse of  $\epsilon'_k$ , this means that at  $\epsilon_F$  the DOS is zero! This prediction by Hartree-Fock theory can easily be checked experimentally. For example, electronic conduction of metals is high. Since the conduction depends on the availability of electrons at the Fermi level we see that HF predicts small metallic conductance, failing miserably...



### iii. Stability of the HEG in Hartree-Fock theory

The density matrix assuming unpolarized gas (all electrons are spin-paired) is given by Eq. (5.2.24). Now, we can calculate the total exchange energy. From Eq. (4.1.17):

$$\begin{aligned}
 K &= -\frac{1}{4} \iint \frac{\rho(\mathbf{r}, \mathbf{r}')^2}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \\
 &= -\frac{V}{4} \int_0^{V^{1/3}} \left( \frac{\sin k_F R - k_F R \cos k_F R}{\pi^2 R^3} \right)^2 4\pi R dR = \quad (5.2.28) \\
 &= -\frac{V k_F^4}{\pi^3} \int_0^{k_F V^{1/3}} \left( \frac{\sin x - x \cos x}{x^3} \right)^2 x dx
 \end{aligned}$$

Notice that  $\int_0^x \left( \frac{\sin x - x \cos x}{x^3} \right)^2 x dx = \frac{\cos 2x + 2x \sin 2x - 1 - 2x^2}{8x^4} + \frac{1}{4}$ . Thus for  $x \rightarrow \infty$  which is the limit of large volume, we have:

$$\int_0^\infty \left( \frac{\sin x - x \cos x}{x^3} \right)^2 x dx = \frac{1}{4} \quad (5.2.29)$$

Thus for large volume:

$$K = -\frac{V k_F^4}{4\pi^3} \quad (5.2.30)$$

From (2.1.15)  $n = \frac{k_F^3}{3\pi^2}$  so:

$$K = -\frac{3k_F}{4\pi} N \quad (5.2.31)$$

The exchange energy per particle is:

$$\epsilon_x = -\frac{3k_F}{4\pi} = -C_X n^{1/3} \quad (5.2.32)$$

With  $C_X = \frac{3}{4} \left( \frac{3}{\pi} \right)^{1/3} = 0.73856$ . Let us recall the kinetic energy per particle for the HEG (Eq. (2.1.18)):

$$\epsilon_{HF} = t + \epsilon_x = C_{TF}n^{2/3} - C_Xn^{1/3} \quad (5.2.33)$$

With  $C_{TF} = 2.871$ . We see that for high density the energy is primarily kinetic and rises with density. For low density the energy drops as the density is increased. Thus there is an equilibrium point. This can be seen in Figure V-8.

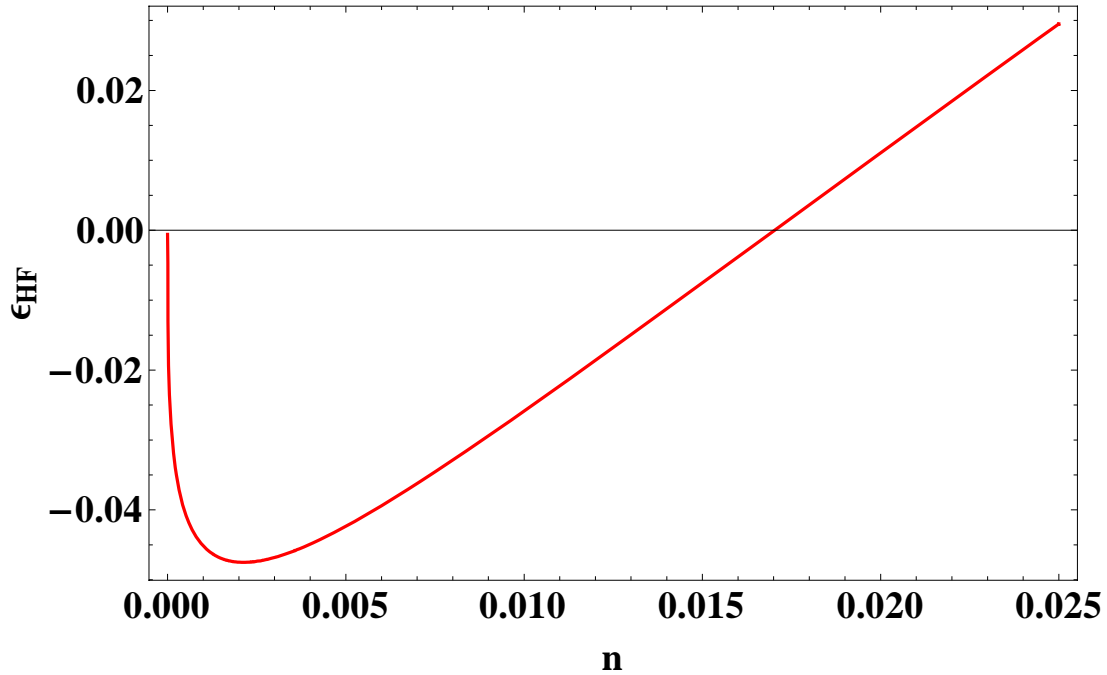


Figure V-8: The total energy, according to the Hartree-Fock approximation, of Jellium, as a function of the density  $n$ . All quantities – in atomic units.

We see that Jellium is most stable at density  $n^*$  calculated from:

$$\epsilon'_{HF} = \frac{2}{3}C_{TF}n^{-1/3} - \frac{1}{3}C_Xn^{-2/3} = 0 \rightarrow n^* = \left(\frac{C_X}{2C_{TF}}\right)^3 \quad (5.2.34)$$

This gives:  $n^* = 0.0021a_0^{-3}$  and  $\epsilon^* = -0.048E_h = -1.29 \text{ eV}$ .

For comparison, sodium, which is a monovalent metal with properties similar to Jellium has a total energy per atom (valence electron) of:  $\epsilon_{Na} = -1.13 \text{ eV}$  and  $n_{Na} = 0.0038 a_0^{-3}$ .

## VI. The Hohenberg-Kohn density 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 wavefunction. 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.

### *A. The first HK theorem*

In electronic structure theory the Hamiltonian is given as:

$$\hat{H} = \hat{T} + \hat{U} + \int v(\mathbf{r})\hat{n}(\mathbf{r})d^3r \quad (6.1.1)$$

Where all symbols have been defined in XXX. The identity of the molecular system is captured in the external potential  $v(\mathbf{r})$ . The other terms are “universal” i.e. the same for all molecules. In DFT they have a special symbol:

$$\hat{F} = \hat{T} + \hat{U}. \quad (6.1.2)$$

Different Hamiltonians differ only by their external potential-density term:

$$\hat{H} - \hat{H}' = \int (v(\mathbf{r}) - v'(\mathbf{r}))\hat{n}(\mathbf{r})d^3r \quad (6.1.3)$$

This observation has a fundamental implication. Suppose we have two electronic wave functions  $\Psi_1$  and  $\Psi_2$  of  $N_e$  electrons which have the same density, i.e. for all  $\mathbf{r}$ :

$$\langle \Psi_1 | \hat{n}(\mathbf{r}) | \Psi_1 \rangle = \langle \Psi_2 | \hat{n}(\mathbf{r}) | \Psi_2 \rangle \equiv n(\mathbf{r}). \quad (6.1.4)$$

Now, if

$$\langle \Psi_1 | \hat{H} | \Psi_1 \rangle < \langle \Psi_2 | \hat{H} | \Psi_2 \rangle \quad (6.1.5)$$

Then:

$$\langle \Psi_1 | \hat{F} | \Psi_1 \rangle < \langle \Psi_2 | \hat{F} | \Psi_2 \rangle \quad (6.1.6)$$

Thus the inequality is independent of the position of the nuclei: only the wave functions affect it through the universal operator  $\hat{F}$ . The external potential term drops out because both wave functions have the same density.

An interesting natural conclusion is that if Eq. (6.1.5) holds for one Hamiltonian then it holds for all Hamiltonians::

$$\langle \Psi_1 | \hat{H}' | \Psi_1 \rangle < \langle \Psi_2 | \hat{H}' | \Psi_2 \rangle \quad (6.1.7)$$

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

**Theorem (Hohenberg-Kohn):** When two Hamiltonians differing only by a single particle potential term  $\hat{H} - \hat{H}' \equiv \int [v(\mathbf{r}) - v'(\mathbf{r})] \hat{n}(\mathbf{r}) d^3r$  have non degenerate ground states which integrate to the same density then these Hamiltonians are identical up to a constant (i.e.  $v(\mathbf{r}) = v'(\mathbf{r}) + const$ ).

**Proof of the HK theorem:** Assume otherwise:  $\Psi$  is the GS of  $\hat{H}$  and  $\Psi'$  that of  $\hat{H}'$ , both wave functions assumed real and have the same expectation values for the density at all points in space. The variational principle for  $\Psi$  dictates  $\langle \Psi | \hat{H} | \Psi \rangle < \langle \Psi' | \hat{H} | \Psi' \rangle$  which, as discussed in Eq. (6.1.7) means that  $\langle \Psi | \hat{H}' | \Psi \rangle < \langle \Psi' | \hat{H}' | \Psi' \rangle$  holds as well. But this latter inequality contradicts the variational principle for  $\Psi'$  as the ground state of  $\hat{H}'$ , unless it differs from  $\hat{H}$  by at most a constant. ♦

This theorem allows one to think of the potential as a functional dependent on the density. Thus, in addition to the “usual”

$$v \rightarrow \hat{H} \rightarrow \psi \rightarrow n \quad (6.1.8)$$

We now have:

$$n \rightarrow v \rightarrow \hat{H} \rightarrow \psi \quad (6.1.9)$$

Thus "everything" about the molecule (all its properties) is in the above sense a functional of the ground-state density.

A generalization of the theorem, giving an inequality is:

**Theorem:** If  $n_i(\mathbf{r})$  is the density of the non-degenerate N-particle groundstate of  $\hat{H}_i = \hat{F} + \int v_i(\mathbf{r})\hat{n}(\mathbf{r})d^3r$ , where  $i = 1,2$ . Then, denoting  $\Delta X = X_1 - X_2$ , we have:

$$\Delta n \neq 0 \implies \int \Delta n(\mathbf{r})\Delta v(\mathbf{r})d^3r < 0 \quad (6.1.10)$$

**Proof:** Suppose  $\Delta n \neq 0$ . Then  $\Psi_1 \neq \Psi_2$  and because of non-degeneracy the following inequality is strict:

$$\langle \Psi_1 | \hat{H}_1 | \Psi_1 \rangle < \langle \Psi_2 | \hat{H}_1 | \Psi_2 \rangle = \langle \Psi_2 | \hat{H}_2 | \Psi_2 \rangle + \int [v_1(\mathbf{r}) - v_2(\mathbf{r})]n_2(\mathbf{r})d^3r \quad (6.1.11)$$

Denoting  $E_i = \langle \Psi_i | \hat{H}_i | \Psi_i \rangle$ , we find:

$$E_1 < E_2 + \int \Delta v(\mathbf{r})n_2(\mathbf{r})d^3r \quad (6.1.12)$$

And exchanging the indices  $1 \leftrightarrow 2$ :

$$E_2 < E_1 - \int \Delta v(\mathbf{r})n_1(\mathbf{r})d^3r \quad (6.1.13)$$

Adding the two inequalities and cleaning up gives Eq. (6.1.10), QED♦

### ***B. The HK functional***

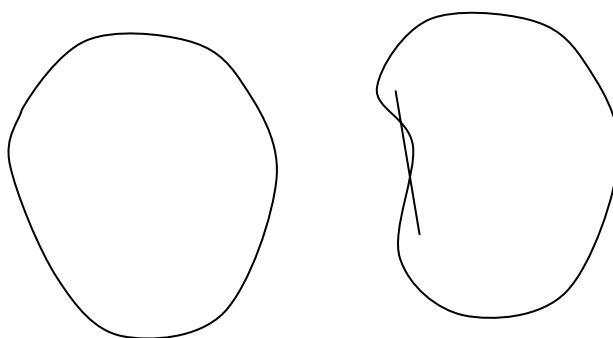
Since "everything" is a functional of the density, we can assert that the ground-state kinetic energy is a functional  $T[n]$  of the density and so is the

electron repulsion energy  $U[n]$ . We can thus define the Hohenberg-Kohn functional of v-rep densities as follows:

$$F_{HK}[n] = T[n] + U[n] = \langle \Psi_{gs}[n] | \hat{T} + \hat{U} | \Psi_{gs}[n] \rangle \quad (6.2.1)$$

Where  $\Psi_{gs}[n]$  is a ground-state wave function with density  $n$ .  $F_{HK}[n]$  is a universal functional, it is not limited to any particular molecular system. It is valid for *all* systems. We have of course no practical way to calculate  $F_{HK}[n]$  in general.

Even the domain of definition of  $F_{HK}$  is difficult to characterize. In fact, and perhaps unexpectedly, this domain is not even convex, as discovered by Levy and Perdew.



**Figure VI-1: A convex set is a set of points such that if  $A$  and  $B$  are in the set then any point  $C$  on the straight line joining  $A$  and  $B$  is in the set as well. The left set is convex while the right set is not.**

This means, that if  $n_0$  and  $n_{\pi/2}$  are two densities which are in the domain, it is not guaranteed that the convex sum  $n_\theta \equiv (\cos^2 \theta)n_0 + (\sin^2 \theta)n_{\pi/2}$  (where  $\theta$  is some parameter in the range  $[0, \frac{\pi}{2}]$ ) is in it. Levy and Perdew showed an example where convexity fails by considering the case where both are densities are densities of two degenerate eigenstates of the same system (we discuss this in more detail below). Technically, we say that the domain of definition is not convex. However, suppose we do have three densities  $n_0$ ,  $n_{\pi/2}$  and  $n_\theta$ , where the latter is the convex sum of the two former densities.

Suppose further that all three belong to the domain of definition of  $F_{HK}$ . Then  $F_{HK}[n_\theta] \leq (\cos^2 \theta)F_{HK}[n_0] + (\sin^2 \theta)F_{HK}[n_{\pi/2}]$ , i.e.  $F_{HK}$  is a convex functional of the density (we discussed this concept in Chapter XXX). This is seen as a result of the variational theorem. Suppose  $\Psi_\theta$  is the ground-state wave function,  $v_\theta$  the potential and  $\hat{H}_\theta = \hat{T} + \hat{U} + \hat{V}_\theta$  the Hamiltonian corresponding to  $n_\theta$  then:

$$\begin{aligned} \langle \Psi_\theta | \hat{H}_\theta | \Psi_\theta \rangle &\leq \langle \Psi_0 | \hat{H}_\theta | \Psi_0 \rangle \\ \langle \Psi_\theta | \hat{H}_\theta | \Psi_\theta \rangle &\leq \langle \Psi_{\pi/2} | \hat{H}_\theta | \Psi_{\pi/2} \rangle \end{aligned} \quad (6.2.2)$$

From this we have:

$$\begin{aligned} F_{HK}[n_\theta] &= \langle \Psi_\theta | \hat{H}_\theta | \Psi_\theta \rangle - \int n_\theta(\mathbf{r})v_\theta(\mathbf{r})d^3r \\ &\leq (\cos^2 \theta) \left[ \langle \Psi_0 | \hat{H}_\theta | \Psi_0 \rangle - \int n_0(\mathbf{r})v_\theta(\mathbf{r})d^3r \right] \\ &\quad + (\sin^2 \theta) \left[ \langle \Psi_{\pi/2} | \hat{H}_\theta | \Psi_{\pi/2} \rangle - \int n_{\pi/2}(\mathbf{r})v_\theta(\mathbf{r})d^3r \right] \\ &= (\cos^2 \theta)F_{HK}[n_0] + (\sin^2 \theta)F_{HK}[n_{\pi/2}] \end{aligned} \quad (6.2.3)$$

We will see that the convexity property of  $F_{HK}[n]$  has an important implication for the variational property of density functional theory, that we discuss in the next section.

### ***C. Minimum principle for density functional theory***

The second HK theorem establishes a minimum principle involving the density, and it can be used to "find" the density without direct reference to the concept of a "wave function".

Given a potential  $v(\mathbf{r})$ , consider the following functional for  $N$  electrons:

$$E_v[n] = F_{HK}[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r \quad (6.3.1)$$

The HK theorem II states that the density  $n_v$  which is the ground-state of  $v(\mathbf{r})$  minimizes this functional. For  $n_v$  we know of course that:

$$E_v[n_v] = \langle \Psi_{gs}[n_v] | \hat{H} | \Psi_{gs}[n_v] \rangle = E_{gs} \quad (6.3.2)$$

Where  $\hat{H} = \hat{T} + \hat{V} + \hat{U}$ . HK theorem II states that for any other v-rep density  $n'$  of  $N$  electrons:

$$E_v[n_v] < E_v[n'] \quad (6.3.3)$$

The proof is an immediate consequence of the quantum mechanical variational principle:

$$E_v[n'] = \langle \Psi_{gs}[n'] | \hat{H} | \Psi_{gs}[n'] \rangle > \langle \Psi_{gs}[n_v] | \hat{H} | \Psi_{gs}[n_v] \rangle = E_{gs} \quad (6.3.4)$$

This theorem allows one to speak of a "minimum-principle" concerning the density. If we have an approximate  $E_v[n]$  we can find an approximation to the ground-state density  $n_v$ , simply by finding the minimum.

The fact that  $F_{HK}[n]$  is convex (in the limited sense, since there is still the issue of domain of definition discussed below), as proved in the previous section, implies that  $E_v[n]$  is convex as well, since by adding a linear term to a function one cannot change its convex/non-convex character. The convexity is desirable since it assures that the minimum is not only global, but that there are no local minima as well, since a convex function (and functional) have no truly local minima. There is still the problem of the *convexity of the domain* and this spoils this useful conclusion. Below we show how to construct a functional with a convex domain. But then, we lose the convex property of the functional.

We now derive the basic equation of DFT (ignoring for the time being the problem of convexity of the domain). We need to minimize  $E_v[n]$  under the constraint that  $\int n(\mathbf{r})d^3r = N_e$ . Thus we need to minimize the Lagrangian:



$$L_{v,N_e}[n] = E_v[n] - \mu \left[ \int n(\mathbf{r}) d^3r - N_e \right] \quad (6.3.5)$$

Where the value of  $\mu$  is varied until the constraint is respected. The final minimizing density obeys:

$$0 = \frac{\delta L_{v,N_e}[n]}{\delta n(\mathbf{r})} = \frac{\delta E_v[n]}{\delta n(\mathbf{r})} - \mu = \frac{\delta F_{HK}[n]}{\delta n(\mathbf{r})} + v(\mathbf{r}) - \mu \quad (6.3.6)$$

This is the basic equation of DFT. It has no direct mention of the wavefunction. Once we find an approximation for  $E_v[n]$  we can get an approximation for  $n_v(\mathbf{r})$  from this equation.

The two theorems of HK put some rigor into the Thomas-Fermi approximation. In this KH theory  $E_{TF}[n]$  is an approximation to  $E_v[n]$  and the TF equations are an application of (6.3.5). Still, we know that TF theory is very poor for chemistry. This means that despite the added rigor, the TF approximation is too cumbersome for quantum chemistry.

#### ***D. An interesting observation on the variational principle of non-interacting electrons***

Consider a system of  $N$  non-interacting particles in a potential  $v(\mathbf{r})$ . Usually we may assume that the ground state of this system is a Slater wave function  $\Phi = \det[\phi_1 \dots \phi_N]$ . The variational theorem states that their ground state energy is given by minimizing:

$$E[v] = \sum_{n=1}^N \langle \phi_n | \hat{T}_1 + \hat{v} | \phi_n \rangle - \sum_{n=1}^N \epsilon_n (\langle \phi_n | \phi_n \rangle - 1) \quad (6.4.1)$$

Where  $\hat{T}_1 = -\frac{1}{2}\nabla^2$ ,  $\epsilon_n$  are Lagrange multipliers imposing the unit norm of each orbital. Minimizing leads to the equations:

$$(\hat{T}_1 + \hat{v})\phi_n = \epsilon_n \phi_n \quad (6.4.2)$$

After solving for the  $N$  lowest energy eigenstates the energy is  $E[v] = \sum_{n=1}^N \epsilon_n$ , the kinetic energy is  $T[v] = \sum_{n=1}^N \langle \phi_n | \hat{T}_1 | \phi_n \rangle$  the density is  $n(\mathbf{r}) = \sum_{n=1}^N |\phi_n(\mathbf{r})|^2$ .

Now consider a different problem. Suppose the density is given and one is required to find the system with this density that has the minimum kinetic energy  $T$ . This leads to the Lagrangian:

$$T[n] = \sum_{n=1}^N \langle \phi_n | \hat{T}_1 | \phi_n \rangle + \int v(\mathbf{r}) \left( \sum_{n=1}^N |\phi_n(\mathbf{r})|^2 - n(\mathbf{r}) \right) - \sum_{n=1}^N \epsilon_n (\langle \phi_n | \phi_n \rangle - 1) \quad (6.4.3)$$

Now  $v(\mathbf{r})$  are Lagrange multipliers and must be searched for in order that the constraint  $n(\mathbf{r}) = \sum_{n=1}^N |\phi_n(\mathbf{r})|^2$  be fulfilled. After minimizing we obtain the equations:

$$(\hat{T}_1 + \hat{v})\phi_n = \epsilon_n \phi_n \quad (6.4.4)$$

This equation rises from the attempt to compute  $T[n]$  and is the same as Eq. (6.4.2) which rose as when  $v$  was given and  $E[v]$  was calculated. This shows that minimizing the kinetic energy under a given density invokes “the same” equations as minimizing the energy when the potential is given. This fact will be used in the Kohn-Sham method.

## ***E. The set of V-representable densities***

### **i. V-rep densities correspond to ground states wave function of some potential well**

The HK theorem shows that the ground state density of a system uniquely determines the one body potential. This is a uniqueness statement: there is at

most one potential associated with a density. An interesting twist is the reverse question: what are the conditions that a given density is the ground state density (GSD) of some system? Of course there are some preliminary conditions on the density: it must be non-negative and it must integrate to a positive integer:

$$\begin{aligned} n(\mathbf{r}) &\geq 0 \\ \int n(\mathbf{r})d^3r &= N_e \end{aligned} \tag{6.5.1}$$

But in general, we have no good criterion for deciding whether a given density  $n(\mathbf{r})$  is a GSD of some potential  $v(\mathbf{r})$ . Densities which are GSDs of a potential are called "v-representable". In short v-rep.

When we say that everything is a functional of the density, we mean everything is a functional of a v-rep density.

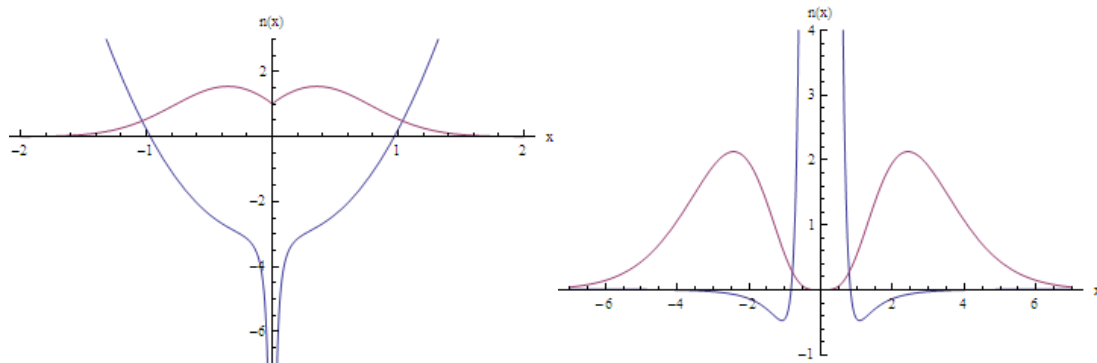
## ii. Some non-v-representability issues

We have seen in XXX that ground state wave functions of single particles is nodeless. A corollary from the above analysis is that a density of one particle with a node is not v-representable.

However, the density does not need to actually develop a zero for the density to be non-vrep. Consider the example by Englisch and Englisch:

$$n(x) \propto (1 + (x^2)^\alpha)^2 e^{-\sqrt{x^2}} \quad \frac{1}{4} \leq \alpha < \frac{1}{2}$$

If  $\psi(x)$  is a wavefunction then the potential is given by  $v(x) = \frac{1}{2} \frac{\psi''(x)}{\psi(x)}$ . Using the above form for  $\alpha = \frac{1}{3}$ , we find that the potential is infinite at the origin:



**Figure VI-2: Examples, following Englisch and Englisch, of a non-vrep density (in purple) and the corresponding potential which is singular. The left panel is a 1D example which the right panel is the radial 3D example.**

In 3 dimensions simply replace  $x$  by  $r$  and  $\psi''(x)$  by  $\nabla^2\psi(r) = \frac{1}{r^2}(r^2\psi(r))''$

thus:  $v(r) = \frac{1}{2} \frac{(r^2\psi(r))''}{r^2\psi(r)} - \frac{1}{r^2}$ . The first term of the potential is seen to be dominated by a term going as  $\frac{1}{r^2}$  when  $r \rightarrow 0$ . This is a centrifugal barrier with  $\ell = 1$ . Subtracting this barrier exposes the bare potential, shown in Figure VI-2.

### iii. The set of v-rep densities of a given electron number is not convex

Degenerate Hamiltonians can generate non-vrep densities quite easily. Thus, non-vrep densities are much more abundant than one may suspect. Let us see why there is a problem, in the following analysis due to Levy and Lieb (developed separately at the same time more or less).

Suppose  $\hat{H} = \hat{T} + \int v(\mathbf{r})\hat{n}(\mathbf{r})d^3r + \hat{U}$  is a Hamiltonian with degenerate ground states, of energy  $E$  and full degeneracy  $Q$ . Thus,  $\Psi_i$ ,  $i = 1, \dots, Q$  are  $Q$  ground state wave functions with  $\langle \Psi_i | \Psi_j \rangle = \delta_{ij}$ . There are infinitely many ways to define the  $\Psi_i$ . We select one arbitrarily. Denote:  $n_i(\mathbf{r}) = \langle \Psi_i | \hat{n}(\mathbf{r}) | \Psi_i \rangle$ . Then, consider the density built as a convex sum of these degenerate-state densities:

$$n(\mathbf{r}) = \sum_{i=1}^Q c_i n_i(\mathbf{r}) \quad c_i > 0 \quad \sum_{i=1}^Q c_i = 1 \quad (6.5.2)$$

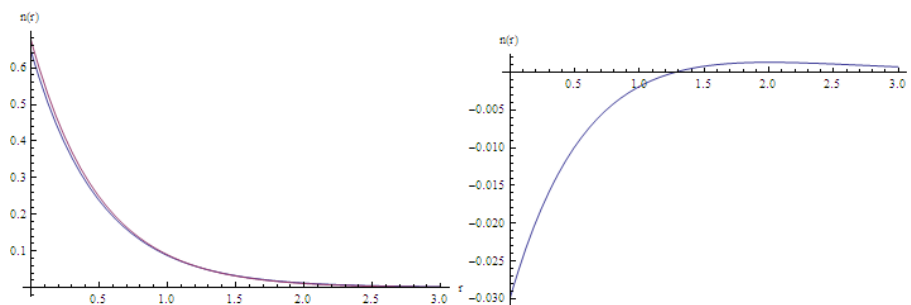
We now show that this density is usually non-v-representable. Of course, it may happen (in rare cases) that a linear combination of the degenerate wave functions  $\Psi = \sum_{i=1}^Q b_i \Psi_i$  yields  $n(\mathbf{r}) = \langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle$ . This case however is not ordinary and we consider the cases where this does not happen. We thus proceed to show that  $n(\mathbf{r})$  cannot be the groundstate density of *any other Hamiltonian* as well. We do this by contradiction and assume existence of a wavefunction  $\Psi'$  which is the GS of some  $\hat{H}'$  and such that  $\langle \Psi' | \hat{n}(\mathbf{r}) | \Psi' \rangle = n(\mathbf{r})$ . The variational principle states  $\langle \Psi_i | \hat{H}' | \Psi_i \rangle > \langle \Psi' | \hat{H}' | \Psi' \rangle$  for each  $i$ . Multiply by the *positive*  $c_i$  (so the inequality is not spoiled) and sum over  $i$ , using  $\sum_i^Q c_i = 1$ , and obtain  $\sum_i^Q c_i \langle \Psi_i | \hat{H}' | \Psi_i \rangle > \langle \Psi' | \hat{H}' | \Psi' \rangle$ . Now use the same reasoning that led to Eq. (6.1.7) and replace  $\hat{H}'$  by  $\hat{H}$ , obtaining:  $\sum_i^Q c_i \langle \Psi_i | \hat{H} | \Psi_i \rangle > \langle \Psi' | \hat{H} | \Psi' \rangle$ . However,  $\langle \Psi_i | \hat{H} | \Psi_i \rangle = E$  and so this leads to  $\langle \Psi' | \hat{H} | \Psi' \rangle < E$  which contradicts of the variational principle for  $\Psi_i$ . Hence,  $\Psi'$  cannot be the ground-state of any Hamiltonian.

---

Example: The above theorem is general and holds for any system of particles. We consider non-interacting electrons. Consider the density of non-interacting electrons in the potential well created by a Lithium nucleus. The ground state is 4-fold degenerate  $\Psi_1 = |1s\bar{1}s2s|$ ,  $\Psi_2 = |1s\bar{1}s2p_x|$ ,  $\Psi_3 = |1s\bar{1}s2p_y|$ ,  $\Psi_4 = |1s\bar{1}s2p_z|$  with  $\psi_{1s} = \frac{e^{-r}}{\sqrt{\pi}}$ ,  $\psi_{2s} = \frac{e^{-\frac{r}{2}}(r-2)}{\sqrt{32\pi}}$ ,  $\psi_{2p_x} = \frac{e^{-\frac{r}{2}}x}{\sqrt{32\pi}}$ ,  $\psi_{2p_y} = \frac{e^{-\frac{r}{2}}y}{\sqrt{32\pi}}$ ,  $\psi_{2p_z} = \frac{e^{-\frac{r}{2}}z}{\sqrt{32\pi}}$ . The 4 densities are:  $n_1 = 2n_{1s} + n_{2s}$ ,  $n_2 = 2n_{1s} + n_{2p_x}$ ,  $n_3 = 2n_{1s} + n_{2p_y}$ ,  $n_4 = 2n_{1s} + n_{2p_z}$ . The average is given by:

$$n(r) = \frac{1}{4} \sum_{i=1}^4 n_i(r) = 2 \frac{e^{-2r}}{\pi} + \frac{e^{-r}}{128\pi} ((r-2)^2 + r^2)$$

This density is spherically symmetric and is plotted below, together with the density  $n_1(r)$  for comparison:



**Figure VI-3: Left panel: The average density of  $n(r)$  plotted together with the almost identical density  $n_1(r)$  orbital. Both densities seem almost identical yet the latter is v-rep while the former not. Right Panel: The difference between the two densities**

This density is very naïve and there is no visible indications suggesting this is not (non-interacting) v-representable.

We have found that such convex sum densities are not v-representable in the sense that there is no Hamiltonian for which a ground-state yields the density. However, clearly these densities are associated with the potential  $v(r)$  which created the degenerate states. One can extend the Hohenberg-Kohn functional definition to deal with these convex sum densities in the following way. Given a density which is not v-representable in the usual way, we now call it non-pure-state v-representable. Then we will say that such a density is ensemble-v-representable and assume it is associated with the potential  $v(r)$  from which the degenerate ground-states  $\Psi_i$  were formed. Then for such a density the HK functional is written as:

$$F_{HK}[n] = \sum_{i=1}^q c_i \langle \Psi_i | \hat{T} + \hat{U} | \Psi_i \rangle \quad n(\mathbf{r}) = \sum_{i=1}^q c_i \langle \Psi_i | \hat{n}(\mathbf{r}) | \Psi_i \rangle \quad (6.5.3)$$

Here, we implicitly made the identification  $n \rightarrow v \rightarrow c_i$ . The HK uniqueness theorem guaranteeing that only one potential  $v$  can form a given  $n$  can be extended also to this case (i.e. that  $n \rightarrow v$  is meaningful) still holds.

Furthermore, the minimum principle, i.e. that  $E_v[n]$  with the more general  $F_{HK}$  obtains its minimum at the density corresponding to  $v(\mathbf{r})$  remains true.

---

### ***F. Levy-Lieb generalization of the HK functional***

The minimum principle of HK is of crucial importance for density functional theory. Yet, the basic equation (Eq. (6.3.6)) which is the basis for development of DFT into a practical approach is somewhat problematic from the mathematical point of view. In order to define the functional derivative, one must ensure that there is an “open” neighborhood of the ground state density  $n_v(\mathbf{r})$  for which  $F_{HK}[n]$  can be defined for *any* density. As we saw, there is a fundamental problem is that of v-representability. The ensemble v-representability solved one kind of problems but it is not clear if there may be additional classes of non-v-representability. It is not possible to assume that any density that is positive and integrates to an integer is a density of a ground state wave function of a Hamiltonian. Furthermore, we cannot even assume that around a v-representable density there is an “open neighborhood” of v-representable densities.

In general one cannot waive the possibility that there may be densities which are not the ground state densities of any system and still are arbitrarily close to any  $n_v(\mathbf{r})$  of interest. In such a case the functional derivative of  $F_{HK}$  is formally undefined.

A way around these problems, developed separately by Levy and by Lieb is to formulate a functional  $F_{LL}[n]$  defined for *any* density on the one hand and equal to  $F_{HK}[n]$  for v-representable densities on the other which still allows for the same type of minimum principle as does the HK functional. We first

note an important property of  $F_{HK}[n]$ . We say that the wave function  $\Psi$  realizes the density  $n(\mathbf{r})$  (and symbolize this as  $\Psi \rightarrow n$ ) if:

$$\langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle = n(\mathbf{r}) \quad \Leftrightarrow \quad \Psi \rightarrow n \quad (6.5.4)$$

We will now show that for any wavefunction  $\Psi$  which realizes  $n(\mathbf{r})$

$$\Psi \rightarrow n \quad \Rightarrow \quad F_{HK}[n] \leq \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle \quad (6.5.5)$$

This is a direct result of the variational principle:  $F_{HK}[n] = E_{gs}[n] - \int v(\mathbf{r})n(\mathbf{r})d^3r \leq \langle \Psi | \hat{H} | \Psi \rangle - \int v(\mathbf{r})n(\mathbf{r})d^3r = \langle \Psi | \hat{H} - \hat{V} | \Psi \rangle = \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle$ .

From this development we see also that the wave-function which realizes this minimum is the ground-state wave function and is denoted  $\Psi_{gs}[n]$ . One has therefore:

$$F_{HK}[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle \quad (6.5.6)$$

Levy and Lieb decided to use this relation, valid only for v-rep densities as the *definition* of their functional, valid for any density:

$$F_{LL}[n] \equiv \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle \quad (6.5.7)$$

Exercise: Show that for any density there is at least one wave function which realizes it

Solution: For a one-electron case this is trivial, since the density is non-negative we can take the wave function as  $\Psi(\mathbf{r}) = \sqrt{n(\mathbf{r})}$ . For  $N$  electrons, just "slice"  $n(\mathbf{r})$  to  $N$  non-overlapping parts each: positive, integrating to 1: (1)  $n(\mathbf{r}) = \sum_{k=1}^N n_k(\mathbf{r})$ , (2)  $n_k(\mathbf{r}) > 0$ , (3)  $\phi_i(\mathbf{r})\phi_j(\mathbf{r}) = \delta_{ij}n_j(\mathbf{r})$  where  $\phi_k(\mathbf{r}) = \sqrt{n_k(\mathbf{r})}$ , and (4)  $\int n_k(\mathbf{r}) = 1$ . This can be done in countless ways (with a good sharp knife). Then the determinant  $\Psi = |\phi_1\phi_2 \dots \phi_{N_e}|$ , is a wavefunction which creates  $n(\mathbf{r})$ .



The above exercise shows, that the definition in Eq. (6.5.7) makes sense for practically all densities. One can now use  $F_{LL}[n]$  instead of  $F_{HK}[n]$  in Eq. (6.3.1) and convert the Hohenberg-Kohn minimum principle to a variational-minimum principle. The search for the constrained minimum can be done using a Lagrange multiplier approach. We formulate the Lagrangian:

$$\tilde{F}_{LL}[\Psi, n, v] \equiv \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle + \int v(\mathbf{r}) [\langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle - n(\mathbf{r})] d^3r \quad (6.5.8)$$

Given  $n(\mathbf{r})$ , one now find suitable Lagrange multipliers  $v(\mathbf{r})$  such that when minimizing  $F_{LL}$  with respect to an antisymmetric wave function  $\Psi$  leads to the constraint  $\langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle = n(\mathbf{r})$  at any  $\mathbf{r}$ . The search for a minimizing wave function will cease when the gradient is zero, so a necessary condition is:

$$\begin{aligned} \frac{\delta}{\delta \Psi} \tilde{F}_{LL}[\Psi, n, v] \Big|_{\Psi=\Psi_{LL}, v=v_{LL}} = 0 &= \left( \hat{T} + \hat{U} + \int v_{LL}(\mathbf{r}) \hat{n}(\mathbf{r}) d^3r \right) \Psi_{LL} \\ &= \hat{H}_{LL} \Psi_{LL} \end{aligned} \quad (6.5.9)$$

Thus, when  $n(\mathbf{r})$  is given, the Lagrange multiplier function  $v_{LL}(\mathbf{r})$  that imposes the condition defines a Hamiltonian  $\hat{H}_{LL} = \hat{T} + \int v_{LL}(\mathbf{r}) \hat{n}(\mathbf{r}) d^3r + \hat{U}$ . Note that  $v_{LL}(\mathbf{r})$  is determined by the procedure up to a constant. If  $n(\mathbf{r})$  is v-rep then necessarily, by Kohn's theorem,  $F_{LL}[n] = F_{HK}[n]$  and  $\Psi_{LL}$  is the groundstate wave function. However, we are not in general assured that  $\Psi_{LL}$  is always the ground state, since  $n(\mathbf{r})$  is not necessarily v-rep. Indeed, from (6.5.9), all we can say is that  $\Psi_{LL}$  must be some eigenstate of  $\hat{H}_{LL}$ .

Once the minimum is attained, one can take the functional derivative of  $F_{LL}$  by  $n(\mathbf{r})$ . As with any constrained minimization, when the derivative of the optimized solution is taken with respect to the constraints, one obtains the Lagrange multiplied (see (1.11.23)):

$$\frac{\delta}{\delta n(\mathbf{r})} \tilde{F}_{LL}[\Psi, n, v] = -v_{LL}(\mathbf{r}) \quad (6.5.10)$$

Now, once the constrained minimum  $\tilde{F}$  has been found, one can plug  $\tilde{F}$  into the HK scheme and consider the Lagrangian corresponding to  $E_v = \tilde{F}_{LL} + \int v(\mathbf{r})n(\mathbf{r})d^3r$ :

$$L_v[n] = \tilde{F}_{LL} + \int v(\mathbf{r})n(\mathbf{r})d^3r - \mu \left[ \int n(\mathbf{r})d^3r - N \right] \quad (6.5.11)$$

If now one searches for the density that minimizes this functional one immediately obtains:

$$0 = \frac{\delta L_v}{\delta n(\mathbf{r})} = -v_{LL}(\mathbf{r}) + v(\mathbf{r}) - \mu \quad (6.5.12)$$

Showing that the minimum of  $E_v[n]$  is obtained when the density found admits, as Lagrange multiplier  $v_{LL}(\mathbf{r})$ , the same potential as was give to start with  $v(\mathbf{r})$ , up to a constant.

Even with the LL functional one problem is still not solved: convexity. While the domain of definition is now finally convex, the functional has lost convexity. A counterexample for convexity can be given, using the degenerate system used above to prove that the domain of definition of the HK functional is non-convex. Indeed, for that system, which had a potential  $v(\mathbf{r})$  and degenerate eigensstates  $\Psi_i$  ( $i = 1, \dots, Q$ ) all of energy  $E$ :

$$\begin{aligned}
F_{LL} \left[ \sum_i c_i n_i \right] &= \langle \Psi_{LL} | \hat{T} + \hat{U} | \Psi_{LL} \rangle \\
&= \langle \Psi_{LL} | \hat{H} | \Psi_{LL} \rangle - \sum_i \int c_i n_i(\mathbf{r}) v(\mathbf{r}) d^3r \\
&\geq \sum_i c_i E - \int c_i n_i(\mathbf{r}) v(\mathbf{r}) d^3r \\
&= \sum_i c_i \left[ \langle \Psi_i | \hat{T} + \hat{U} + V | \Psi_i \rangle - \int n_i(\mathbf{r}) v(\mathbf{r}) d^3r \right] \\
&= \sum_i c_i F_{LL}[n_i]
\end{aligned} \tag{6.5.13}$$

Showing that for these densities the functional is concave (while for the case where  $n_i$  and  $\sum_i c_i n_i$  are all v-rep the functional is convex). Therefore overall the functional is not convex.

Once again, it is possible to circumvent this problem by extending the definition to ensemble densities.

---

Exercise: Discuss the Levy approach for a given density  $n(\mathbf{r})$  of a system of non-interacting electrons.

In this case the Hamiltonian is  $\hat{H} = \hat{T} + \int v(\mathbf{r}) \hat{n}(\mathbf{r}) d^3r$  i.e. a 1-body Hamiltonian. Suppose  $\psi_m$ ,  $m = 1, 2, \dots$  are the 1-particle states which are eigenstates of  $\hat{H}_1 = \hat{T}_1 + \int v(\mathbf{r}) \hat{n}(\mathbf{r}) d^3r$ :

$$\begin{aligned}
-\frac{\hbar^2}{2m_e} \nabla^2 \phi_m(\mathbf{r}) + v(\mathbf{r}) \phi_m(\mathbf{r}) \\
= \epsilon_m \phi_m(\mathbf{r})
\end{aligned} \tag{6.5.14}$$

Let us suppose orbital energy ordering, so that  $\epsilon_1 \leq \epsilon_2 \leq \epsilon_3 \dots$ . If the density is "non-interacting v-rep" then it must be due to the Slater wave function  $D_{gs} = |\phi_1 \dots \phi_{N_e}|$  of the lowest energy  $N_e$  orbitals. In this case, the kinetic energy is given by:

$$T_{LL}[n] = \sum_{m=1}^{N_e} \langle \phi_m | \hat{T} | \phi_m \rangle \tag{6.5.15}$$

And the functional derivative is:

$$\frac{\delta}{\delta n(\mathbf{r})} T_{LL}[n] = -v(\mathbf{r}) \quad (6.5.16)$$

If we are given a non-interacting non-v-rep density then we might find not a ground state but an excited determinant, for example:  $D_{ex} = |\phi_1 \cdots \phi_{N_e-1} \phi_{N_e+1}|$ . This leads to “holes” in the non-interacting system.

### ***G. The dilation inequality for the HK functional***

We have already seen the the HF functional is convex. More exact properties are desirable, so that when one derives approximations to this functional, they can perhaps be forced to obey the known exact relations. One approach is to use dilation considerations. This will lead to an interesting inequality.

We will use the following symbol convention:

$$\begin{aligned} \Psi_\gamma(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) &\equiv \gamma^{3N_e/2} \Psi(\gamma\mathbf{r}_1, \gamma\mathbf{r}_2, \dots, \gamma\mathbf{r}_{N_e}) \\ n_\gamma(\mathbf{r}) &\equiv \gamma^3 n(\gamma\mathbf{r}) \end{aligned} \quad (6.6.1)$$

It is quite straightforward to show the following relations:

#### *Exercise*

Derive the following relation

$$\begin{aligned} \langle \Psi_\gamma | \hat{n}(\mathbf{r}) | \Psi_\gamma \rangle &= \gamma^3 \langle \Psi | \hat{n}(\gamma\mathbf{r}) | \Psi \rangle \\ \langle \Psi_\gamma | \hat{T} | \Psi_\gamma \rangle &= \gamma^2 \langle \Psi | \hat{T} | \Psi \rangle \\ \langle \Psi_\gamma | \hat{U} | \Psi_\gamma \rangle &= \gamma \langle \Psi | \hat{U} | \Psi \rangle \end{aligned} \quad (6.6.2)$$

From the first relation in this equation, one can deduce that if  $\Psi \rightarrow n(\mathbf{r})$  (i.e.  $\Psi$  is a many-electron wave function exhibiting the spatial density  $n(\mathbf{r})$ ) then:

$$\langle \Psi_\gamma | \hat{n}(\mathbf{r}) | \Psi_\gamma \rangle = \gamma^3 \langle \Psi | \hat{n}(\gamma\mathbf{r}) | \Psi \rangle = \gamma^3 n(\gamma\mathbf{r}) = n_\gamma(\mathbf{r}) \quad (6.6.3)$$

One can now also use this relation to get a dilation inequality. Indeed, since  $\Psi[n]_\gamma$  realizes the density  $n_\gamma(\mathbf{r})$  (see the first of Eqs. (6.6.2)) we have

$F_{HK}[n_\gamma] \leq \langle \Psi[n]_\gamma | \hat{T} + \hat{U} | \Psi[n]_\gamma \rangle = \gamma^2 \langle \Psi[n] | \hat{T} | \Psi[n] \rangle + \gamma \langle \Psi[n] | \hat{U} | \Psi[n] \rangle$ , thus we find:

$$F_{HK}[n_\gamma] \leq \gamma^2 T[n] + \gamma U[n] \quad (6.6.4)$$

## VII. The Kohn-Sham method

Kohn and Sham noticed that the HK theory is valid for both interacting and non-interacting electrons. Now, they ask, what happens if for any system of interacting electrons, with density  $n$  there is a non-interacting system of the same density? It is clear that if such two systems exist they are unique. The non-interacting system has one advantage over the interacting system: we can find its ground-state rather easily, since the many-body wavefunction is a Slater wave function. So, the problem is: how to perform such a mapping.

### A. Non-interacting electrons

If non-interacting electrons are tractable, let then study their density functionals. First, if we are given  $n(\mathbf{r})$  we assume it is non-interacting  $v$ -representable. That is we assume there exists a potentials  $v_s(\mathbf{r})$  such that the ground-state of the resulting non-interacting Hamiltonian  $\hat{H}_s = \hat{T} + \hat{V}_s$  admits a ground-state having  $n(\mathbf{r})$  as density. We can later use the Levy-Lieb approach to patch up the  $v$ -representability requirement.

Let us denote the ground-state wave function of non-interacting electrons that have the density  $n(\mathbf{r})$  by  $\Phi_{gs}[n]$ . The Hohenberg-Kohn functional for non-interacting electrons is reduced to just the kinetic energy, i.e. we define:

$$T_S[n] = \langle \Phi_{gs}[n] | \hat{T} | \Phi_{gs}[n] \rangle \quad (7.1.1)$$

An interesting question is – can we devise a method to compute  $T_S[n]$ . For most densities, the non-interacting wave function is simply a Slater wave function. The minimum principle for  $T_S[n]$  (derived from Eq. (6.5.5)) is:

$$\Psi \rightarrow n \quad \Rightarrow \quad T_S[n] \leq \langle \Psi | \hat{T} | \Psi \rangle \quad (7.1.2)$$

One should not treat this equation lightly and realize its non-trivial consequences: that the ground-state wave function of non-interacting particles associated with  $n(\mathbf{r})$  minimizes the kinetic energy! Let us see some consequence this minimum principle.

Let  $\Phi_{gs}[n]$  be the ground state wave function of the system of non interacting electrons realizing the density  $n(\mathbf{r})$ . From the first equation in Eq. (6.6.2) one sees that  $\Phi_{gs}[n]_\gamma$  realizes  $n_\gamma(\mathbf{r})$ , thus one can plug it into the right hand side of Eq. (7.1.2), with  $n_\gamma(\mathbf{r})$  plugged into  $T_S$ :

$$T_S[n_\gamma] = \langle \Phi_{gs}[n_\gamma] | \hat{T} | \Phi_{gs}[n_\gamma] \rangle \leq \langle \Phi_{gs}[n]_\gamma | \hat{T} | \Phi_{gs}[n]_\gamma \rangle \quad (7.1.3)$$

Since now the right hand side is an expression of a scaled wave-function, one can use Eq. (6.6.2) and obtain:

$$T_S[n_\gamma] \leq \langle \Phi_{gs}[n]_\gamma | \hat{T} | \Phi_{gs}[n]_\gamma \rangle = \gamma^2 \langle \Phi_{gs}[n] | \hat{T} | \Phi_{gs}[n] \rangle = \gamma^2 T_S[n] \quad (7.1.4)$$

And so  $T_S[n_\gamma] \leq \gamma^2 T_S[n]$ . But we can also use this equation to scale  $n_\gamma$  back to  $n$  by dilating by  $1/\gamma$ , since:  $(n_\gamma)_{1/\gamma} = n$ . Then, the same rule applies and we get:  $T_S[n] \leq \gamma^{-2} T_S[n_\gamma]$ , and so  $T_S[n_\gamma] \geq \gamma^2 T_S[n]$ . We obtained two contradicting equations which can agree only if both are reduced to equality. Thus:

$$T_S[n_\gamma] = \gamma^2 T_S[n]. \quad (7.1.5)$$

This should be compared with the analogous result of Eq. (6.6.4) which is an inequality. The dilation effects for non-interacting electrons is obviously much simpler. One important corollary of (7.1.5) is that the non-interacting wave

functions scale with the density. This can be seen if one remembers that  $\Phi_{gs}[n]$  minimizes the kinetic energy. Therefore any function  $\Phi[n]$  for which  $\langle \Phi[n] | \hat{T} | \Phi[n] \rangle$  equals  $T_S[n]$  must be the ground state of the non-interacting system with density  $n$ . We saw that  $\Phi_{gs}[n]_\gamma$  gives the correct kinetic energy of for the system with density  $n_\gamma$ ,  $T_S[n_\gamma]$ , and therefore, necessarily:

$$\Phi_{gs}[n]_\gamma = \Phi_{gs}[n_\gamma]. \quad (\text{non - interacting particles}) \quad (7.1.6)$$

### Exercise

- 1) We can define a functional called the Hartree energy, which is the classical electrostatic energy associated with the charge distribution  $n(\mathbf{r})$ :

$$E_H[n] = \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.1.7)$$

Prove the following dilation relation:

$$E_H[n_\gamma] = \gamma E_H[n] \quad (7.1.8)$$

- 2) Now define the exchange energy functional (see also Eq. (4.1.17) for a definition based on the orbitals):

$$K[n] \equiv \langle \Phi_{gs}[n] | \hat{U} | \Phi_{gs}[n] \rangle - E_H[n] \quad (7.1.9)$$

Use (7.1.6). Prove:

$$K[n_\gamma] = \gamma K[n] \quad (7.1.10)$$

What is the relation between the potential  $v_S(\mathbf{r})$  for which  $n(\mathbf{r})$  is a non-interacting ground state density and  $v_{S,\gamma}(\mathbf{r})$  for which  $n_\gamma(\mathbf{r})$  is a non-interacting ground state? We can use the basic DFTEquation to answer this.

From the basic definition of functional derivation:

$$\mu_\gamma - v_{S,\gamma}(\mathbf{r}') = \left. \frac{\delta T_S}{\delta n(\mathbf{r}')} \right|_{n_\gamma} = \lim_{\eta \rightarrow 0} \frac{(T_S[n_\gamma(\mathbf{r}) + \eta \delta(\mathbf{r} - \mathbf{r}')] - T_S[n_\gamma(\mathbf{r})])}{\eta} \quad (7.1.11)$$

The 3D delta-function has the density dilation structure:  $\delta(\mathbf{r} - \mathbf{r}') = \lambda^3 \delta(\lambda \mathbf{r} - \lambda \mathbf{r}')$ , so:

$$\mu_\gamma - v_{S,\gamma}(\mathbf{r}') = \lim_{\eta \rightarrow 0} \frac{(T_S[\gamma^3(n(\gamma \mathbf{r}) + \eta \delta(\gamma \mathbf{r} - \gamma \mathbf{r}'))] - T_S[\gamma^3 n(\gamma \mathbf{r})])}{\eta} \quad (7.1.12)$$

Then using Eq.:

$$\begin{aligned} \mu_\gamma - v_{S,\gamma}(\mathbf{r}') &= \lim_{\eta \rightarrow 0} \frac{\gamma^2 (T_S[n(\mathbf{r}) + \eta \delta(\mathbf{r} - \gamma \mathbf{r}')] - T_S[n(\mathbf{r})])}{\eta} \\ &= \gamma^2 (\mu - v_S(\gamma \mathbf{r}')) \end{aligned} \quad (7.1.13)$$

By dilating in the reverse direction we can easily see that:

$$\begin{aligned} \mu_\gamma &= \gamma^2 \mu \\ v_{S,\gamma}(\mathbf{r}') &= \gamma^2 v_S(\gamma \mathbf{r}') \end{aligned} \quad (7.1.14)$$

We could have obtained this result directly from the Schrödinger equation. Suppose  $\Psi(\{\mathbf{r}\})$  is a many-body eigenfunction of Hamiltonian  $\hat{H} = \hat{T} + V(\{\mathbf{r}\})$ , i.e.  $\hat{T}\Psi = (E - V)\Psi$ . Define a scaled wavefunction:

$$\Psi_\gamma(\mathbf{r}_1, \dots, \mathbf{r}_N) = \gamma^{-\frac{3N}{2}} \Psi(\gamma \mathbf{r}_1, \dots, \gamma \mathbf{r}_N) \quad (7.1.15)$$

Then clearly:

$$\hat{T}\Psi_\gamma(\{\mathbf{r}\}) = \gamma^2 \gamma^{-\frac{3N}{2}} (\hat{T}\Psi)(\{\gamma \mathbf{r}\}) = \gamma^2 (E - V(\{\gamma \mathbf{r}\})) \Psi_\gamma(\{\mathbf{r}\}) \quad (7.1.16)$$

And so  $\Psi_\gamma$  obeys the S. E.  $\hat{T}\Psi_\gamma = (E_\gamma - V_\gamma)\Psi_\gamma$  with energy and potential given by:

$$\begin{aligned} V_\gamma(\{\mathbf{r}\}) &= \gamma^2 V(\{\gamma \mathbf{r}\}) \\ E_\gamma &= \gamma^2 E \end{aligned} \quad (7.1.17)$$



For a ground state of non-interacting the first equation means that  $v_{s,\gamma}(\mathbf{r}) = \gamma^2 v_s(\gamma\mathbf{r})$  is the one-particle potential for the scaled determinant, and thus for the scaled density (since the scaled determinant realizes the scaled density).

### ***B. Orbitals for the non-interacting electrons***

Consider the ground state wave function for the non-interacting electrons  $\Phi_{gs}[n]$ . Suppose the non-interacting electrons reside in the potential  $v_s$  (the subscript  $S$  is in honor of Slater) then  $(\hat{T} + \hat{V}_S)\Phi_{gs}[n] = E_S\Phi_{gs}[n]$ . In most cases this wave function is a normalized Slater wave function. The only exception might be the existence of a degeneracy and then there may be several independent Slater wave functions and  $\Phi_{gs}$  is a linear combination. However, one can always *assume* that he is looking for one of the determinants and not for a linear combination (this means that one must be careful not to impose possible symmetries of the Hamiltonian). Once one does that, we can introduce  $N_e$  orthonormal spin-orbitals  $\phi_q(\mathbf{x})$   $q = 1 \dots N_e$ , from which  $\Phi_{gs}[n]$  is built. These orbitals are excited states of the potential well in which the non-interacting electrons reside. Thus the orbitals must each obey the single-electron Schrödinger equation:

$$-\frac{\hbar^2}{2m_e}\nabla^2\phi_q(\mathbf{x}) + v_s(\mathbf{r})\phi_q(\mathbf{x}) = \epsilon_q\phi_q(\mathbf{x}) \quad (7.1.18)$$

The orbitals correspond to the  $N_e$  lowest eigenvalues  $\epsilon_q$ . The fact that  $\Phi_{gs}$  realizes the density  $n(\mathbf{r})$  is expressed as :

$$n(\mathbf{r}) = \sum_{q=1}^{N_e} |\phi_q(\mathbf{x})|^2 \quad (7.1.19)$$

The non-interacting kinetic energy  $T_S[n]$  is then:

$$T_S[n] = \sum_{q=1}^{N_e} \int \phi_q(\mathbf{x}) \left( -\frac{\hbar^2}{2m_e} \nabla^2 \right) \phi_q(\mathbf{x}) d\mathbf{x} \quad (7.1.20)$$

When one wants to find the functional derivative of  $T_S$  with respect to  $n(\mathbf{r})$  one can turn to basic equation, Eq. (6.3.6) which in case of non-interacting electrons becomes:

$$\frac{\delta T_S[n]}{\delta n(\mathbf{r})} + v_S(\mathbf{r}) = \mu \quad (7.1.21)$$

We will see that this equation is important for the method known as the Kohn-Sham method.

Note that the discussion of dilation in the previous subsection can be carried on to the orbitals. The only additional information to convey is that the orbitals scale as the density and each of the orbital energies scale as the total energy:

$$\epsilon_{q,\gamma} = \gamma^2 \epsilon_q \quad (7.1.22)$$

### ***C. The correlation energy functional: definition and some formal properties***

The ground state energy of the an system of electrons in density  $n(\mathbf{r})$  can be written in terms of the non-interacting (Slater wave function) wave function  $\Phi_{gs}[n]$ :

$$E[n] = \langle \Phi_{gs}[n] | \hat{H} | \Phi_{gs}[n] \rangle + E_C[n] \quad (7.2.1)$$

This equation is actually a definition of a new density functional, the correlation energy functional  $E_C[n]$ . If we suppose for the time being that  $E_C[n]$  is known, this expression can be used to define a working procedure for DFT known as the Kohn-Sham method. From our studies of the Hartree-

Fock theory, we know that the expectation value of  $\hat{H}$  within a determinant can be written as:

$$\langle \Phi_{gs}[n] | \hat{H} | \Phi_{gs}[n] \rangle = E_H[n] + K[n] \equiv U_S[n] \quad (7.2.2)$$

Where  $E_H$  is the Hartree energy:

$$E_H[n] = \frac{e^2}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.2.3)$$

And  $K$  is given in terms of the orbitals from which  $\Phi_{gs}[n]$  is composed (Eq. (4.1.17)):

$$K[n] = -\frac{e^2}{2} \iint \frac{|\rho(\mathbf{x}, \mathbf{x}')|^2}{|\mathbf{r} - \mathbf{r}'|} d^3x d^3x' \quad (7.2.4)$$

Based on this, we rewrite  $F_{HK}[n]$  as:

$$F_{HK}[n] = \langle \Psi_{gs}[n] | \hat{F} | \Psi_{gs}[n] \rangle = \langle \Phi_{gs}[n] | \hat{F} | \Phi_{gs}[n] \rangle + E_C[n] \quad (7.2.5)$$

This allows us to write Eq. (7.2.5) compactly as:

$$F_{HK}[n] = T_S[n] + U_S[n] + E_C[n] \quad (7.2.6)$$

Clearly we have also the equivalent equation:

$$E_C[n] = T[n] - T_S[n] + U[n] - U_S[n] \quad (7.2.7)$$

Physical intuition concerning molecules and solids tells us that  $E_C[n]$  is a small quantity when compared to  $T_S[n]$  or  $U_S[n]$ . Thus, it is reasonable to look for approximations to this quantity. Approximation to the correlation energy functional is the most important issue in DFT. It is an open question, still being worked upon.

We shall deal with approximations later. Meanwhile, let us ask what can be safely said about the correlation functional. We prove here several important

inequalities. First, consider the difference  $T_C[n] \equiv T[n] - T_S[n]$ , the correlation kinetic energy.

### Exercise

Show  $T_C[n]$  is a positive quantity. This is actually intuitively clear: the interacting electrons must have much more complicated "paths" in the interacting case because they want to avoid "bumping into" other electrons. Anything with more swirls must have higher kinetic energy.

### Solution

Using the variational theorem:

$$\begin{aligned} T_S[n] + V_S[n] &= \langle \Phi_{gs}[n] | \hat{T} + \hat{V}_S | \Phi_{gs}[n] \rangle \leq \langle \Psi_{gs}[n] | \hat{T} + \hat{V}_S | \Psi_{gs}[n] \rangle \\ &= T[n] + V_S[n] \end{aligned} \quad (7.2.8)$$

Where  $\Psi_{gs}[n]$  is the interacting ground-state wave function determined by  $n$ . Comparing the two sides we have:

$$T_S[n] \leq T[n] \quad (7.2.9)$$

Next, we can show that the exchange correlation energy is always negative. This comes about from our experience with expectation values of determinants:

$$F_{HK}[n] + V[n] = E_{gs} \leq \langle \Phi_{gs}[n] | \hat{H} | \Phi_{gs}[n] \rangle = T_S[n] + V[n] + U_S[n] \quad (7.2.10)$$

Using Eq. (7.2.6) we find:

$$E_C[n] < 0 \quad (7.2.11)$$

Furthermore  $U_C[n] \equiv U[n] - U_S[n]$  is negative as can be seen from the fact that  $E_C$  is negative and  $T_C$  is positive:

$$U_C[n] = E_C[n] - T_C[n] \leq 0 \quad (7.2.12)$$

An additional property is the dilating relations. We have proved that  $F_{HK}[n_\gamma] \leq \gamma^2 T[n] + \gamma U[n]$  and  $T_S[n_\gamma] = \gamma^2 T_S[n]$ ,  $U_S[n_\gamma] = \gamma U_S[n]$ . Thus:

$$E_C[n_\gamma] = T_C[n_\gamma] + U_C[n_\gamma] \leq \gamma^2 T_C[n] + \gamma U_C[n] \quad (7.2.13)$$

One way to proceed is to substitute  $U_C$  with  $E_C - T_C$ . The other is to substitute  $T_C$  with  $E_C - U_C$ . We obtain from each possibility:

$$E_C[n_\gamma] \leq \gamma(\gamma - 1)T_C[n] + \gamma E_C[n] \quad (7.2.14)$$

$$E_C[n_\gamma] \leq \gamma^2 E_C[n] + \gamma(1 - \gamma)U_C[n]$$

And, since  $T_C$  is always positive and  $U_C$  negative, we find that:

$$E_C[n_\gamma] \leq \gamma E_C[n] \quad \gamma < 1 \quad (7.2.15)$$

$$E_C[n_\gamma] \leq \gamma^2 E_C[n] \quad \gamma < 1$$

Obviously, the second relation is contained in the first (since  $E_C < 0$ ) and so only the first relation is important; it can be used to derive complementary inequalities. Indeed, applying them for  $\gamma \rightarrow \gamma^{-1}$  we find:

$$E_C[n_{1/\gamma}] \leq \gamma^{-1} E_C[n] \quad \gamma > 1 \quad (7.2.16)$$

This holds for any  $n$ , so we stick in (7.2.16)  $n_\gamma$  instead of  $n$  and using the fact that  $(n_\gamma)_{1/\gamma} = n$  to obtain:

$$E_C[n_\gamma] \geq \gamma E_C[n] \quad \gamma > 1 \quad (7.2.17)$$

## ***D. The Kohn Sham equations***

### **i. The Kohn-Sham equation from a system of non-interacting problem**

Let us now turn to the issue of determining  $v_s(\mathbf{r})$ , required for the mapping between the interacting non-interacting systems. Let us start from the basic equation Eq. (6.3.6) which becomes, using Eq. (7.2.6):

$$\frac{\delta T_S[n]}{\delta n(\mathbf{r})} + v_{HX}(\mathbf{r}) + v_C(\mathbf{r}) + v(\mathbf{r}) = \mu \quad (7.3.1)$$

Where:

$$v_{HX}(\mathbf{r}) = \frac{\delta U_S[n]}{\delta n(\mathbf{r})} = v_H(\mathbf{r}) + v_X(\mathbf{r}) \quad (7.3.2)$$

is the Hartree-Exchange potential. Notice that:

$$v_H(\mathbf{r}) = \frac{\delta E_H[n]}{\delta n(\mathbf{r})} = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (7.3.3)$$

and:

$$v_X(\mathbf{r}) = \frac{\delta K[n]}{\delta n(\mathbf{r})} \quad (7.3.4)$$

Thus, from Eq. (7.1.21) and Eq. (7.3.1) (up to a constant):

$$v_S(\mathbf{r}) = v(\mathbf{r}) + v_{HX}(\mathbf{r}) + v_C(\mathbf{r}) \quad (7.3.5)$$

This equation gives us the potential of the non-interacting system. Thus, we have made a definite connection between the interacting and non-interacting systems.

Now, an important observation allows us to set up a simple method for obtaining the ground-state of an interacting system of electrons. We need to find a density that obeys two conditions:

- a) It is the density of the non-interacting electrons so it is the sum of square orbitals (Eq. (7.1.19)) that obey the Schrödinger equations (7.1.18) with potential  $v_S(\mathbf{r})$ .
- b) The potential  $v_S(\mathbf{r})$  must be related to the interacting system by Eq. (7.3.5)

This leads to a simple SCF procedure, called "The Kohn-Sham method" reminiscent of the Hartree-Fock algorithm:

1. Guess  $n(\mathbf{r})$ .
2. Build  $v_{HX}(\mathbf{r})$ ,  $v_C(\mathbf{r})$  from  $n(\mathbf{r})$  (Eq. (7.3.3) and (7.3.4)).
3. Obtain the orbitals  $\phi_q(\mathbf{x})$  from the lowest  $N_e$  eigenvalues of Eq. (7.1.18).
4. Compute the density from Eq. (7.1.19)
5. Redo from step 2 using the new density until you converge – i.e. until the density changes no more.

When the process converges, we have the exact ground-state density  $n_v(\mathbf{r})$ . It can be used to compute the ground-state energy by plugging it into the energy functional of Eq. (6.3.1):

$$E_v[n] = T_s[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r + U_S[n] + E_C[n] \quad (7.3.6)$$

## ii. Systems with partially occupied orbitals

Partially occupied orbitals arise in the realm of ensemble-DFT, i.e. when the density searched for is not “non-interacting  $v$ -representable”. We still assume the density is a sum of orbital densities as follows:

$$n(\mathbf{r}) = \sum_q |\phi_q(\mathbf{x})|^2, \quad (7.3.7)$$

but unlike Eq. (7.1.19) we do not impose exactly  $N_e$  orbitals. The orbitals are almost completely unconstrained, except for the Fermionic condition, that their norm must be not be great than 1:

$$\langle \phi_q | \phi_q \rangle = n_q \leq 1, \quad (7.3.8)$$

but the total number of electrons is still  $N_e$ :

$$\int n(\mathbf{r})d^3r = \sum_q n_q = N_e \quad (7.3.9)$$

And the kinetic energy is written similarly to Eq. (7.1.20):

$$T_S[n] = \sum_q \int \phi_q(\mathbf{x}) \left( -\frac{\hbar^2}{2m_e} \nabla^2 \right) \phi_q(\mathbf{x}) dx \quad (7.3.10)$$

We generalized Eqs. (7.1.19) and (7.1.20) any number of orbitals leading to the following orbital Lagrangian:

$$\begin{aligned}
L_{orb}[\{\phi_q\}, \{n_q\}] &= E_v \left[ \sum_q |\phi_q(\mathbf{x})|^2 \right] - \sum_q \epsilon_q (\langle \phi_q | \phi_q \rangle - n_q) \\
&\quad - \mu \left( \sum_q n_q - N_e \right)
\end{aligned} \tag{7.3.11}$$

$\epsilon_q$  ( $q = 1, 2, \dots$ ) and  $\mu$  are *Lagrange multipliers* assuring orbital normality and total number of particles. We minimize the Lagrangian with respect to the orbitals  $\phi_q(\mathbf{r})$  and the occupation numbers (constraint to be non-negative and not greater than 1). The constraint minimum procedure thus requires that:

$$\frac{\delta L_{orb}}{\delta \phi_q(\mathbf{x})} = 0 \quad \text{and} \quad \frac{\delta L_{orb}}{\delta n_q} = 0 \tag{7.3.12}$$

The first equation leads to the KS equations:

$$-\frac{\hbar^2}{2m_e} \nabla^2 \phi_q(\mathbf{x}) + v_S(\mathbf{r}) \phi_q(\mathbf{x}) = \epsilon_q \phi_q(\mathbf{x}) \tag{7.3.13}$$

For convenience, we order the indexing so that the series of orbital energies is ascending:  $\epsilon_1 \leq \epsilon_2 \leq \dots$ . The orbitals are now eigenfunctions of a Hermitian Hamiltonian, and so we can assume they are orthogonal:

$$\langle \phi_q | \phi_{q'} \rangle = n_q \delta_{qq'}, \tag{7.3.14}$$

We define as a short-hand notation  $v_S(\mathbf{r}) \equiv v(\mathbf{r}) + v_{HX}(\mathbf{r}) + v_C(\mathbf{r})$  (see Eq. (7.3.5)). By multiplying Eq. (7.3.13) by  $\phi_q(\mathbf{x})$  integrating on  $\mathbf{x}$  and summing on  $q$  we find:

$$T_S + \int v_S(\mathbf{r}) n(\mathbf{r}) d^3r = \sum_q \epsilon_q n_q \tag{7.3.15}$$

The second condition for minimum of  $L_{orb}$  holds only in cases that the minimum is attained with non-integer electron number,  $n_q < 1$ . The cases



$n_q = 1$  or  $n_q = 0$  are the boundary of the constraints and the derivative need not be zero there. From this second part of the equation, we have:

$$\epsilon_q = \mu, \quad 0 < n_q < 1 \quad (7.3.16)$$

This shows that all incompletely occupied orbitals have the same orbital energy equal to the chemical potential  $\mu$ , the Lagrange multiplier imposing the number of particle constraint of the Hohenberg-Kohn theory (Eq. (6.3.6)).

Typically, all orbitals with  $\epsilon_q < \mu$  are fully occupied ( $n_q = 1$ ) and those with  $\epsilon_q > \mu$  are fully unoccupied ( $n_q = 0$ ). The interacting electron energy is obtained using Eqs. (7.3.6) and (7.3.15) as

$$E_v = \sum_q \epsilon_q n_q + E_{HX}[n] + E_C[n] - \int n(\mathbf{r})(v_{HX}(\mathbf{r}) + v_C(\mathbf{r}))d^3r \quad (7.3.17)$$

We should note that the development here assumed non-interacting v-representability. In cases where this is not valid other occupation rules may apply.

### **iii. Is the ground state wave function of non-interacting particles always a Slater wave function**

The standard KS approach to the non-interacting kinetic energy is by defining the functional  $T_S$  as a minimum principle on the manifold of single determinants:

$$T_S[n] = \min_{\Phi \rightarrow n(\mathbf{r})} \langle \Phi | \hat{T} | \Phi \rangle \quad (7.3.18)$$

Setting up a Lagrangian and searching for the constrained minimum yields  $N$  occupied orbitals. If the density is not v-representable one or more of the low energy orbitals may have 0 occupation numbers.

Another way to define a non-interacting kinetic energy functional is by an extended minimum search over more general wave functions:

$$T_{NI}[n] = \min_{\Psi \rightarrow n(r)} \langle \Psi | \hat{T} | \Psi \rangle \quad (7.3.19)$$

Usually this search ends with  $\Psi$  being a single Slater wave function and for these case  $T_{NI}$  and  $T_S$  given are the same. Yet, this may not always be the case. Let us assume that we are searching through all  $\Psi = \cos \theta \Phi_A + \sin \theta \Phi_B$  where  $\Phi_A$  or  $\Phi_B$  are Slater wave functions. In general, the kinetic energy is then:

$$T_{NI}[n] = \min_{\Psi \rightarrow n(r)} \langle \Phi_A | \hat{T} | \Phi_A \rangle \cos^2 \theta + \sin 2\theta \langle \Phi_A | \hat{T} | \Phi_B \rangle + \langle \Phi_B | \hat{T} | \Phi_B \rangle \sin^2 \theta \quad (7.3.20)$$

Now, if  $\Phi_A$  and  $\Phi_B$  differ by only one orbital (say  $\phi_A^k = \phi_B^k$  for  $k = 1, \dots, N - 1$ ) then  $\Psi$  is actually a Slater wave function: its orbitals are the  $N - 1$   $\phi_A^k$ 's and then a new orbital  $\phi_N^{new} = \cos \theta \phi_A^N + \sin \theta \phi_B^N$  is added. Next, if  $\Phi_A$  and  $\Phi_B$  differ by two (or more) orbitals the cross term in (7.3.20) is zero and the kinetic energy is simply the sum of orbitals kinetic energies with occupation numbers given by  $\cos^2 \theta$  and  $\sin^2 \theta$ . The orbitals shared by the two determinants will have unity occupation number (since  $\cos^2 \theta + \sin^2 \theta = 1$ ) while orbitals in the  $A$  determinant but not in the  $B$  determinant will have occupation number  $\cos^2 \theta$  and those in the  $B$  determinant but not in the  $A$  determinant will have occupation number  $\sin^2 \theta$ . The orbitals coming out from the minimization will all solve a Schrödinger equation with the same potential. The 4 odd orbitals (2 from each determinant) will all have the same orbital energy equal to  $\mu$ . When there are more than 2 Slater wavefunctions a similar treatment will result and even more orbitals will be degenerate at the

chemical potential. Even if  $n$  is  $v$ -representable this type of wave function can arise.

A third way to define the kinetic energy is as an ensemble average. Instead of a wave function, one uses a mixed density matrix  $D = \sum_n c_n |\Phi_n\rangle\langle\Phi_n|$ . The constants  $c_n$  are positive and sum to 1. The density is the convex sum  $\sum_n c_n n_n(\mathbf{r})$  of densities from each participating determinant. A similar expression will arise for the kinetic energy. This approach is designed to solve the problems of non-interacting  $v$ -representability emanating from convex sums of degenerate wave functions (see section XXX).

#### iv. Janak's Theorem

A very general theorem was noted by Janak [13], based on earlier work of Slater and Wood[14] concerning the meaning of orbital energies. Let us return to the functional of Eq. (7.3.11) and assume now that the occupation numbers  $n_q$  are given and they are all non-negative and not larger than 1 and that they sum up to the number of electrons. Thus, for a given set of occupation numbers we can search for the orbitals that minimize the following functional:

$$L[\{\phi_q\}] = E_v \left[ \sum_q |\phi_q(\mathbf{x})|^2 \right] - \sum_q \epsilon_q (\langle\phi_q|\phi_q\rangle - n_q) \quad (7.3.21)$$

The equations that the orbitals must obey are still derived from Eq. (7.3.12) leading to the same equations as in (7.3.13), the KS equation. Now, let us ask: what happens to the energy when we change the occupation number of one of the orbitals  $n_q$  by an infinitesimal amount  $\delta n_q$ ? When we do this the "total number of electrons"  $N_e$  changes by this amount as well. This is not a physical change (since electrons cannot change by non-integer amounts) but still

mathematically speaking the change can be studied. Since  $n_q$  are the constraints and  $\epsilon_q$  the Lagrange multipliers in a minimization problem we can use the general result of Eq. (1.11.23) that the rate at which the minimized function changes when the constraints change is equal to the Lagrange multiplier:

$$\frac{\partial E_v}{\partial n_q} = \epsilon_q \quad (7.3.22)$$

This relation, giving some meaning to the orbital energies is called Janak's theorem. This theorem is quite general but relies on some analytical assumptions of the energy functional. For example, when the occupation number is 1 the change can only be by a negative amount and when it is zero – only positive. For approximate functionals, that are analytical with respect to any  $n(\mathbf{r})$  and  $\phi_q(\mathbf{r})$  this relation holds. Such is the case for the often used local, semilocal and even most hybrid functionals, including Hartree-Fock theorem.

In the context of Hartree-Fock theory this result is a restatement of Koopmans' theorem, by which  $-\epsilon_q$  is the unrelaxed ionization energy from orbital  $q$ . The orbital relaxation is a second order effect and thus negligible when occupation numbers change infinitesimally.

### ***E. "Virial Theorem" related identities in DFT***

The following development is inspired by the virial theorem treatment. It continues to consider dilation relations. Taking the derivative of  $n_\gamma(\mathbf{r})$  with respect to  $\gamma$ , we have  $\frac{d}{d\gamma} n_\gamma(\mathbf{r}) = \gamma^3 \mathbf{r} \cdot \nabla n(\gamma \mathbf{r}) + 3\gamma^2 n(\gamma \mathbf{r})$  and since,  $\nabla \cdot (\mathbf{r} f(\mathbf{r})) = 3f(\mathbf{r}) + \mathbf{r} \cdot \nabla f(\mathbf{r})$ , we find

$$\frac{d}{d\gamma}n_\gamma(\mathbf{r}) = \gamma^2[\nabla_{\mathbf{s}} \cdot (s\mathbf{n}(s))]_{s=\gamma\mathbf{r}} \quad (7.4.1)$$

For a general density functional,  $A[n]$ , we have, using the chain rule for derivatives,  $a[n](\mathbf{r}) = \delta A[n]/\delta n(\mathbf{r})$  and Eq. (7.4.1) then:

$$\begin{aligned} \frac{d}{d\gamma}A[n_\gamma] &= \int a[n_\gamma](\mathbf{r})\gamma^2[\nabla_{\mathbf{s}} \cdot (s\mathbf{n}(s))]_{s=\gamma\mathbf{r}}d^3r = \int \gamma^{-1}a[n_\gamma](\gamma^{-1}\mathbf{s})\nabla_{\mathbf{s}} \cdot (s\mathbf{n}(s))d^3s \\ &= -\gamma^{-1} \int (\mathbf{r} \cdot \nabla_{\mathbf{r}}a[n_\gamma](\mathbf{r}))_{r=\gamma^{-1}s} n(s)d^3s \end{aligned}$$

From which a completely general virial-dilation relation holds for any functional:

$$\frac{dA[n_\gamma]}{d\gamma} + \gamma^2 \int (\mathbf{r} \cdot \nabla_{\mathbf{r}}a[n_\gamma](\mathbf{r}))n(\gamma\mathbf{r})d^3r = 0 \quad (7.4.2)$$

We will be especially interested in the case of  $\gamma = 1$ . Thus the basic relation:

$$\left(\frac{dA[n_\gamma]}{d\gamma}\right)_{\gamma=1} + \int (\mathbf{r} \cdot \nabla_{\mathbf{r}}a[n](\mathbf{r}))n(\mathbf{r})d^3r = 0 \quad (7.4.3)$$

We further find that

$$A[n] + \int (\mathbf{r} \cdot \nabla_{\mathbf{r}}a[n](\mathbf{r}))n(\mathbf{r})d^3r = \left(\frac{d(\gamma A[n] - A[n_\gamma])}{d\gamma}\right)_{\gamma=1} \quad (7.4.4)$$

Let us apply this for the Hartree energy. From Eq. (7.1.8) we find the following relation, valid for any  $\gamma$ :  $\gamma E_H[n] - E_H[n_\gamma] = 0$ , so:

$$E_H[n] + \int \mathbf{r} \cdot \nabla v_H[n](\mathbf{r}) n(\mathbf{r})d^3r = 0 \quad (7.4.5)$$

A similar relation, namely  $\gamma K_X[n] - K_X[n_\gamma] = 0$  hold also for the exchange energy  $K_X[n]$  (see Eq. (7.1.10)), and so:

$$K_X[n] + \int \mathbf{r} \cdot \nabla v_X[n](\mathbf{r}) n(\mathbf{r})d^3r = 0 \quad (7.4.6)$$

Where  $v_X[n](\mathbf{r}) = \delta K_X[n]/\delta n(\mathbf{r})$  is the exchange potential.

Now, what about the correlation energy? Let us consider the KS DFT functional:

$$E_v[n] = T_S[n] + \int v(\mathbf{r}) n(\mathbf{r}) d^3r + U_S[n] + E_C[n] \quad (7.4.7)$$

Suppose  $n(\mathbf{r})$  is the density minimizing minimizer of  $E_v[n]$  and now plug into the latter the scaled density. This will give a  $\gamma$  dependent energy:

$$\tilde{E}(\gamma) = E_v[n_\gamma] \quad (7.4.8)$$

And clearly, the minimum is at  $\gamma = 1$  so:

$$\tilde{E}'(1) = 0 \quad (7.4.9)$$

Now, let us evaluate  $\tilde{E}'(\gamma)$  using Eq. (7.1.5):

$$\tilde{E}'(\gamma) = 2\gamma T_S[n] + \frac{d}{d\gamma} \left( \int v(\mathbf{r}) n_\gamma(\mathbf{r}) d^3r \right) + U_S[n] + \frac{d}{d\gamma} E_C[n_\gamma], \quad (7.4.10)$$

and plug in  $\gamma = 1$ . We obtain:

$$\frac{d}{d\gamma} \left\{ E_C[n_\gamma] + \int v(\mathbf{r}) n_\gamma(\mathbf{r}) d^3r \right\}_{\gamma=1} = -2T_S[n] - U_S[n]. \quad (7.4.11)$$

or, using Eq. (7.4.3):

$$\int (\mathbf{r} \cdot \nabla_{\mathbf{r}} (v_C[n](\mathbf{r}) + v(\mathbf{r}))) n(\mathbf{r}) d^3r = -2T_S[n] - U_S[n] \quad (7.4.12)$$

The second term in the parenthesis can be related to the interacting system.

Indeed we have, using Eqs. (6.6.2):

$$\begin{aligned} E[\Psi_\gamma] &= T[\Psi_\gamma] + U[\Psi_\gamma] + \int v(\mathbf{r}) n_\gamma(\mathbf{r}) d^3r \\ &= \gamma^2 T[\Psi] + \gamma U[\Psi] + \int v(\mathbf{r}) n_\gamma(\mathbf{r}) d^3r \end{aligned} \quad (7.4.13)$$

Taking the derivative with respect to  $\gamma$ , remembering that  $E'[\Psi_\gamma]_{\gamma=1} = 0$  we find, for  $\gamma = 1$ :

$$0 = 2T[\Psi] + U[\Psi] + \frac{d}{d\gamma} \left( \int v(\mathbf{r}) n_\gamma(\mathbf{r}) d^3r \right)_{\gamma=1} \quad (7.4.14)$$

(As a sidenote, you can see that substituting this relation in Eq. (7.4.1) gives after trivial manipulation the virial theorem  $2T[\Psi] + U[\Psi] = \int (\mathbf{r} \cdot \nabla v(\mathbf{r})) n(\mathbf{r}) d^3r$ ). Continuing the above development, using Eqs. (7.4.11) and (7.4.14) we finally find:

$$\frac{d}{d\gamma} E_C[n_\gamma] |_{\gamma=1} = 2T_C + U_C, \quad (7.4.15)$$

which can be written equivalently as:

$$E_C[n] - \frac{d}{d\gamma} E_C[n_\gamma] |_{\gamma=1} = -T_C[n], \quad (7.4.16)$$

or, using Eq. (7.4.3):

$$E_C[n] + \int (\mathbf{r} \cdot \nabla_r v_C[n](\mathbf{r})) n(\mathbf{r}) d^3r = -T_C[n] \leq 0. \quad (7.4.17)$$

The Hartree and exchange energies give zero while the correlation energy gives a negative quantity equal exactly to  $-T_C[n]$ .

This result shows that the correlation energy functional and the correlation potential are enough for determining the correlation kinetic energy (and from it, by  $T = T_S + T_C$ ) the kinetic energy itself.

This latter result is related to the virial theorem of Slater, which shows that one can derive the kinetic energy of the electrons from the Born-Oppneheimer potential surface itself (Eq. (1.9.30)). Since DFT gives, in principle the Born-Oppenheimer potential surface, one can access the kinetic energy (and the potential energy) from the DFT calculation.

### ***F. Galilean invariance***

A basic property of the electron-electron interaction is that if the coordinates of all electrons are shifted by a constant  $\mathbf{L}$ :

$$\mathbf{r}'_n = \mathbf{r}_n - \mathbf{L} \quad (7.5.1)$$

the e-e interaction energy does not change

$$U' = \frac{1}{2} \sum_{n \neq m} \frac{1}{|\mathbf{r}'_n - \mathbf{r}'_m|} = \frac{1}{2} \sum_{n \neq m} \frac{1}{|\mathbf{r}_n - \mathbf{r}_m|} = U \quad (7.5.2)$$

This property shows that the e-e interaction is translationally invariant. This property is also called Galilean invariance. The same property holds when the coordinates of each electron are rotated around some axis. This rotation can be described by a 3×3 orthogonal matrix  $O$ , where  $O^T O = O O^T = I$ :

$$\mathbf{r}''_n = O \mathbf{r}_n \quad (7.5.3)$$

The lengths of vectors are preserved under a rotation:

$$|\mathbf{r}''_n|^2 = \mathbf{r}''_n{}^T \mathbf{r}''_n = (O \mathbf{r}_n)^T O \mathbf{r}_n = \mathbf{r}_n^T O^T O \mathbf{r}_n = \mathbf{r}_n^T \mathbf{r}_n = |\mathbf{r}_n|^2 \quad (7.5.4)$$

Thus the e-e interaction energy does not change

$$\begin{aligned} U'' &= \frac{1}{2} \sum_{n \neq m} \frac{1}{|\mathbf{r}''_n - \mathbf{r}''_m|} = \frac{1}{2} \sum_{n \neq m} \frac{1}{|O \mathbf{r}_n - O \mathbf{r}_m|} = \frac{1}{2} \sum_{n \neq m} \frac{1}{|O(\mathbf{r}_n - \mathbf{r}_m)|} \\ &= \frac{1}{2} \sum_{n \neq m} \frac{1}{|\mathbf{r}_n - \mathbf{r}_m|} = U \end{aligned} \quad (7.5.5)$$

Thus the e-e energy is rotational invariant.

These relations indeed hold for the density functional  $E_H[n]$  since it is a reflection of the e-e functional. Indeed, define the shifted density:

$$n'(\mathbf{r}') = n(\mathbf{r}) = n(\mathbf{r}' + \mathbf{L}) \quad (7.5.6)$$

Then:



$$\begin{aligned}
E_H[n'] &= \frac{1}{2} \iint \frac{n'(\mathbf{r}'_1)n'(\mathbf{r}'_2)}{|\mathbf{r}'_1 - \mathbf{r}'_2|} d^3r'_1 d^3r'_2 \\
&= \frac{1}{2} \iint \frac{n(\mathbf{r}'_1 + \mathbf{L})n(\mathbf{r}'_2 + \mathbf{L})}{|\mathbf{r}'_1 - \mathbf{r}'_2|} d^3r'_1 d^3r'_2
\end{aligned} \tag{7.5.7}$$

We now make a variable replacement:  $\mathbf{r}'_n \rightarrow \mathbf{r}_n - \mathbf{L}$  and obtain:

$$\begin{aligned}
E_H[n'] &= \frac{1}{2} \iint \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|(\mathbf{r}_1 - \mathbf{L}) - (\mathbf{r}_2 - \mathbf{L})|} d^3r_1 d^3r_2 \\
&= \frac{1}{2} \iint \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d^3r_1 d^3r_2 = E_H[n]
\end{aligned} \tag{7.5.8}$$

This condition, that  $E_H[n] = E_H[n']$  is called Galilean invariance. It is easy to show that  $E_H[n]$  functional is also rotational invariant.

The exchange energy is also Galilean invariant, since the translation of the density will cause a translation of the density matrix:

$$\rho[n'](\mathbf{r}'_1, \mathbf{r}'_2) = \rho[n](\mathbf{r}_1, \mathbf{r}_2) \tag{7.5.9}$$

It is easy based on this to show that  $E_X[n] = E_X[n']$ .

Finally, the same will hold for the kinetic energies  $T[n]$ ,  $T_S[n]$ . All this shows that we must demand this invariance of the correlation energy:

$$E_C[n''] = E_C[n'] = E_C[n] \tag{7.5.10}$$

One consequence is the property of Galilean covariance of the potentials for each of the above energy functionals. For example, for the correlation energy we have the following result. Suppose we shift the density by a small displacement:

$$n'(\mathbf{r}) = n(\mathbf{r}' + \delta\mathbf{L}) = n(\mathbf{r}') + \delta\mathbf{L} \cdot \nabla n(\mathbf{r}')$$

Thus:

$$E_C[n'] = E'_C[n + \delta\mathbf{L} \cdot \nabla n] = E_C[n] + \delta\mathbf{L} \cdot \int \frac{\delta E_C}{\delta n(\mathbf{r}')} \nabla n(\mathbf{r}') d^3 r' \quad (7.5.11)$$

Since we demand Galilean invariance  $E_C[n'] = E_C[n]$  we find:

$$\delta\mathbf{L} \cdot \int \frac{\delta E_C}{\delta n(\mathbf{r}')} \nabla n(\mathbf{r}') d^3 r' = 0 \quad (7.5.12)$$

Since  $\frac{\delta E_C}{\delta n(\mathbf{r}')} = v_C[n](\mathbf{r})$  and  $\delta\mathbf{L}$  is arbitrary:

$$\int v_C[n](\mathbf{r}) \nabla n(\mathbf{r}) d^3 r = 0 \quad (7.5.13)$$

For finite systems where the density drops to zero at infinity we can move the nabla sign to the potential:

$$\int -\nabla v_C[n](\mathbf{r}) n(\mathbf{r}) d^3 r = 0 \quad (7.5.14)$$

$-\nabla v_C[n](\mathbf{r})$  is the force derived from the correlation potential. This shows that the total correlation force is always zero. Another consequence from Galilean invariance of the correlation energy is the Galilean covariance:

$$v_C[n'](\mathbf{r}') = v_C[n](\mathbf{r}) = v_C[n](\mathbf{r}' + \mathbf{L}) \quad (7.5.15)$$

Similar conditions can be proved from rotational invariance. For example the torque:

$$\int \mathbf{r} \times \nabla v_C[n](\mathbf{r}) n(\mathbf{r}) d^3 r = 0 \quad (7.5.16)$$

## ***G.Holes and the adiabatic connection***

### **i. The exchange and correlation holes**

Let us now take a step back and return to wave function theory. We examine the electron-electron interaction energy

$$U = \langle \Psi_{gs} | \hat{U} | \Psi_{gs} \rangle \quad (7.6.1)$$

Which we write using the following operator:

$$\hat{U} = \frac{e^2}{2} \sum_{i \neq j} \frac{1}{r_{ij}} = \frac{e^2}{2} \iint \frac{\hat{n}_2(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.2)$$

In the first term, we sum over the pairs of r-vectors of each of the  $N$  coordiantes (indices  $i$  and  $j$ ). In the second term we use the definition:

$$\hat{n}_2(\mathbf{r}, \mathbf{r}') = \sum_{i \neq j} \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \delta(\mathbf{r} - \hat{\mathbf{r}}_j), \quad (7.6.3)$$

which is the pair density *operator*. Note the relation between the one and two densities:

$$\begin{aligned} \hat{n}_2(\mathbf{r}, \mathbf{r}') &= \sum_{i,j} \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \delta(\mathbf{r}' - \hat{\mathbf{r}}_j) - \sum_i \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \delta(\mathbf{r}' - \hat{\mathbf{r}}_i) \\ &= \hat{n}(\mathbf{r}) \hat{n}(\mathbf{r}') - \delta(\mathbf{r} - \mathbf{r}') \hat{n}(\mathbf{r}) = \hat{n}(\mathbf{r}) [\hat{n}(\mathbf{r}') - \delta(\mathbf{r} - \mathbf{r}')]. \end{aligned} \quad (7.6.4)$$

With  $\hat{n}(\mathbf{r}) = \sum_{i=0}^N \delta(\mathbf{r} - \mathbf{r}_i)$ . With this definition, we have:

$$\begin{aligned} \int \hat{n}(\mathbf{r}) d^3r &= N \\ \int \hat{n}_2(\mathbf{r}, \mathbf{r}') d^3r' &= (N - 1) \hat{n}(\mathbf{r}) \\ \int \hat{n}_2(\mathbf{r}, \mathbf{r}') d^3r' d^3r &= N(N - 1) \hat{n}(\mathbf{r}) \end{aligned} \quad (7.6.5)$$

The expectation value of  $\hat{n}_2(\mathbf{r}, \mathbf{r}')$  is the "two electron density function":

$$\Gamma(\mathbf{r}, \mathbf{r}') = \langle \Psi_{gs} | \hat{n}_2(\mathbf{r}, \mathbf{r}') | \Psi_{gs} \rangle \quad (7.6.6)$$

With the two-electron density function, the interaction energy is:

$$U = \frac{e^2}{2} \iint \frac{\Gamma(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.7)$$

This pair density function has the symmetry, positivity and normalization properties given by:

$$\begin{aligned}
 \Gamma(\mathbf{r}, \mathbf{r}') &= \Gamma(\mathbf{r}', \mathbf{r}) \\
 \Gamma(\mathbf{r}, \mathbf{r}') &\geq 0 \\
 \int \Gamma(\mathbf{r}, \mathbf{r}') d^3 r' &= (N - 1)n(\mathbf{r}) \\
 \int \Gamma(\mathbf{r}, \mathbf{r}') d^3 r &= (N - 1)n(\mathbf{r}') \\
 \iint \Gamma(\mathbf{r}, \mathbf{r}') d^3 r d^3 r' &= (N - 1)N
 \end{aligned}
 \tag{7.6.8}$$

The normalization allows interpretation of  $\frac{\Gamma(\mathbf{r}, \mathbf{r}')}{N(N-1)}$  as the probability density to find an electron at  $\mathbf{r}$  and *another* electron at  $\mathbf{r}'$ . One property that is intuitively expected of  $\Gamma(\mathbf{r}, \mathbf{r}')$  is as the limit  $|\mathbf{r} - \mathbf{r}'| \rightarrow \infty$  is approached electrons will gradually uncorrelate and  $\Gamma(\mathbf{r}, \mathbf{r}')$  collapse to the density product  $n(\mathbf{r}_1)n(\mathbf{r}_2)$ . Indeed, this insight bears out in most cases:

$$\Gamma(\mathbf{r}, \mathbf{r}') \rightarrow n(\mathbf{r})n(\mathbf{r}') \quad (|\mathbf{r} - \mathbf{r}'| \rightarrow \infty, \text{ in most cases}), \tag{7.6.9}$$

but not always. For example, the ground state wavefunction of the Carbon atom  $\Psi_C(\mathbf{r}_1, \mathbf{r}_2, \dots)$  (we are neglecting to write spin indices for sake of notational simplicity) in the large  $r_1$  limit: for minimal energy reasons the remaining electrons will lower their energy to a maximal extent thus, the wave function should obey:  $\Psi_C(\mathbf{r}_1, \mathbf{r}_2, \dots) \rightarrow \sqrt{\frac{n(\mathbf{r}_1)}{N}} \Psi_{C^+, r_1}(\mathbf{r}_2, \dots)$  where  $\Psi_{C^+, r_1}(\mathbf{r}_2, \dots)$  is the ground state wave function of the cation  $C^+$  and  $\sqrt{\frac{n(\mathbf{r}_1)}{N}}$  is the root of the propability to find an electron far at  $\mathbf{r}_1$ . Note however that this cation has a 3-fold degeneracy in its groundstate energy and thus for any finite  $r_1$ , no matter how large, the  $\Psi_{C^+, r_1}(\mathbf{r}_2, \dots)$  wave function is that

degenerate wave function distorts in a certain fashion in correlation with the direction  $\mathbf{r}_1/r_1$  (for more details see Phys. Rev. A **49**, 809 (1994) or J. Chem. Phys. **105**, 2798 (1996)).

We can also look at the *conditional probability density* to find an electron at  $\mathbf{r}$  *given* that there is one at  $\mathbf{r}'$  (this latter probability is  $n(\mathbf{r}')/N_e$ ), given by

$$P(\mathbf{r}|\mathbf{r}') = \frac{\Gamma(\mathbf{r}, \mathbf{r}')}{(N_e - 1)n(\mathbf{r}')} \quad (7.6.10)$$

Obviously, if one integrates on  $\mathbf{r}$  one gets unity. We can thus view:

$$n_{cond}(\mathbf{r}|\mathbf{r}') = (N_e - 1)P(\mathbf{r}|\mathbf{r}') = \frac{\Gamma(\mathbf{r}, \mathbf{r}')}{n(\mathbf{r}')} \quad (7.6.11)$$

as a “conditional” density, the density at  $\mathbf{r}$  of  $N_e - 1$  electrons: all “other” electrons except that one electron is known to be at  $\mathbf{r}'$ . Indeed, upon integration over  $\mathbf{r}$ , we get, irrespective of  $\mathbf{r}'$  :

$$\int n_{cond}(\mathbf{r}|\mathbf{r}')d^3r = (N_e - 1) \quad (7.6.12)$$

Furthermore, we have for  $|\mathbf{r} - \mathbf{r}'| \rightarrow \infty$ :

$$n_{cond}(\mathbf{r}|\mathbf{r}') \rightarrow n(\mathbf{r}) \quad (|\mathbf{r} - \mathbf{r}'| \rightarrow \infty, \text{ in most cases}), \quad (7.6.13)$$

This shows that the density far from the localized electron is unperturbed. Now, let us subtract from this conditional density the total  $N_e$ -electron density  $n(\mathbf{r})$  and obtain the Fermi-Coulomb hole function:

$$n_{FC}(\mathbf{r}|\mathbf{r}') \equiv n_{cond}(\mathbf{r}|\mathbf{r}') - n(\mathbf{r}) \quad (7.6.14)$$

Since we localized an electron at  $\mathbf{r}'$  the rest of the electrons will “rearrange” so as to be repelled from the stationary source. This will give us a “missing density” or “hole density”, i.e. the charge density at  $\mathbf{r}$  expelled by an electron

at  $\mathbf{r}'$ . We expect the total charge of the hole is  $-1$ . Indeed plugging Eq. (7.6.14) into (7.6.12) we find:

$$\int n_{FC}(\mathbf{r}|\mathbf{r}')d^3r = -1. \quad (7.6.15)$$

Furthermore far from the hole center we have from Eqs. (7.6.13) and (7.6.14) :

$$\frac{n_{FC}(\mathbf{r}|\mathbf{r}')}{n(\mathbf{r})} \rightarrow 0 \quad (|\mathbf{r} - \mathbf{r}'| \rightarrow \infty, \text{ in most cases}), \quad (7.6.16)$$

This shows that the FC hole decays faster than the density far from the system: for localized systems it is, in most cases, a highly localized overall singly charged distribution.

Now, because  $\Gamma(\mathbf{r}, \mathbf{r}') = n(\mathbf{r}')n_{cond}(\mathbf{r}|\mathbf{r}') = n(\mathbf{r}')(n(\mathbf{r}) + n_{FC}(\mathbf{r}|\mathbf{r}'))$ , the Coulomb interaction energy can be written as:

$$U = \frac{e^2}{2} \iint \frac{n(\mathbf{r}')(n(\mathbf{r}) + n_{FC}(\mathbf{r}|\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.17)$$

And in terms of the FC hole:

$$U = E_H + \frac{e^2}{2} \iint \frac{n(\mathbf{r}')n_{FC}(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.18)$$

Thus the part of the interaction energy beyond the Hartree energy is the sum of all interaction energies  $\epsilon_{FC}(\mathbf{r}') = \frac{e^2}{2} \int \frac{n_{FC}(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r$  between an electron at  $\mathbf{r}'$  and its Fermi-Coulomb hole  $n_{FC}(\mathbf{r}|\mathbf{r}')$ . We will shortly see that the correlation energy admits a similar analysis only with slightly modified quantities.

By using non-interacting electrons we can also pull out of the integral the exchange energy and write:

$$U = U_S + \frac{e^2}{2} \iint \frac{n(\mathbf{r}')n_C(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.19)$$

This leaves a Coulomb hole which is overall neutral. It too is localized. We discuss this in the section after next.

## ii. The Fermi-Coulomb hole for harmonic electrons

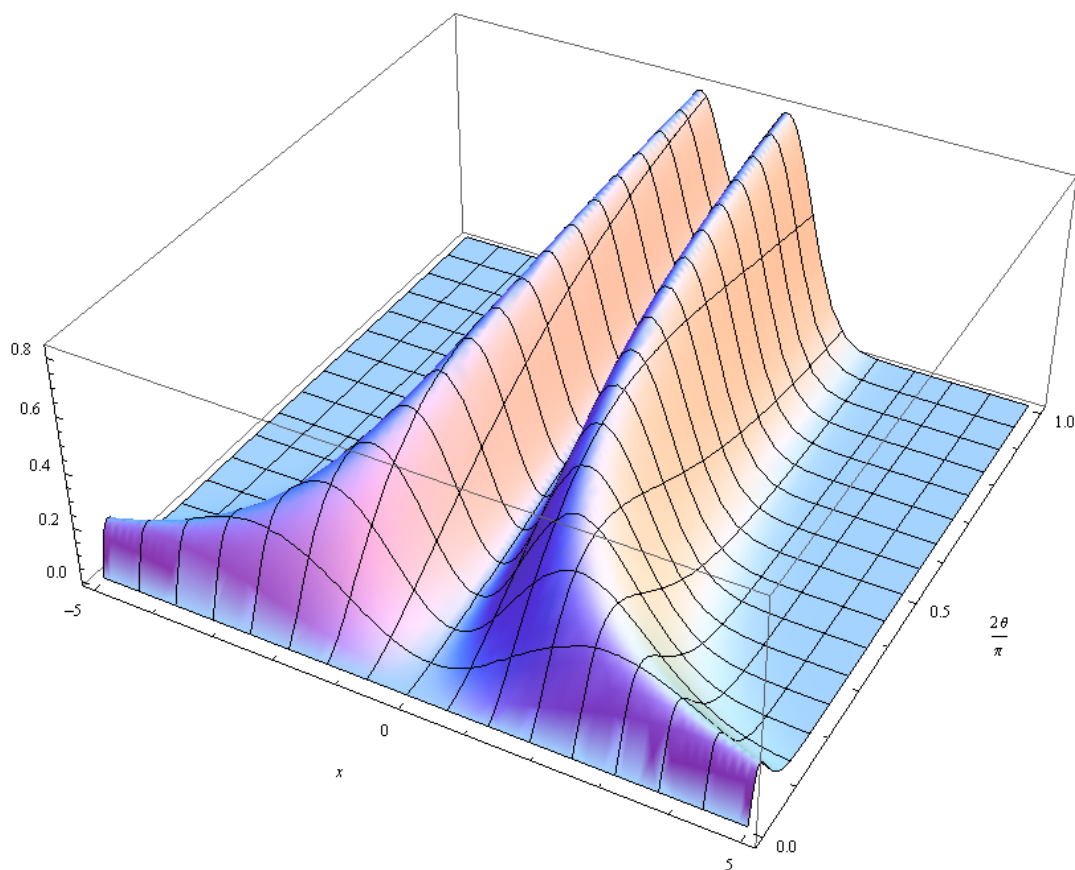
Let us calculate these quantities for our 2-harmonic electrons in their ground state triplet (so we have both exchange and correlation). The pair density and density for the wave function in (1.6.18) is:

$$\Gamma(x_1, x_2) = \frac{2}{\pi} \sqrt{\Omega_2^3 \Omega_1} e^{-\frac{\Omega_1(x_1+x_2)^2 + \Omega_2(x_1-x_2)^2}{2}} (x_1 - x_2)^2 \quad (7.6.20)$$

The density can be obtained by integrating:

$$n(x) = \sqrt{\frac{8}{\pi}} \frac{\sqrt{\Omega_2^3 \Omega_1}}{(\Omega_1 + \Omega_2)^{5/2}} e^{-2 \frac{\Omega_1 \Omega_2}{\Omega_1 + \Omega_2} x^2} (\Omega_1 + \Omega_2 + 4x^2 \Omega_2) \quad (7.6.21)$$

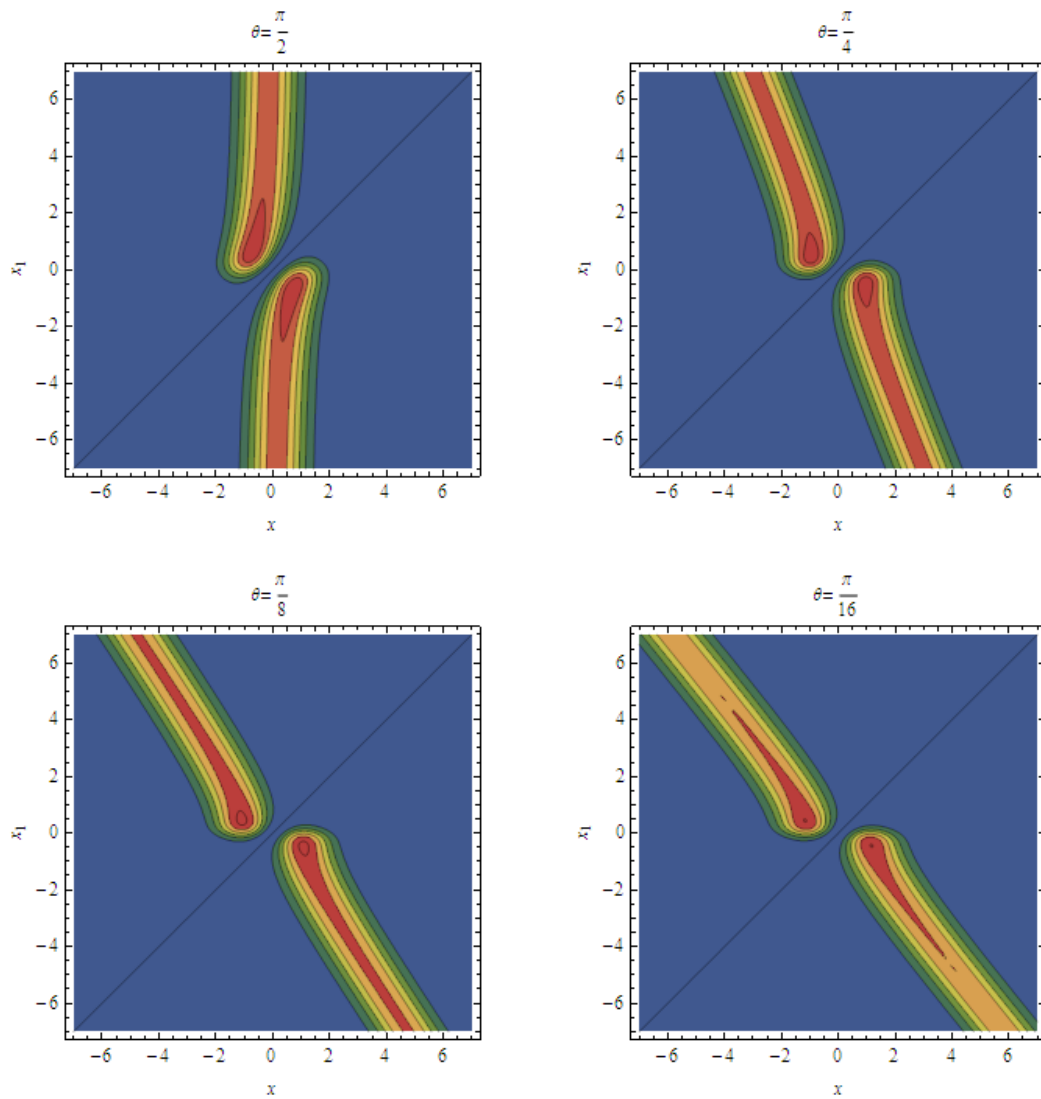
The density is plotted for several values of the correlation constant  $\theta$ ,  $\cos \theta = \left(\frac{\gamma}{\omega}\right)^2$ . This calculation is for  $\omega = 1$ :



**Figure VII-1: The 1-particle density, for a system of two harmonic fermions placed in a harmonic well in their triplet ground state for various interaction strengths. When  $\theta = \frac{\pi}{2}$  there is no interaction and the dip in  $x = 0$  is due to the “Pauli repulsion”. As interaction grows the dip becomes deeper and broader.**

We plot the conditional probability  $p(x|x_1)$  for this system in Figure VII-2.





**Figure VII-2: Contour plots for the conditional probability distribution  $p(x|x_1)$  for a system of two Fermions in their triplet ground state for various interaction strengths. When  $\theta = \frac{\pi}{2}$  there is no interaction and the only correlation is due to the Pauli principle. As interaction grows the probability distribution rotates by  $45^\circ$ .**

The XC hole is plotted next in Figure VII-3.

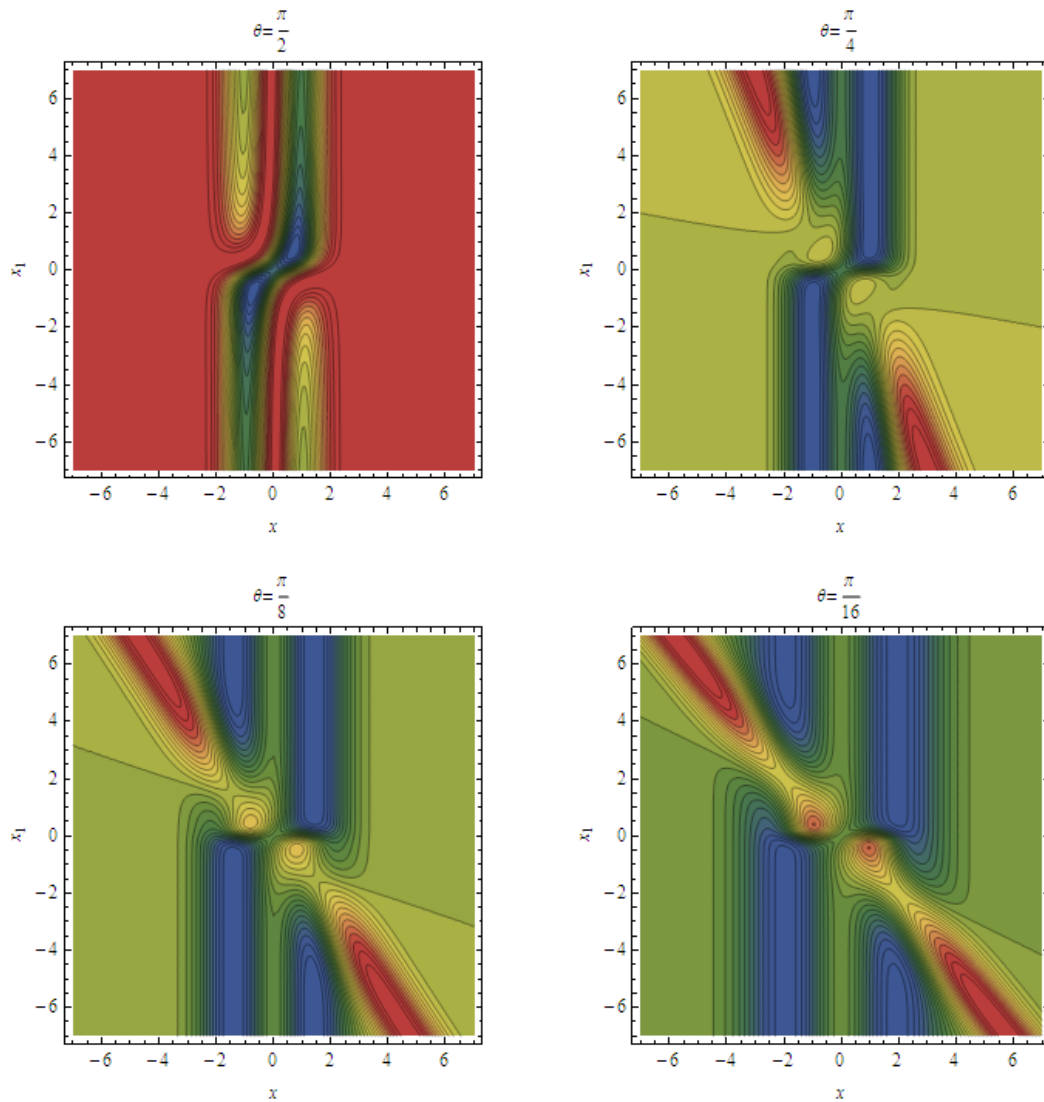


Figure VII-3: Contour plots for the FC hole  $n_{XC}(x, x_1)$  for a system of two Fermions in their triplet ground state for various interaction strengths. When  $\theta = \frac{\pi}{2}$  there is no interaction and the only correlation is due to the Pauli principle.

### iii. The Fermi hole in the non-interacting system

Let us consider now the FC hole in the non-interacting system. Since there is no correlation in absence of interactions, we attribute the hole only to the exchange (Fermi) effects. A non-interacting system having the density  $n(\mathbf{r})$  that has a closed shell Kohn-Sham determinant, composed of orthonormal orbitals  $\phi_a(\mathbf{r})$ , where 1 indicates spin up and -1 spin down.

$$\begin{aligned}
\Gamma_S(\mathbf{r}, \mathbf{r}') &= \langle \Phi[n]_{gs} | \hat{n}(\mathbf{r}, \mathbf{r}') | \Phi[n]_{gs} \rangle \\
&= \sum_{a \neq b} [\phi_a(\mathbf{r})^2 \phi_b(\mathbf{r}')^2 - \phi_a(\mathbf{r}) \phi_b(\mathbf{r}) \phi_b(\mathbf{r}') \phi_a(\mathbf{r}')] \\
&= \sum_{a, b} [\phi_a(\mathbf{r})^2 \phi_b(\mathbf{r}')^2 - \phi_a(\mathbf{r}) \phi_b(\mathbf{r}) \phi_b(\mathbf{r}') \phi_a(\mathbf{r}')] \\
&= n(\mathbf{r})n(\mathbf{r}') - P(\mathbf{r}, \mathbf{r}')^2
\end{aligned} \tag{7.6.22}$$

Where the sum is over the orbitals in  $\Phi_{gs}$  (the occupied KS orbitals) and we defined the density non-interacting matrix  $P(\mathbf{r}, \mathbf{r}') = \sum_a \phi_a(\mathbf{r}) \phi_a(\mathbf{r}')$ .

Exercise: Prove that

$$\int P(\mathbf{r}, \mathbf{r}')^2 d^3r = n(\mathbf{r}') \tag{7.6.23}$$

Exercise: As a check, integrate  $\Gamma_S(\mathbf{r}, \mathbf{r}')$  over  $\mathbf{r}$  and find:

$$\int \Gamma_S(\mathbf{r}, \mathbf{r}') d^3r = n(\mathbf{r}') [N_e - 1] \tag{7.6.24}$$

The Fermi-conditional density is:

$$n(\mathbf{r}|\mathbf{r}') = n(\mathbf{r}) - \frac{P(\mathbf{r}, \mathbf{r}')^2}{n(\mathbf{r}')} \tag{7.6.25}$$

And the Fermi-hole is:

$$n_F(\mathbf{r}|\mathbf{r}') = -\frac{P(\mathbf{r}, \mathbf{r}')^2}{n(\mathbf{r}')} \tag{7.6.26}$$

It can be shown[15] that in most cases the density matrix  $P(\mathbf{r}, \mathbf{r}')$  decays exponentially as  $|\mathbf{r} - \mathbf{r}'| \rightarrow \infty$ , although this could be much slower than  $n(\mathbf{r})$ .

Thus we may say:

$$n_F(\mathbf{r}|\mathbf{r}') \rightarrow 0 \quad (|\mathbf{r} - \mathbf{r}'| \rightarrow \infty, \text{ in most cases}), \tag{7.6.27}$$

This is weaker than Eq. (7.6.16) for the total FC hole. This shows that  $n_C = n_{FC} - n_F$  decays to zero in a similar but opposite way than  $n_F(\mathbf{r}|\mathbf{r}')$ :

$$n_F(\mathbf{r}|\mathbf{r}') \rightarrow -n_C(\mathbf{r}|\mathbf{r}') \quad (|\mathbf{r} - \mathbf{r}'| \rightarrow \infty, \text{ in most cases}), \quad (7.6.28)$$

Based on Eq. (7.6.23) the Fermi hole carries all the charge of the FC hole:

$$\int n_F(\mathbf{r}|\mathbf{r}') d^3r = -1 \quad (7.6.29)$$

This allows one to say that it is the Fermi or “exchange”-hole in the non-interacting system that “carries the charge” of the exchange correlation-hole in the interacting system. Once the interacting system has been mapped onto the non-interacting system the Fermi-hole is easily calculated. This can be used to define the Coulomb hole by:

$$n_C(\mathbf{r}|\mathbf{r}') = n_{FC}(\mathbf{r}|\mathbf{r}') - n_F(\mathbf{r}|\mathbf{r}') \quad (7.6.30)$$

It has no total charge:

$$\int n_C(\mathbf{r}|\mathbf{r}') d^3r = 0 \quad (7.6.31)$$

The interaction energy can be written now as:

$$U = U_S + \frac{1}{2} \iint \frac{n(\mathbf{r}') n_C(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.32)$$

Exercise: Compute the Fermi-hole function of the homogeneous electron gas

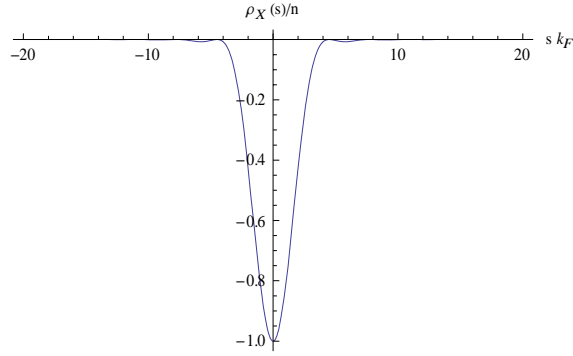
Solution: We already determined the density matrix (see Eq. (5.2.24))

$$P(s) = 3n \frac{j_1(sk_F)}{sk_F} \quad (7.6.33)$$

where  $s = |\mathbf{r} - \mathbf{r}'|$  and  $j_1(x) = \frac{\sin x - x \cos x}{x^2}$ . The x-hole is of course independent of  $\mathbf{r}$ :

$$n_F^{HEG}(\mathbf{r}; \mathbf{r} + \mathbf{s}) = -\frac{P(s)^2}{n} = -9n \frac{j_1(sk_F)^2}{(sk_F)^2} \quad (7.6.34)$$

Plotting shows the form of the HEG x-hole function:



Here  $k_F^3 = 3\pi^2 n$  (unpolarized gas). Since  $4\pi \int_0^\infty \frac{j_1(x)^2}{x} dx = \pi$ , we find HEG exchange energy is, :

$$\begin{aligned} \frac{U_F}{N} &= \frac{1}{N} \frac{1}{2} \iint \frac{n(\mathbf{r}) n_F^{HEG}(\mathbf{r}; \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' = \frac{1}{2} 4\pi \int_0^\infty \frac{n_F^{HEG}(s)}{s} s^2 ds \\ &= -\frac{9n}{2k_F^2} 4\pi \int_0^\infty \frac{j_1(x)^2}{x} dx = -\frac{3}{2\pi} k_F = -3 \left( \frac{3}{8\pi} n \right)^{1/3} \end{aligned} \quad (7.6.35)$$

This is indeed the LDA exchange energy per particle.

#### iv. The Adiabatic Connection

Having written down the relation between the interaction energy  $U$  and the XC hole, we still have no such relation for the correlation energy  $E_C$ . We now derive such a relation. Remember the correlation energy is defined as:

$$E_C[n] = T[n] - T_S[n] + (U[n] - U_S[n]) = T_C[n] + U_C[n] \quad (7.6.36)$$

Given a ground-state electron density  $n(\mathbf{r})$ , consider a family of  $N_e$ -"electron" systems, with parameter  $0 \leq \lambda \leq 1$ , where:

$$\hat{H}_\lambda = \hat{T} + \int v_\lambda(\mathbf{r}) \hat{n}(\mathbf{r}) d^3r + \lambda \hat{U} \quad (7.6.37)$$

The ground-state is denoted  $\Psi_\lambda$ . The potential  $v_\lambda(\mathbf{r})$  is chosen in such a manner that the density at the ground-state wave function is  $n(\mathbf{r})$ , i.e.:

$$\langle \Psi_\lambda | \hat{n}(\mathbf{r}) | \Psi_\lambda \rangle = n(\mathbf{r}) \quad (7.6.38)$$

This is a generalization of the idea by Kohn and Sham, that the interacting electron system is mapped onto a non-interacting electron system with the same density. Except that now we map our system to a system of electrons with interaction  $\lambda\hat{U}$ . When  $\lambda = 0$  we have the non-interacting system and  $v_{\lambda=0}(\mathbf{r})$  is the Kohn-Sham potential  $v_s(\mathbf{r})$  and we have:

$$E_0[n] = T_S[n] + \int v_0(\mathbf{r})n(\mathbf{r})d^3r \quad (7.6.39)$$

where  $T_S[n]$  is the kinetic energy of non-interacting electrons. When  $\lambda = 1$  we have the fully interacting system and  $v_{\lambda=1}(\mathbf{r}) = v(\mathbf{r})$  is the actual external potential on the electron system and the energy is:

$$E_1[n] = E_v[n] = T[n] + \int v_1(\mathbf{r})n(\mathbf{r})d^3r + U[n] \quad (7.6.40)$$

We can also define the obvious quantities:

$$E_\lambda[n] = T_\lambda[n] + \int v_\lambda(\mathbf{r})n(\mathbf{r})d^3r + U_\lambda[n] \quad (7.6.41)$$

From Eq. (7.6.36):

$$\begin{aligned} E_1 - E_0 &= \int [v_1(\mathbf{r}) - v_0(\mathbf{r})]n(\mathbf{r})d^3r + U_S[n] + E_C[n] \\ E_\lambda - E_0 &= \int [v_\lambda(\mathbf{r}) - v_0(\mathbf{r})]n(\mathbf{r})d^3r + \lambda U_S[n] + E_{C,\lambda}[n] \end{aligned} \quad (7.6.42)$$

With:

$$E_{C,\lambda}[n] = \langle \Psi_\lambda | \hat{T} + \lambda\hat{U} | \Psi_\lambda \rangle - \langle \Psi_0 | \hat{T} + \lambda\hat{U} | \Psi_0 \rangle$$

Now, the ground-state energy of the intermediately interacting electrons obeys, by Hellmann-Feynman's theorem:

$$\frac{d}{d\lambda} E_\lambda = \left\langle \Psi_\lambda \left| \frac{d}{d\lambda} \hat{H}_\lambda \right| \Psi_\lambda \right\rangle = \int \left[ \frac{d}{d\lambda} v_\lambda(\mathbf{r}) \right] n(\mathbf{r})d^3r + \langle \Psi_\lambda | \hat{U} | \Psi_\lambda \rangle \quad (7.6.43)$$

From the second equation then:

$$\frac{d}{d\lambda} E_{c,\lambda}[n] = \langle \Psi_\lambda | \hat{U} | \Psi_\lambda \rangle - U_S[n] = U_{c,\lambda}[n] \quad (7.6.44)$$

This expression is the differential form of the adiabatic connection. If we integrate it with respect to  $\lambda$  from 0 to 1, we find:

$$E_c[n] = \int_0^1 \langle \Psi_{\lambda'} | \hat{U} | \Psi_{\lambda'} \rangle d\lambda' - U_S[n]. \quad (7.6.45)$$

This formula is called the "adiabatic connection" formula for the XC energy [16]. We may write:  $E_{c,\lambda} = T_{c,\lambda} + \lambda U_{c,\lambda}$ . Then  $\frac{d}{d\lambda} E_{c,\lambda}[n] = \frac{d}{d\lambda} T_{c,\lambda} + U_{c,\lambda} + \lambda \frac{d}{d\lambda} U_{c,\lambda}$  and so

$$\frac{dT_{c,\lambda}[n]}{d\lambda} + \lambda \frac{dU_{c,\lambda}[n]}{d\lambda} = 0. \quad (7.6.46)$$

We can rewrite  $U_{c,\lambda}$  in terms of the correlation hole. Indeed, if  $n_c^\lambda(\mathbf{r}|\mathbf{r}')$ , is the correlation hole for the  $\lambda$  system then using (7.6.32):

$$U_{c,\lambda}[n] = \frac{1}{2} \iint \frac{n(\mathbf{r}') n_c^\lambda(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (7.6.47)$$

From which:

$$E_c[n] = \frac{1}{2} \iint \frac{n(\mathbf{r}') \bar{n}_c(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' d^3r \quad (7.6.48)$$

And we see that the correlation energy can be obtained from the the  $\lambda$ -averaged Coulomb hole, called the correlation hole (since it is associated with the correlation energy):

$$\bar{n}_c(\mathbf{r}|\mathbf{r}') = \int_0^1 n_c^\lambda(\mathbf{r}|\mathbf{r}') d\lambda \quad (7.6.49)$$

Note that because  $\int n_c^\lambda(\mathbf{r}, \mathbf{r}') d^3r = 0$ , we have also:

$$\int \bar{n}_c(\mathbf{r}|\mathbf{r}')d^3r = 0 \quad (7.6.50)$$

It is interesting that the correlation energy, like to Coulomb energy, can be represented as a Coulomb interaction of the density and a hole as in Eq. (7.6.48). Note however that the relevant hole as a coupling-constant  $(\lambda)$  averaged correlation hole and not the Coulomb hole itself.

Let us discuss one of the important consequences of Eq. (7.6.50) i.e. that the total charge of the correlation hole is zero for localized charge systems. If we rewrite the correlation energy as:

$$E_C[n] = \frac{1}{2} \iint \frac{n(\mathbf{r}') \bar{n}_c(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' d^3r \quad (7.6.51)$$

We see that the correlation energy can be written as

$$E_C[n] = \int \epsilon_c(\mathbf{r}')n(\mathbf{r}')d^3r' \quad (7.6.52)$$

where:

$$\epsilon_c(\mathbf{r}') = \frac{e^2}{2} \int \frac{\bar{n}_c(\mathbf{r}|\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r \quad (7.6.53)$$

(Note that this is just a suggestion since adding to  $\epsilon_c(\mathbf{r}')$  any function  $\Delta\epsilon_c(\mathbf{r}')$  for which  $\int \Delta\epsilon_c(\mathbf{r}')n(\mathbf{r}')d^3r' = 0$  will give the same correlation energy). Because for a fix  $\mathbf{r}'$   $\bar{n}_c(\mathbf{r}|\mathbf{r}')$  is an overall neutral charge density in  $\mathbf{r}$  space its "Coulombic potential"  $\frac{\epsilon_c(\mathbf{r}')}{e}$  is expected to decay relatively fast for  $r'$ (faster than  $1/r'$ ).

### ***H. Derivative Discontinuity in the exchange correlation potential functional***



## VIII. Approximate correlation energy functionals

While the correlation energy in atoms and molecules is only a small fraction of the total electronic energy, it is found that it is in fact a very large percentage when one computes energy differences, such as energy of atomization (i.e. the difference between the energy of the atomic constituents to the energy of the molecule) or relative energies of different conformations, as those determining the shape of the Born-Oppenheimer potential surface. In essence, the exchange correlation energy is the chemical bonding energy. It is therefore crucial to model this energy accurately. We describe below some of the basic approximations for density functional theory.

### *A. The local density approximation (LDA)*

The mapping of the interacting electron system onto the non-interacting system, encapsulated in Eq. XX, is of formal interest only, unless we devise a way to approximate the correlation potential. One way is to consider the correlation energy per electron  $\epsilon_c(n)$  in the homogeneous electron gas of density  $n$ . This energy can be computed with relatively high precision using Monte Carlo methods. Under this approximation we can write the correlation energy as  $\int \epsilon_c(n(\mathbf{r}))n(\mathbf{r})d^3r$ . However, this does not yield in practice good enough results and is thus seldom used. A more successful way was devised by Kohn and Sham. They considered both the exchange and correlation per an electron in the homogeneous gas,  $\epsilon_{xc}(n)$ . In this case the correlation energy comes out:

$$E_{XC}^{LDA}[n] = \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d^3r - E_X[n] \quad (8.1.1)$$

This approximation is called the local density approximation (LDA)[17]. The functional  $E_{XC}^{LDA}[n]$ . It leads to the following LDA approximation for the energy functional:

$$E_v^{LDA}[n] = T_S[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r + E_H[n] + \int \epsilon_{XC}(n(\mathbf{r}))n(\mathbf{r})d^3r \quad (8.1.2)$$

The minimization of this functional, by the Kohn-Sham approach leads to the LDA approximation of DFT. This approach is highly successful and is considered in DFT as the basis for most of the developments of other functionals.

Note however that in LDA the correlation energy is extremely awkward looking because of the term  $-E_X[n]$ . The presence of this term has detrimental effects which harm some of the predictions of DFT.

### **i. The exchange energy per electron in the HEG**

In section XXX we discussed in some detail the Hartree-Fock theory of the homogeneous electron gas. We defined a Jellium as a smeared positive background of Volume  $V$  at density  $n = \frac{N_e}{V}$  together with  $N_e$  electrons. We showed that the Jellium self energy, the Jellium-electron attraction energy and the electron Hartree energy all cancel exactly in the HEG. Thus, the energy per particle is given by:

$$\epsilon(n) = t_s(n) + \epsilon_{XC}(n) = t_s(n) + \epsilon_X(n) + \epsilon_C(n) \quad (8.1.3)$$

We already calculate exchange energy, using a Hartree-Fock treatment of the HEG and we saw that  $\epsilon_X = -C_X n^{1/3}$  (see Eq. (5.2.32)) with  $C_X = \frac{3}{4} \left(\frac{3}{\pi}\right)^{1/3} E_h a_0$ . In terms of the Wigner-Sietz radius, which is a dimensionless quantity given by:

$$r_s = \left( \frac{3}{4\pi n} \right)^{1/3} a_0 \quad (8.1.4)$$

is the radius of a Jellium sphere containing the charge of an electron. Thus:

$$\epsilon_x = - \left( \frac{243}{256\pi^2} \right)^{1/3} \frac{E_h}{r_s} = - \frac{0.45817}{r_s} E_h \quad (8.1.5)$$

If we have a way of computing  $\epsilon(n)$  and  $t(n)$ , we can then find  $\epsilon_{xc}(n)$  from Eq. (8.1.3). We can then also compute  $\epsilon_c(n)$  from Eqs. (8.1.5).

## ii. Correlation energy of the HEG: the high density limit

The calculation of  $\epsilon_c$  can presently be done analytically in two limits. One is the high density limit  $r_s \rightarrow 0$  where the kinetic energy dominates and the Coulomb interaction can be treated as a perturbation. In this limit the kinetic energy is that of non-interacting particles. Thus in the perturbative approach one can take  $\hat{H}_0 = \sum_{n=1}^{N_e} -\frac{\hbar^2}{2m_e} \nabla_n^2$ . The unperturbed wave function is the Slater wave functional wave function  $\Psi_0$  composed of the plane wave orbitals  $\frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{V}}$  with  $\mathbf{k}$  taking the  $N_e$  lowest momentum vectors  $k < k_F$ . The energy of the unperturbed state is  $E_0 = \sum_{k < k_F} \frac{\hbar^2 k^2}{2m_e}$ . The first order contribution of the the e-e Coulomb repulsions is  $\langle \Psi_0 | \hat{U} | \Psi_0 \rangle$ . This quantity is equal to the direct and exchange contribution (see Eq. (4.1.12)). The direct part is nullified by the other electrostatic interactions, so the 1<sup>st</sup> order contribution is essentially the exchange energy of the HEG which we already included  $\epsilon$  (see Eq. (8.1.3)). Thus to continue and determine the interaction energy beyond exchange, i.e. the correlation energy, we must move to at least second order perturbation theory. When one does this, one finds that the usual second order perturbation theory yields infinite terms. These are associated with low wave length excitations where a pair of electrons having the momentum state  $|\mathbf{k}_1\rangle$

and  $|\mathbf{k}_2\rangle$ , are excited by the Coulomb interaction  $V_q = 4\pi/q^2$  to states  $|\mathbf{k}_1 + \mathbf{q}\rangle$  and  $|\mathbf{k}_2 - \mathbf{q}\rangle$  (such that  $\frac{1}{2}(\mathbf{k}_1 + \mathbf{q})^2 > \epsilon_F$  and  $\frac{1}{2}(\mathbf{k}_2 - \mathbf{q})^2 > \epsilon_F$ ). One can show that for small  $q$  this process gives a term proportional to  $\ln q$  in the expression for the 2<sup>nd</sup> order perturbation energy. This term is singular at low  $q$ . A method of performing perturbation theory which is non-singular and goes beyond second order, was devised in 1957[18]. This theory is essentially exact at the high density limit and leads to the following relation:

$$\epsilon_c = A \ln r_s + C + O(r_s) \quad (8.1.6)$$

Where  $r_s = \frac{1}{a_0} \left( \frac{3}{4\pi n} \right)^{-1/3}$  is the Wigner-Seitz radius, namely  $a_0 r_s$  is the radius of a sphere in the Jellium which scoops an amount of charge equal to  $1e$ , where  $e$  is the elementary quantum of charge (the electron charge).  $A = 0.0311E_h$  and  $C = -0.048 \pm 0.001E_h$ . Later work refined these constants:  $A = 0.031091E_h$  and  $C = -0.046644E_h$ .

#### Exercise VIII-1

Using Eq. (1.9.43) and the Hellman-Feynman theorem prove that

$$t_c \equiv t - t_s = 3v_{XC}^{HEG} - 4\epsilon_{XC}^{HEG} = 3v_C^{HEG} - 4\epsilon_C^{HEG} \quad (8.1.7)$$

Hints:

(a) Show that  $e^2 \frac{\partial \epsilon^{HEG}}{\partial e^2} = u_{ee} - u_H$  where  $u_{ee}$  and  $u_H$  are, respectively, the total electronic repulsion energy per electron and the Hartree energy per electron.

(b) From the the fact that  $\epsilon_x \propto n^{1/3}$  show that the last equality is correct.

### iii. Correlation energy of the HEG: the low density limit and the Wigner crystal

The second limit is that of low density, where the electrons form a crystal. We give the development of this limit, originally proposed by Wigner[19] when he devised a theory for the electron density in metallic sodium.

Wigner assumed that at low energy the homogeneous electron gas forms a crystal. Now that may sound strange: how can the density be uniform and at the same time the electrons form a crystal? Thanks to Quantum Mechanics this is actually not a contradiction, as the following example shows.

Exercise: Show that for 2 particle in a 3D box of volume  $V$  with periodic boundary conditions, if the Hamiltonian is:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + u(\mathbf{r}_1 - \mathbf{r}_2) \quad (8.1.8)$$

Then

- 1) The eigenstates have a homogeneous 1-particle density  $n = \frac{2}{V}$
- 2) The pair correlation function has structure.  $(g(x_1, x_2) = \frac{P(x_1|x_2)}{P_{class}(x_1|x_2)} = \frac{\Gamma(x_1, x_2)}{n(x_1)n(x_2)})$

In the low density regime the electron kinetic energy (per electron) can be neglected since, as seen in XXX it is proportional to  $r_s^{-2}$  while the repulsion energy between the electrons per electron is proportional to  $r_s^{-1}$ . At low density the Pauli exclusion principle is non-operative, since electrons do not overlap. Thus the quantum nature of the electron is gone at this limit and we can think of the electron as a classical particle that localizes. This is because non-localization of particles in quantum mechanics arises only from their need to reduce kinetic energy. The electrons will arrange themselves in the lowest energy state by forming a close packed crystal. Each electron is then as far as possible from each other electron, while still filling 3D space with

average density  $n$ . Let us calculate the energy of such a crystal. Consider one of the electrons in the crystal. We imagine it together with a cell containing 1 unit of positive charge. This cell shape depends on the crystal symmetry. Following Wigner, we neglect the crystal structure and assume each electron is surrounded by a *sphere* of positive charge completely neutralizing it. Our approximation then neglects the volume of the space between the spheres. The radius of the sphere is  $r_s$  and it is filled with smeared positive charge and with one negative charged point-electron at its center. The spheres do not interact since they are neutral and have no electric moments.

The total energy per electron is the energy  $\epsilon_{sp}(r_s a_0)$  to assemble the Jellium sphere and the energy  $\epsilon_{el}(r_s a_0)$  needed to bring the electron from infinity into the center of the sphere.

Let  $\epsilon_{sp}(R)$  be the energy to assemble a sphere of charge density  $n$  and radius  $R$ . Suppose we now enlarge it by adding a shell of radius  $dR$ . The electric potential at distance  $r > R$  outside the sphere is  $Q/r$  where  $Q = \frac{4\pi}{3} enR^3$  is the charge in the sphere. The charge in the shell is  $dq = n4\pi R^2 dR$  and bringing it from infinity, where the potential is zero to its place on top of the existing sphere involves the energy  $d\epsilon_{sp} = dq \frac{Q}{R} = \frac{(4\pi)^2}{3} R^4 e^2 n^2 dR$ . Thus, by integration from 0 to  $R$ , we find:  $\epsilon_{sp}(R) = \left(\frac{4\pi n}{3}\right)^2 \frac{3}{5} R^5$ . And so at  $R = r_s a_0$ :

$$\epsilon_{sp}(r_s) = \frac{3}{5} \frac{1}{r_s} E_h \quad (8.1.9)$$

Next, we want to calculate the energy to bring an electron from infinity to the center of the sphere. This will be done in two stages, first bringing the electron from infinity to the rim of the sphere, a distance  $a_0 r_s$  from its center and then from the rim to the center. Accordingly write  $\epsilon_e = \epsilon_{rim} + \epsilon_{center}$ . The first part

is easy since we already know the potential, and it is negative since energy is released by this process, so:

$$\epsilon_{rim} = -\frac{1}{r_s} E_h \quad (8.1.10)$$

Inside the sphere, at a distance  $R$  from the center there exists an electric field due to the Jellium, which according to Gauss' is  $\mathbf{E} = \frac{Q(R)\mathbf{R}}{R^3} = \frac{4\pi}{3} n\mathbf{R}$ . This force is a Harmonic force, with force constant  $k_H = \frac{4\pi}{3} n = \frac{1}{(a_0 r_s)^3}$ . The work to move an electron in this field to the center is:  $\epsilon_{center} = \int_{r_s}^0 \mathbf{E}(R) \cdot d\mathbf{R} = -\frac{1}{2r_s}$ . The energy for the second stage is therefore  $\epsilon_{el}(r_s) = -\frac{3}{2} \frac{E_h}{r_s}$  and the total energy per electron in the crystal is:

$$\epsilon(r_s) = -\frac{9}{10} \frac{1}{r_s} E_h \quad (8.1.11)$$

This then is the exchange-correlation energy for low density. We neglected the volume between the spheres. The exchange energy we already know from (8.1.5), is  $\epsilon_X(r_s) = -\frac{0.45817}{r_s} E_h$ . Thus, Wigner's approximation for the correlation energy in the low density limit is:

$$\epsilon_C(r_s) = \epsilon(r_s) - \epsilon_X(r_s) = -\frac{0.44183}{r_s} E_h \quad (\text{low density, } r_s \rightarrow \infty) \quad (8.1.12)$$

Wigner also considered the correction due to the finite kinetic energy when  $r_s$  is finite. Since we saw that the electron inside the spherical Jellium drop is a Harmonic potential, one can reduce the correlation energy by the 3D Harmonic zero-point potential,  $\frac{3}{2} \sqrt{k_H} = \frac{3}{2r_s^{3/2}}$ . The correlation energy is then:

$$\epsilon_C(r_s) = -\frac{0.44183}{r_s} E_h + \frac{3}{2r_s^{3/2}} \quad (\text{low density, } r_s \rightarrow \infty) \quad (8.1.13)$$

#### iv. Monte-Carlo determination of the correlation energy for the HEG

Between these high and low limits there is no analytical theory, in general and a numerical computation can be made based on quantum Monte Carlo methods. The results of the calculation are then fitted to an analytical form which respects the limits

#### v. The polarized HEG; local spin-density approximation (LSDA)

Up to now we have assumed that the electron gas is unpolarized, i.e. the total  $z$  component of spin  $S_z$  per electron is zero. However,  $S_z$  is a good quantum number and it can vary continuously from  $-\frac{1}{2}$  to  $\frac{1}{2}$ . The extreme case is the fully polarized case. In general one may define the density of spin-up electron  $n_\uparrow(\mathbf{r})$  and that of spin down  $n_\downarrow(\mathbf{r})$ . Then:

$$\begin{aligned}n(\mathbf{r}) &= n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r}) \\ \zeta(\mathbf{r}) &= \frac{n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r})}{n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r})}\end{aligned}\tag{8.1.14}$$

For a fully polarized gas  $\zeta = 1$  and the difference is first of all in the Fermi energy. For the HEG, since every momentum state can populate only one electron, we find by a similar analysis as in the unpolarized case:

$$\begin{aligned}k_{F\uparrow\downarrow} &= k_F(1 \pm \zeta)^{1/3} \\ \epsilon_{F\uparrow\downarrow} &= \frac{\hbar^2 k_{F\uparrow\downarrow}^2}{2m_e} = \epsilon_F(1 \pm \zeta)^{2/3} \\ t_{\uparrow\downarrow} &= \frac{T_{\uparrow\downarrow}}{N_{\uparrow\downarrow}} = \frac{3}{5} \epsilon_{F\uparrow\downarrow}\end{aligned}\tag{8.1.15}$$

The total kinetic energy is the sum of up and down contributions:  $t = \frac{n_\uparrow t_\uparrow + n_\downarrow t_\downarrow}{n}$ , which is evaluated to be  $t = \frac{3}{5} \frac{\epsilon_{F\uparrow} n_\uparrow + \epsilon_{F\downarrow} n_\downarrow}{n}$  using the expressions for  $n_{\uparrow\downarrow}$  we obtain



$$t = \frac{3}{5} \epsilon_F \frac{(1 + \zeta)^{5/3} + (1 - \zeta)^{5/3}}{2} \quad (8.1.16)$$

As for exchange energy, since exchange interaction occurs only between like spins, we XXXXX

## vi. Successes and failures of LSDA

## vii. Plausible reasons for the success of LSDA

One of the uses of Eq. (7.6.48) expression is to explain the success of a simple theory such as LDA[20]. TO see this, let us expand the XC hole in terms of moments around  $\mathbf{r}$ :

$$n_{XC}^\lambda(\mathbf{r}, \mathbf{r} + \mathbf{R}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l n_{lm}^\lambda(\mathbf{r}; R) Y_{lm}(\theta, \phi) \quad (8.1.17)$$

where:  $\mathbf{R} = R\hat{\mathbf{u}} \equiv (R \sin \theta \cos \phi, R \sin \theta \sin \phi, R \cos \theta)$  and

$$n_{lm}^\lambda(\mathbf{r}; R) = \int_0^\pi \int_0^{2\pi} n_{XC}^\lambda(\mathbf{r}, \mathbf{r} + R\hat{\mathbf{u}}) Y_{lm}(\theta, \phi) \sin \theta d\theta d\phi \quad (8.1.18)$$

Then consider the XC energy, it can be written as:

$$\begin{aligned} E_{XC}[n] &= \frac{1}{2} \iint \frac{n(\mathbf{r}) \bar{n}_{XC}(\mathbf{r}, \mathbf{r} + \mathbf{R})}{R} d^3R d^3r \\ &= \frac{1}{2\sqrt{4\pi}} \iint \frac{n(\mathbf{r}) \bar{n}_{00}(\mathbf{r}, R)}{R} d^3r R^2 dR \end{aligned} \quad (8.1.19)$$

Were:

$$\bar{n}_{00}(\mathbf{r}, R) = \int_0^1 n_{00}^\lambda(\mathbf{r}, R) d\lambda \quad (8.1.20)$$

And:

$$\sqrt{4\pi} \int_0^\infty \bar{n}_{00}(\mathbf{r}, R) R^2 dR = -1 \quad (8.1.21)$$

Thus, only the 00 moment of the  $\lambda$  dependent XC holes enters the expression. Therefore, in a sense the angular shape of the average XC hole gets averaged over and only the radial dependence affects the XC energy. This is used to explain some of the success of LDA.

We see that the XC anisotropy of the XC hole around  $\mathbf{r}$  is averaged over. Only this average enters the XC energy formula. In LDA we use the homogeneous electron gas to compute the HEG. Of course this leads to an isotropic XC hole. Yet, since only the spherical average of the hole enters into the XC energy, this drastic approximation gives a reasonably good XC energy.

Exercise: Calculate the spherically averaged X-hole for a 1-electron system (H atom for example)

Solution: The orbital is  $\psi(\mathbf{r})$ , the density is  $n(\mathbf{r}) = \psi(\mathbf{r})^2$  and the DM is  $P(\mathbf{r}, \mathbf{r}') = \psi(\mathbf{r})\psi(\mathbf{r}')$  thus:

$$n_x(\mathbf{r}; \mathbf{r}') = -n(\mathbf{r}') \quad (8.1.22)$$

The hole is independent of the reference point  $\mathbf{r}$ .

One of the important results shown below is that only the spherically averaged hole enters the XC energy. Thus we only need the spherically averaged hole:

$$n_x^{SA}(\mathbf{r}; s) = \frac{1}{4\pi} \int n_x(\mathbf{r}; \mathbf{r} + \mathbf{s}) d\Omega_s \quad (8.1.23)$$

Which becomes:

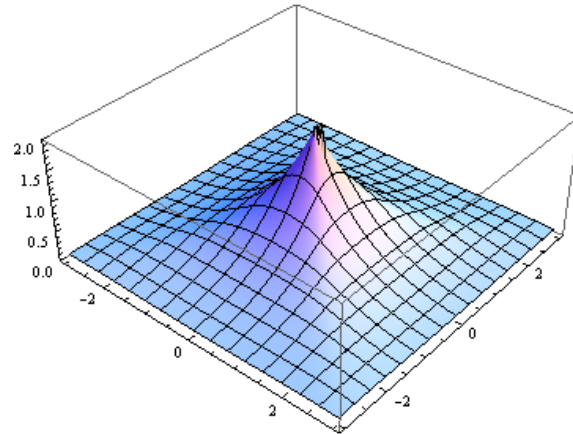
$$n_x^{SA, 1el}(\mathbf{r}; s) = \frac{1}{4\pi} \int n(\mathbf{r} + \mathbf{s}) d\Omega_s \quad (8.1.24)$$

For the H-atom  $n(\mathbf{r}) = A^2 e^{-\alpha r}$ , defining  $\beta = \alpha \sqrt{r^2 + s^2}$   $\gamma = \frac{2rs}{(r^2+s^2)}$  therefore

$\beta \sqrt{1 \pm \gamma} = \alpha |r \pm s|$  and  $\gamma \beta^2 = \alpha^2 2rs$  :

$$\begin{aligned}
 n_X^{SA,1el}(\mathbf{r}; s) &= \frac{1}{4\pi} A^2 \int e^{-\alpha|r+s|} d\Omega_s = \frac{1}{2} A^2 \int_0^\pi e^{-\alpha \sqrt{r^2+s^2+2rs\cos\theta}} d\cos\theta \\
 &= \frac{A^2}{2\gamma} \int_{-\gamma}^{\gamma} e^{-\beta \sqrt{1+y}} dy \\
 &= \frac{A^2}{\alpha^2 2rs} [(1 + \alpha|r-s|)e^{-\alpha|r-s|} \\
 &\quad - (1 + \alpha|r+s|)e^{-\alpha|r+s|}]
 \end{aligned} \tag{8.1.25}$$

The form of this spherically averaged hole function as function of s and r is shown here:



The hole has a cusp at the r-s origin.

## *B. Semilocal functionals and the generalized gradient approximation*

# **IX. Generalized Kohn-Sham approaches**

### *A. The generalized Kohn-Sham framework*

The Kohn-Sham approach is based on pure density functionals for the exchange correlation energy. This approach is based on the Hohenberg-Kohn universal density functional  $F_{HK}[n]$  is written as:

$$F_{HK}[n] = T[n] + U[n] \quad (9.1.1)$$

We noted one problem with  $F_{HK}$  is that it is not defined for all densities. This problem can be cured by the Levy-Lieb procedure, where the  $F_{HK}$  is replaced by a functional based on a constrained minimum principle:

$$F_{LL}[n] = \min_{\Psi \rightarrow n} \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle \quad (9.1.2)$$

Similarly, for the non-interacting system, one finds:

$$T_S[n] = \min_{\Phi \rightarrow n} \langle \Phi | \hat{T} | \Phi \rangle = \min_{\Phi \rightarrow n} S[\Phi] \quad (9.1.3)$$

Where here the search is over all determinants. Here we defined here a new determinantal functional  $S[\Phi]$  with the intent of generalizing the KS approach, as we do now. Using the minimizing determinant  $\Phi_S$ , the Kohn-Sham correlation energy:

$$E_C[n] = F_{LL} - T_S - U_S \quad (9.1.4)$$

Is written using  $T_S = \langle \Phi_S | \hat{T} | \Phi_S \rangle$  and  $U_S = \langle \Phi_S | \hat{U} | \Phi_S \rangle$ .

Since a Slater wave function is given in terms of orthonormal one-electron orbitals  $\phi_n(\mathbf{x})$ , it is best to view  $S$  as a functional of the orbitals and we write  $S[\{\phi\}]$ . We can now generalize and demand that  $S[\{\phi\}]$  not be just the kinetic

energy but a more general functional of the Slater wave function. We denote the functional derivative of  $S$  as follows:

$$\frac{\delta S[\Phi]}{\delta \phi_n(\mathbf{r})} = \hat{O}_S \phi_n(\mathbf{r}) \quad (9.1.5)$$

Where  $O_S$  is some convenient operator.

We now define the energy functional for a given potential  $v_S(\mathbf{r})$  and determinant  $\Phi$ :

$$E_S[v_S, \Phi] = S[\Phi] + \int v_S(\mathbf{r}) n[\Phi](\mathbf{r}) d^3r \quad (9.1.6)$$

Where:

$$n[\Phi](\mathbf{r}) = \langle \Phi | \hat{n}(\mathbf{r}) | \Phi \rangle \quad (9.1.7)$$

We then minimize with respect to the orbitals of  $\Phi$ , keeping them normalized. This constrained minimization results in the orbital equations:

$$\left( \hat{O}_S + v_S(\mathbf{r}) \right) \phi_n = \epsilon_n \phi_n. \quad (9.1.8)$$

Furthermore, let us define the density functional  $F_S$  resulting from minimizing  $S$  with constrained of a given density:

$$F_S[n] = \min_{\Phi \rightarrow n} S[\Phi]. \quad (9.1.9)$$

From this we further define the density functional  $R_S[n]$ :

$$F_{LL}[n] = F_S[n] + R_S[n] \quad (9.1.10)$$

In actual implementations the functional  $R_S$  must of course be approximated. Now we put all ingredients together. The HK DFT energy is developed as a train of equalities as follows:

$$\begin{aligned}
E_v &= \min_{n \rightarrow N} \left\{ F_{LL}[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r \right\} \\
&= \min_{n \rightarrow N} \left\{ F_S[n] + R_S[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r \right\} \\
&= \min_{n \rightarrow N} \left\{ \min_{\Phi \rightarrow n} S[\Phi] + R_S[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r \right\} \\
&= \min_{\Phi \rightarrow N} \left\{ S[\Phi] + R_S[n[\Phi]] + \int v(\mathbf{r})n[\Phi](\mathbf{r})d^3r \right\}
\end{aligned} \tag{9.1.11}$$

In the last step we changed the order of the minimum procedure, assuming this is OK. Note that the final procedure is minimization with respect to orbitals instead of to density. The minimization, under normality of the orbitals  $\phi_n(\mathbf{x})$  from which the Slater wave function is built is gives the generalized Kohn-Sham equation:

$$\left( \hat{O}_S + v_R(\mathbf{r}) + v(\mathbf{r}) \right) \phi_n = \epsilon_n \phi_n \tag{9.1.12}$$

Where:

$$v_R(\mathbf{r}) = \frac{\delta R_S}{\delta n(\mathbf{r})} \tag{9.1.13}$$

If we write:

$$v_S(\mathbf{r}) = v_R(\mathbf{r}) + v(\mathbf{r}) \tag{9.1.14}$$

we find that the orbitals also obey equations (9.1.8). Thus, the density  $n(\mathbf{r})$  obtained from the orbitals of the GKS procedure is the minimizing density of the following density functional

$$E_S[v_S] = \min_{n(\mathbf{r})} [F_S[n] + \int v_S(\mathbf{r})n(\mathbf{r})d^3r] \tag{9.1.15}$$

The strength of the GKS scheme is that it incorporates explicit orbital functionals. This greatly enlarges the scope of approximations for DFT.

### ***B. Kohn-Sham from generalized KS***

By choosing:  $O_S[\Phi] = \langle \Phi | \hat{T} + \hat{U} | \Phi \rangle$  one immediately obtained the usual KS theory, with  $R_S[n] = E_C[n]$ .

### ***C. The Hybrid functional of Becke***

Let us return to the adiabatic connection formula Eq. (7.6.45):

$$E_C[n] = \int_0^1 \langle \Psi_\lambda | \hat{U} | \Psi_\lambda \rangle d\lambda - U_S[n] \quad (9.2.1)$$

Becke [21], suggested the following trapezoidal approximation for the integral:

$$\int_0^1 \langle \Psi_\lambda | \hat{U} | \Psi_\lambda \rangle d\lambda \approx \frac{\langle \Psi_0 | \hat{U} | \Psi_0 \rangle + \langle \Psi_1 | \hat{U} | \Psi_1 \rangle}{2} \quad (9.2.2)$$

Leading to the Becke's Half & Half approximation:

$$E_C[n] \approx \frac{1}{2} U_C^1 - \frac{1}{2} U_S = \frac{1}{2} (U_C^1 + E_H) - \frac{1}{2} K \quad (9.2.3)$$

The function  $(U_C^1 + E_H)[n]$  is next approximated as a local density functional:

$$(U_C^1 + E_H)[n] \approx \int u_{XC}(n(\mathbf{r})) n(\mathbf{r}) d^3r \quad (9.2.4)$$

Where  $u_{XC}(n) = \epsilon_{XC}(n) - t(n) + t_S(n)$ . the interaction energy per electron in a homogeneous electron gas is  $u_{XC}(n)$ . This function can be computed using the Quantum Monte-Carlo results concerning the XC energy.

Becke showed that this half and half theory gives a big improvement over LDA and in some cases also over GGA functionals. This is because he has found a way to approximate Hartree-Fock theory and LDA.

The Becke Half & Half orbital (BH&H) was the first example of a hybrid theory, a theory which is a mixture of a Hartree-Fock functional of orbitals

and a DFT functional of density. Becke eventually wrote down a functional which contains 3 parameters [22]:

$$\begin{aligned}
 E_{B3LYP}[\phi_1, \dots, \phi_{N_e}] \\
 \approx E_X^{\text{LSDA}}[n] + a_0(K - E_X^{\text{LSDA}}[n]) + a_x(E_X^{\text{B88}} - E_X^{\text{LSDA}}[n]) \quad (9.2.5) \\
 + a_c(E_C^{\text{PW91}}[n] - E_C^{\text{LSDA}}[n])
 \end{aligned}$$

We see that the B3LYP functional starts from a LSDA functional and adds to it 3 corrections. One is a fraction of the explicit orbital exchange correction ( $K - E_X^{\text{LSDA}}[n]$ ) then a correction ( $E_X^{\text{B88}} - E_X^{\text{LSDA}}[n]$ ) to the LSDA exchange given by a GGA type functional called Becke88 [23] and finally a GGA correction for the correlation energy, given by the PW91 functional [24]. The values of the 3 parameters  $a_0 = 0.20$ ,  $a_x = 0.72$  and  $a_c = 0.81$  were found by optimizing the performance of the functional for 56 atomization energies, 42 ionization potentials (calculated by  $\Delta SCF$  method) and 8 proton affinities.

*D.*

*E. Long-Range self-repulsion and lack of derivative discontinuity*

*F. Range separated hybrids*

*G. Orbital functionals and optimized effective potentials*

*H. Approximate correlation functionals and the Born-Oppenheimer force on nuclei*

The Kohn-Sham density functional theory method (KS-DFT), when applied for electrons of molecules makes sense only for frozen nuclei, because the external potential in DFT is assumed time-independent. However, the nuclei



in molecules are not in general motionless. How do we KS-DFT a useful approach for molecules? Here we use the Born-Oppneheimer approximation, which allows us to divide the electron-nuclear problem into 2 stages: first computing the electronic energy  $E(\{R\})$  for each nuclear configuration  $\{R\}$  and then combining it with the nuclear repulsion energy  $V_{rep}(\{R\})$  to obtain an effective potential for the motion of the nuclei. The nuclear-electron potential is given by:

$$v_{nuc}(\mathbf{r}, \{R\}) = - \sum_I \frac{Z_I e^2}{|\mathbf{r} - \mathbf{R}_I|} \quad (9.3.1)$$

In order to compute the electronic energy we first write down the functional:

$$E_{v(\{R\})}[n] = T_s[n] + \int v_{nuc}(\mathbf{r}, \{R\})n(\mathbf{r})d^3r + E_{HXC}[n] \quad (9.3.2)$$

We then minimize it under the constraint for the number of electrons  $\int n(\mathbf{r})d^3r = N$ . This gives a minimizing density  $n_*(\mathbf{r})$  and we have:

$$E[\{R\}] = E_{v(\{R\})}[n_*] = T_s[n_*] + \int v_{nuc}(\mathbf{r}, \{R\})n_*(\mathbf{r})d^3r + E_{HXC}[n_*] \quad (9.3.3)$$

Since in exact DFT  $E[\{R\}]$  is the exact electronic energy, we are assured that:

$$\mathbf{F}_I = \int -\nabla_I v_{nuc}(\mathbf{r}, \{R\})n_*(\mathbf{r})d^3r \quad (9.3.4)$$

Since this relation is true for the electronic Hamiltonian (as can be seen from Hellman-Feynman theorem).

However, what happens when we use an approximation for  $E_{HXC}$ , as done in all applications of KSDFT? Does this relation still hold? We now show that indeed it does.

The density  $n_*$  is determined by:

$$\left. \frac{\delta E_{v(\{R\})}[n]}{\delta n(\mathbf{r})} \right|_{n_*} = \mu \quad (9.3.5)$$

The electronic force on the nuclei is given by:

$$\begin{aligned} \mathbf{F}_I &= -\nabla_I E = -\nabla_I E_{v(\{R\})}[n_*] \\ &= \int -\nabla_I v_{nuc}(\mathbf{r}, \{R\}) n_*(\mathbf{r}) d^3r \\ &\quad - \int \left. \frac{\delta E_{v(\{R\})}[n]}{\delta n(\mathbf{r})} \right|_{n_*} \nabla_I n_*(\mathbf{r}) d^3r \end{aligned} \quad (9.3.6)$$

The second integral is zero because:

$$\int \left. \frac{\delta E_{v(\{R\})}[n]}{\delta n(\mathbf{r})} \right|_{n_*} \nabla_I n_*(\mathbf{r}) d^3r = \int \mu \nabla_I n_*(\mathbf{r}) d^3r = \mu \nabla_I \int n_*(\mathbf{r}) d^3r = 0 \quad (9.3.7)$$

Thus we find that the relation:

$$\mathbf{F}_I = \int -\nabla_I v_{nuc}(\mathbf{r}, \{R\}) n_*(\mathbf{r}) d^3r \quad (9.3.8)$$

still holds. Even if the approximate XC functional does not yield the exact density, the formal relation between the force and the density is still valid.

## X. More on the DFT correlation energy

### A. Approximations to $E_X$ are approximation to $E_C$

In Kohn-Sham DFT the ground state energy of a system of electrons having the particle-density  $n(\mathbf{r})$  is written as a sum of density functionals[17]:

$$E = T_s[n] + \int v(\mathbf{r})n(\mathbf{r})d^3r + E_H[n] + E_{XC}[n] \quad (10.1.1)$$

Here,  $T_s[n]$  is the kinetic energy of a system of *non-interacting fermions* having the same mass and ground-state density as the electrons,  $v(\mathbf{r})$  the external potential exerted on the electrons,  $E_H = \frac{1}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r d^3r'$  is the so-called Hartree energy of the system and  $E_{XC}[n]$  is the XC energy functional. The explicit form of the latter functional is formally known but practically inaccessible to computation. The basic challenge in DFT is to develop an efficacious approximation to the XC energy functional.

The most widely used functionals of DFT fall into two broad categories. Semilocal functionals which are based locally on the density the orbitals of the non-interacting system or on their derivatives, such as the local density approximation (LDA), generalized gradients approximation (GGA) and Meta-GGAs [23, 25] The second category includes non-local functionals such as the widely used B3LYP[22]. Recently a new brand of hybrid functionals was developed, designed to remove the long range self-interaction in DFT[26-28].

For a given density one can use the wave function  $\Psi_S$  of the non-interacting system and define the “exact exchange energy“  $E_X[n] = \langle \Psi_S | \hat{U} | \Psi_S \rangle - E_H$ , where  $\hat{U} = \frac{1}{2} \sum_{n \neq m} \frac{1}{r_{nm}}$  is the repulsion energy operator. The functional defined by  $E_C[n] = E_{XC}[n] - E_X[n]$  is called the correlation energy functional. Kohn

and Sham[17] recognized that in practical calculations it is beneficial to *approximate* the exchange energy than to use its *exact* form. Indeed it was found that the errors introduced into  $E_{XC}$  by using simple approximations to  $E_c[n]$  are partially cancelled by the approximate treatment of the exchange energy. This error cancellation is the one of the major reasons for success of DFT. Since  $E_X$  is actually known exactly and easy to calculate, it is evident that using an approximate exchange functional  $E_X^{approx}$  in place of the exact  $E_X$  is a way to correct the approximate correlation energy. For example, the correlation energy functional in the LDA is NOT the LDA of the correlation energy  $E_C^{HEG} = \int \epsilon_C^{HEG}(n(\mathbf{r}))n(\mathbf{r})d^3r$ , but instead:

$$E_C^{LDA} = E_C^{HEG} + \left[ \int \epsilon_X^{HEG}(n(\mathbf{r}))n(\mathbf{r})d^3r - E_X \right] \quad (10.1.2)$$

This view emphasizes the fact that LDA of local exchange is a way of correcting the correlation energy for inhomogeneous systems. Similar expressions are to be used in more advanced functionals (GGA etc). Any correction to  $\epsilon_X$  for example, using a GGA type expression  $\int f_X(n(\mathbf{r}), |\nabla n(\mathbf{r})|) d^3r$  should be viewed as an essentially correlation correction.

This way of thinking leads naturally to hybrid functionals. A hybrid functional is a further improvement of LDA (or GGA) obtained simply by multiplying the correction in (10.1.2) by a factor  $\lambda$ :

$$E_C^{HYB,1} = \int \epsilon_C^{HEG}(n(\mathbf{r}))n(\mathbf{r})d^3r + \lambda \left[ \int \epsilon_X^{HEG}(n(\mathbf{r}))n(\mathbf{r})d^3r - E_X \right] \quad (10.1.3)$$

Or, based on BLYP, one of the GGA functionals we can write:

$$E_C^{HYB,2} = \int f_C^{LYP}(n(\mathbf{r}), |\nabla n(\mathbf{r})|)d^3r + \lambda \left[ \int f_X^B(n(\mathbf{r}), |\nabla n(\mathbf{r})|)d^3r - E_X \right] \quad (10.1.4)$$

Becke [22] found by a series of calculations on many molecules that the “natural” assumption  $\lambda = 1$  is not optimal and  $\lambda$  should be somewhat smaller, around 0.8. This led to development of the extremely successful B3LYP functional. However, experience has shown that some systems require lower values of  $\lambda$  while others strive for higher values. An example is reaction barriers where lower values of  $\lambda$  (around 0.5-0.7) were needed. There is presently no general systematic way to set the value of  $\lambda$ .

The hybrid DFT approaches are well established electronic structure methods however in the past few years it has become increasingly clear that several types of calculations, for example, polarizability, electron transfer excitations, Rydberg states and charge allocation in weakly coupled systems are often not well described by the above DFT/TDDFT methods[29-31]. This deficiency, which is not cured by the Becke approach to hybrid functionals, was attributed to spurious self interaction[32] and missing derivative discontinuities[33] - two pervasive problems in density functional theory (DFT) that are intimately related.[34, 35]

One way to mitigate the spurious self interaction *and* to retain a good treatment of correlation is to deploy a range-separated hybrid functional[26-28, 34, 36] In this approach, the exchange term in the Kohn Sham energy functional is split into long-range and short-range terms, e.g., via  $r^{-1} = r^{-1} \text{erf}(\gamma r) + r^{-1} \text{erfc}(\gamma r)$ . The short-range exchange is represented by a local potential derived from a Semilocal functional. The long-range part is treated via an “explicit” or “exact” exchange term. If one assumes that an appropriate choice for  $\gamma$  is system independent, its value can be optimized using a molecular training set for optimizing its value. Such semi-empirical approaches, typically with  $\gamma$  in the range of 0.3-0.5  $a_0^{-1}$ , were shown to

achieve impressive results for the ground state properties of some classes of systems.[27, 37, 38] Furthermore, it was demonstrated on the benchmark model of Dreuw and Head-Gordon[39] - the  $C_2H_4 \cdots C_2F_4$  dimer at large molecular distances - that the range-separated hybrid corrects the principal deficiencies of the charge transfer excitation prediction of TDDFT.[38, 40]

One of the drawbacks of the semilocal DFT and Becke schemes is that the correlation energy of Eqs. (10.1.3) - (10.1.4) is much too nonlocal due to the presence of the nonlocal exchange  $E_X$ . This is unbalanced in view of the presence of  $E_H$  in the total energy. This is the root cause of long-range "self repulsion" which exists in all these functionals. A way to improve the situation is to include in the exchange functional a non-local long-range exchange, which leads to the following form for the correlation[26, 27, 41]:

$$E_C^{HYB,d} = \int f_C^{LYP}(n(\mathbf{r}), |\nabla n(\mathbf{r})|) d^3r + \int \left( f_X^B(n(\mathbf{r}), |\nabla n(\mathbf{r})|) - f_{X,\gamma}(n(\mathbf{r}), |\nabla n(\mathbf{r})|) \right) d^3r - (E_X - E_{X,\gamma}) \quad (10.1.5)$$

The functional  $E_{X,\gamma}$  is the exchange energy of a system of particles at density  $n$  mutually interacting through the repulsive potential  $u_\gamma(r) = \frac{\text{erf}(\gamma r)}{r}$  where  $\gamma$  is a parameter:

$$E_{X,\gamma} = -\frac{1}{2} \iint |\rho(\mathbf{r}, \mathbf{r}')|^2 u_\gamma(r) d^3r d^3r'. \quad (10.1.6)$$

$f_X - f_{X,\gamma}$  is a semilocal density approximation to the complementary interaction  $y_\gamma(r) = \frac{1}{r} - u_\gamma(r) = \frac{\text{erfc}(\gamma r)}{r}$ . The idea behind this appealing approach is to eliminate in the correlation functional the long range dependence on the density.  $\gamma$  is treated as a universal parameter determined by fitting the DFT predictions based on Eq. (10.1.5) to a large database of

experimental data. This forms the basis of several recent semi-empirical approaches all of which use a universal long-range parameter  $\gamma$  determined by semiempirical fitting to known experimental and sometimes high quality ab-initio results [26-28, 34, 36-38]. These approaches have been quite successful for a broad variety of molecules, sometimes beyond the types used in the fitting procedures.

The approach of a universal  $\gamma$  exchange functional with complementary *semilocal* functionals cannot however claim universality. In some systems the exact value of  $\gamma$  is critical and no semilocal DFT method can correct for it. We show examples below. An alternative view was developed in ref. [28], which provides a theory for the long-range parameter  $\gamma$  based on the adiabatic connection theorem, from which the following exact expression for the correlation energy can be deduced:

$$E_C = \langle \Psi | \hat{Y}_\gamma | \Psi \rangle - \langle \Psi_S | \hat{Y}_\gamma | \Psi_S \rangle \quad (10.1.7)$$

Where  $\hat{Y}_\gamma = \frac{1}{2} \sum_{n \neq m} \gamma_\gamma(r_{nm})$ . In ref. [28] arguments were given for the existence of a parameter  $\gamma$  for which Eq. (10.1.7) holds for practically all densities. This equation describes the correlation energy as a difference between potential energies of interaction in the interacting and in the non-interacting systems. Unlike the standard definition of correlation energy as  $E_C = \langle \Psi | \hat{T} | \Psi \rangle - \langle \Psi_S | \hat{T} | \Psi_S \rangle + \langle \Psi | \hat{U} | \Psi \rangle - \langle \Psi_S | \hat{U} | \Psi_S \rangle - E_X$ , there are no explicit contributions of kinetic energy differences in Eq. (10.1.7). These are absorbed by the interaction screening parameter  $\gamma$ . In principle all quantities in this equation, including  $\gamma$ , are density-dependent. We stress that the difference between this exact theory and the usual exposition of RSHs is that the latter do not derive (or even strive to derive) an exact expression for the correlation energy. Eq. (10.1.7) can be applied to the HEG and the exact value

of  $\gamma$  can be determined as a function of density[38, 42]. The result is shown in Figure X-1 and it is seen that  $\gamma$  is in fact strongly density dependent, certainly not a universal constant. In particular, for molecular densities, where  $r_s$  is between 0.5 and 5 the long-range parameter  $\gamma$  changes by an order of magnitude.

When we add the exact exchange energy to the exact correlation energy of Eq (10.1.7) we obtain for the exchange-correlation the following expression[28]:

$$\begin{aligned}
 E_{XC} &= \langle \Psi | \hat{Y}_\gamma | \Psi \rangle \\
 &- \frac{1}{2} \int n(\mathbf{r})n(\mathbf{r}')y_\gamma(|\mathbf{r} - \mathbf{r}'|)d^3rd^3r' + E_{X,\gamma} \\
 &\equiv E_{XC}^\gamma + E_{X,\gamma}
 \end{aligned} \quad (10.1.8)$$

This expression is exact. The right-

hand equivalence defines a new functional, the complementary XC energy functional  $E_{XC}^\gamma$ . It is the focus for approximation in our proposal. Because the potential  $y_\gamma(\mathbf{r})$  is short-range, we expect that a semilocal approximation, depending on density and perhaps gradients, is of sufficient accuracy for describing the “complementary” XC energy.

Except for the HEG, we have no practical way of calculating the expectation values in Eq. (10.1.8) so an approximation must be developed for determining the long-range parameter  $\gamma$  and for determining the complementary XC energy functional  $E_{XC}^\gamma$ . The procedure for determining the variable  $\gamma$  will be

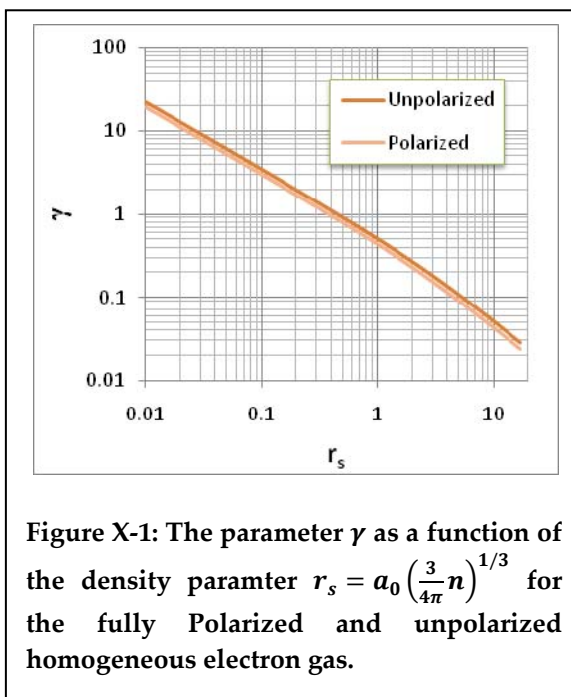


Figure X-1: The parameter  $\gamma$  as a function of the density parameter  $r_s = a_0 \left(\frac{3}{4\pi}n\right)^{1/3}$  for the fully Polarized and unpolarized homogeneous electron gas.



called below “long-range parameter tuning”. We now show some preliminary results which demonstrate the fact that one cannot rely on a universal long-range parameter  $\gamma$ .

### ***B. The necessity of long-range parameter tuning***

The requirement for a system-dependent parameter is not just a theory it comes up in practical calculations. Here we give two examples of unpublished results, one in DFT and one in TDDFT.

#### **i. The symmetric radical cation**

In DFT, a clear example of this issue is seen in the symmetric radical cation systems of the type  $R_2^+ \rightarrow R + R^+$ , where  $R$  is any molecule or atom. Specific examples are  $R = \text{H}, \text{He}$  and  $\text{Ne}$ . It is known that for these systems a semilocal or even nonlocal (B3LYP) DFT gives qualitatively spurious potential surfaces and the big issue is self-repulsion[1, 43]. Yet, even with a standard  $\gamma$ -functional the results are not satisfactory. Take for example the bond dissociation energy (BDE). This can be estimated in two ways. The first is the atomization energy. The second is the depth of the molecular well relative to the asymptote of the potential surface. From both numbers we subtract the zero-point vibrational energy to get the BDE. We can see in Table 1 that the BDE of  $R=\text{He}$  and  $\text{Ne}$  based on atomization energy using a semi-empirical  $\gamma = 0.5 a_0^{-1}$  (denoted BNL functional[38]) gives BDE around 1.5-2 times the experimental value! While in the asymptotic method the calculated BDE is too small by 15%. Such discrepancies cannot be waived by a better semilocal functional. They reflect an inherent imbalance between short and long range exchange inflected by a *wrong* long-range parameter  $\gamma$ . Indeed, by tuning the long-range parameter in the following way we obtain a value  $\gamma^*$  which is specific for each system. The

tuning procedure is “ab initio” in that it does not use any experimental data, and results from the calculations themselves. It is describe in ref. [1] and is equivalent to the procedure depicted in a more general treatment in ref [38].

The resulting functional is denoted BNL\*. This procedure chooses  $\gamma^*$  in such a way as to impose the “energy removal theorem” namely that the energy to remove an electron from the interacting

**Table 1: (Taken from ref [1]) Data for  $R_2^+$ ,  $R=H, He$  and  $Ne$  calculated in the cc-pVTZ basis[2, 3]. All energies in kcal/mole.  $r_{eq}$  is the inter-nuclear distance at which the  $E(r)$  is minimal.  $\omega = \sqrt{2\mu^{-1}E''(r_{eq})}$  is the harmonic frequency and “BDE” is  $E_{atoms} - E(r_{eq}) + \frac{1}{2}\hbar\omega$  where  $E_{atoms}$  is estimated either as the sum of atomic energies (SUM) or by the asymptotic value of the potential energy curve (ASS).  $\alpha_R$  is the calculated polarizability of the atom  $R$  and  $\alpha_{eff} = \lim_{r \rightarrow \infty} \frac{1}{2}E'(r)r^5$  is the polarizability as estimated from the form of the asymptotic potential. Finally  $\alpha_R^{daug}$  is the polarizability of atom  $R$  computed with a large basis-set.**

Property	R	BLYP	B3LYP	HF	BNL	BNL*	Exp.[4]
BDE by $E_{atoms}^{(sum)}$	H	66	65	60.9	60.9	60.9	61
	He	82	75	43	74	59	55
	Ne	75	60	2	59	34	32
BDE by $E_{atoms}^{(ass)}$	H	NA	NA	60.9	50	60.9	61
	He	NA	NA	43	42	59	55
	Ne	NA	NA	2	27	34	32
$r_{eq}$ (Å)	H	1.1	1.1	1.06	1.2	1.06	1.05
	He	1.2	1.1	1.075	1.2	1.078	1.080
	Ne	1.9	1.9	1.7	1.760	1.72	1.765
$\hbar\omega/2$	H	2.7	2.9	3.3	2.9	3.3	3.32
	He	1.7	2.0	2.5	2.1	2.5	2.42
	Ne	0.5	0.6	0.9	0.726	0.8	0.729
$\alpha_R/\alpha_{eff}$	H	NA	NA	1	0.6	1	1
	He	NA	NA	0.98	NA	0.98	1
	Ne	NA	NA	1.01	NA	1.02	1
$a_R^{daug}$ ( $a_0^3$ )	H	5.3	5.6	4.51	5.8	4.51	4.50
	He	1.6	1.5	1.34	1.8	1.41	1.38
	Ne	3.1	2.9	2.4	3.2	2.70	2.66

system is equal exactly to  $-\epsilon_{HOMO}$  (the HOMO energy in the non-interacting system).[44] The parameter  $\gamma^*$ , as we compute it, obeys the equation  $-\epsilon_{HOMO}^{\gamma^*} = E^{\gamma^*}(N) - E^{\gamma^*}(N - 1)$  where  $\epsilon_{HOMO}^{\gamma}$  is the HOMO energy and  $E^{\gamma}$  is the calculated ground-state energy of the  $N$ -electron system. This procedure approximately imposes the energy removal theorem and defines BNL\*. The merit is that the HOMO and the SCF procedures are now well-balanced. Unlike the results of BNL, where  $\gamma$  is not optimal, those of BNL\* are very satisfactory for several observables when compared to experiment, as seen in

Table 1. The comparison with regular functionals is convincing and the improvement over BNL and Hartree-Fock is also large. Even the complete basis set limit of the atomic polarizability as calculated by BNL\* compares accurately with experiment.

## ii. Aromatic donor -TCNE acceptor charge transfer excitation

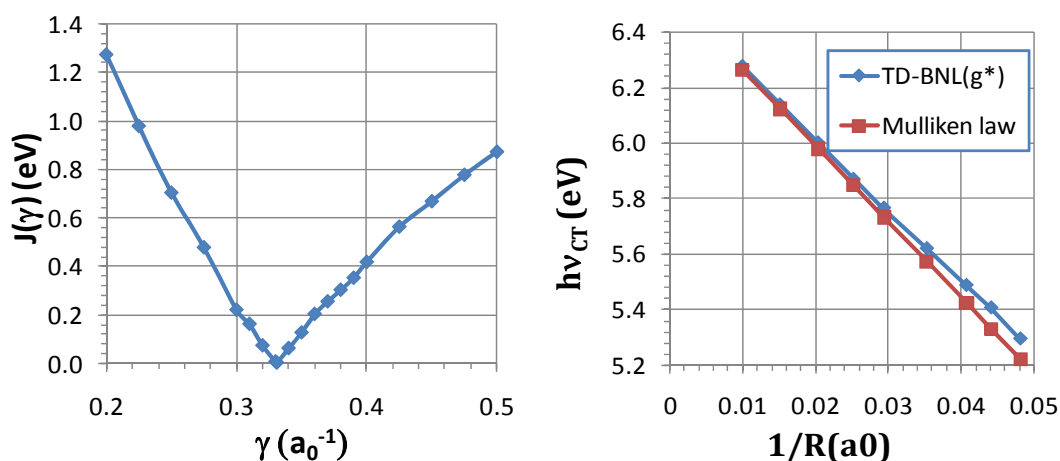
The second example concerns charge-transfer excitations in TDDFT. The results shown here are part of a manuscript recently submitted for publication[45]. It is known that semilocal functionals cannot reasonably describe such excitations[39], so a range-separated hybrid is a welcome cure. We tested our approach on complexes formed by an aromatic donor (Ar=benzene, toluene, o-xylene and naphthalene) and the tetracyanoethylene (TCNE) acceptor, for which optical absorption is available both in gas phase and in solution.[46] All calculations were performed using QCHEM 3.1,[47] modified to include the range-separated BNL functional,[38] using the cc-pVDZ basis set.[3] The internal structure of the molecules in the complex is known to be little-perturbed by complex formation[30, 31, 48, 49] and the equilibrium distance and relative orientation of the  $\pi$ -stacked donor and acceptor determined from the conventional B3LYP[21] hybrid functional is known to compare well with experiment (where available).[30, 49] Therefore, B3LYP-optimized geometry was used throughout. The TDB3LYP excitations energies were much too low and the BNL much too high (see Table 1). In molecular complexes, the lowest photon energy required to induce a CT excitation,  $h\nu_{CT}$ , is given for asymptotically large donor-acceptor distances by the Mulliken rule[50]

$$h\nu_{CT} = IP(D) - EA(A) - 1/R \quad (10.1.9)$$

Where  $IP(D)$  and  $EA(A)$  are the donor ionization potential and acceptor electron affinity, respectively. The last term on the right hand side is the Coulomb energy of attraction between the electron-hole pair formed by the charge transfer, where  $R$  is the inter-molecular separation. For our TDDFT calculation to conform to the Mulliken rule, the ionization energies computed from Eq. (1) must correspond to the HOMO energy of the neutral donor, but also to the HOMO energy of acceptor *anion*. Thus, one needs to generalize Eq. (1) so as to yield, as closely as possible, two limits. We therefore look for  $\gamma$  that minimizes the following  $J(\gamma)$  function

$$J(\gamma) = \sum_{i=D^0, A^-} \left| \epsilon_{HOMO}^{\gamma, i} - \left( E_\gamma(N_i - 1) - E_\gamma(N_i) \right) \right| \quad (10.1.10)$$

For complexes where a range-parameter  $\gamma$  that renders  $J(\gamma)$  very small can be found (which is the case for all complexes we studied), we expect the range-separated hybrid to yield a quantitative description of CT excitations. An example of the way this tuning procedure works and its ability to replicate the Mulliken law for large separations  $R$  is shown in Figure X-2



**Figure X-2: Tuning  $\gamma$  in the Benzene-TCNE complex. Left:  $J$  of Eq. Error! Reference source not found. as a function of  $\gamma$ . The optimal  $\gamma$  is 0.331 for which  $J$  is very small. Right: The Mulliken rule, compared to TDDFT results obtained from the optimal  $\gamma$ , as a function of inverse inter-molecular separation.**

With the asymptotic behavior enforced, we expect proper balance between the semilocal and non-local exchange components. In **Error! Reference source not found.** we compare the calculated and experimental gas-phase results for various Ar-TCNE complexes. It is readily seen that the B3LYP results are unacceptably low and predictive power is absent. Results of generalized-gradient (GGA) calculations (not shown) are even lower than the B3LYP ones. With the range-separated BNL functional,[38] excitation energies determined with an "off-the-shelf"  $\gamma$  value of  $0.5 a_0^{-1}$  are much too high with respect to experiment. But with  $\gamma^*$  quantitative agreement is obtained to within  $\pm 0.2$ eV. For benzene and toluene, the theoretical oscillator strengths are also in good agreement with experiment, but are too weak for xylene and naphthalene. This is likely a basis-set issue, as oscillator strengths are much more sensitive to the basis set than the excitation energies.

## XI . TDDFT

### A. Time-dependent Linear response theory

The linear response of

$$i\hbar \frac{\partial}{\partial t} \psi_k(\mathbf{r}, t) = \left( \hat{H}[\rho(t)] + v^1(\mathbf{r}, t) \right) \psi_k(\mathbf{r}, t) \quad (10.2.1)$$
$$\psi_k(\mathbf{r}, 0) = \psi_k^0(\mathbf{r})$$

Where the TD Hamiltonian, in the *adiabatic* TDDFT approximation, is:

$$\hat{H}[\rho] = \hat{T} + v(\mathbf{r}) + v_H[n(t)](\mathbf{r}) + v_{XC}^Y[n(t)](\mathbf{r}) + \hat{K}^V[\rho] \quad (10.2.2)$$

And the TD density matrix is defined in terms of the TD orbitals:

$$\rho(\mathbf{r}, \mathbf{r}', t) = \sum_{k=1}^N \psi_k(\mathbf{r}, t) \psi_k(\mathbf{r}', t)^* \quad (10.2.3)$$

The initial orbitals are assumed the lowest eigenstates of  $\hat{H}[\rho^0]$ :

$$\hat{H}[\rho^0] \psi_k^0(\mathbf{r}) = \epsilon_k \psi_k^0(\mathbf{r}) \quad (10.2.4)$$

where:

$$\rho^0(\mathbf{r}, \mathbf{r}') = \sum_{k=1}^N \psi_k^0(\mathbf{r}) \psi_k^0(\mathbf{r}')^* \quad (10.2.5)$$

We assume that the perturbation  $v^1(\mathbf{r}, t)$  is weak and we will treat it as a first order quantity. Before we do this however, let us prepare the equation for linearization by defining:

$$\psi_k(\mathbf{r}, t) = e^{-i\epsilon_k t/\hbar} \left( \psi_k^0(\mathbf{r}) + \psi_k^1(\mathbf{r}, t) \right) \quad (10.2.6)$$

and:

$$\begin{aligned}\rho(\mathbf{r}, \mathbf{r}', t) &= \rho^0(\mathbf{r}, \mathbf{r}') + \rho^1(\mathbf{r}, \mathbf{r}', t) \\ n(\mathbf{r}, t) &= n^0(\mathbf{r}) + n^1(\mathbf{r}, t)\end{aligned}\tag{10.2.7}$$

With these definitions we have:

$$\begin{aligned}i\hbar \frac{\partial}{\partial t} \left[ e^{-i\epsilon_k t/\hbar} \left( \psi_k^0(\mathbf{r}) + \psi_k^1(\mathbf{r}, t) \right) \right] \\ = \left( \hat{H}[\rho^0 + \rho^1(t)] \right. \\ \left. + v^1(\mathbf{r}, t) \right) \left[ e^{-i\epsilon_k t/\hbar} \left( \psi_k^0(\mathbf{r}) + \psi_k^1(\mathbf{r}, t) \right) \right]\end{aligned}\tag{10.2.8}$$

From this we find:

$$i\hbar \frac{\partial}{\partial t} \psi_k^1(\mathbf{r}, t) = \left( \hat{H}[\rho^0 + \rho^1(t)] - \epsilon_k + v^1(\mathbf{r}, t) \right) \left( \psi_k^0(\mathbf{r}) + \psi_k^1(\mathbf{r}, t) \right)\tag{10.2.9}$$

This equation is rigorously equivalent to the first equation. It is a non-homogeneous linear equation for the unknown  $\psi_k^1(t)$ .

We now make the linearization step. We neglect all quantities beyond first order. As a first step we have:

$$i\hbar \frac{\partial}{\partial t} \psi_k^1(\mathbf{r}, t) \approx \left( \hat{H}^1 + v^1(\mathbf{r}, t) \right) \psi_k^0(\mathbf{r}) + \left( \hat{H}[\rho^0] - \epsilon_k \right) \psi_k^1(\mathbf{r}, t)\tag{10.2.10}$$

Here, the symbol “ $\approx$ ” means “equal to first order”. We use the definition:

$$\hat{H}^1 \approx \hat{H}[\rho^0 + \rho^1(t)] - \hat{H}[\rho^0]\tag{10.2.11}$$

An efficient method to obtain linear response of a perturbation is as follows. Select a small parameter  $\varepsilon$  and solve the following time dependent equation (using a standard propagator, such as Runge Kutta)[51]:

$$\begin{aligned}i\hbar \frac{\partial}{\partial t} \psi_k^1(\mathbf{r}, t) &= \left( \frac{\hat{H}[\rho^0 + \varepsilon \rho^1(t)] - \epsilon_k}{\varepsilon} + v^1(\mathbf{r}, t) \right) \psi_k^0(\mathbf{r}) \\ &+ \left( \hat{H}[\rho^0] - \epsilon_k \right) \psi_k^1(\mathbf{r}, t)\end{aligned}\tag{10.2.12}$$

With:

$$\rho^1(\mathbf{r}, \mathbf{r}', t) \approx \sum_{k=1}^N \psi_k^0(\mathbf{r}) \psi_k^1(\mathbf{r}', t)^* + \sum_{k=1}^N \psi_k^1(\mathbf{r}, t) \psi_k^0(\mathbf{r}')$$

$$n^1(\mathbf{r}, t) \approx 2\text{Re} \sum_{k=1}^N \psi_k^0(\mathbf{r}) \psi_k^1(\mathbf{r}, t)$$

The size of  $\varepsilon$  must be so small that the orbital amplitudes are linear with  $\varepsilon$ . The solution involves repeated applications of the Hamiltonian. This could be considerably more efficient than the algebraic method described below, requiring the computation of all the eigenvectors of the Hamiltonian. Of course, for small systems with a small basis the algebraic approach may be extremely effective thus we discuss it now.

### ***B. Algebraic Approach to Linear response***

In order to derive an algebraic approach, we need expressions for all the linear quantities. For  $\hat{H}^1$  we find:

$$\hat{H}^1 \approx v_H^1 + v_{XC}^{\gamma 1} + K^{\gamma 1} \quad (10.3.1)$$

And:

$$v_H^1(\mathbf{r}, t) = \int \frac{n^1(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} d^3 r'$$

$$v_{XC}^{\gamma 1}(\mathbf{r}) = \int f_{XC}^{\gamma}(\mathbf{r}, \mathbf{r}') n_1(\mathbf{r}', t) d^3 r' \quad (10.3.2)$$

$$\hat{K}^{\gamma 1} f(\mathbf{r}) = - \int \rho^1(\mathbf{r}, \mathbf{r}', t) u_{\gamma}(|\mathbf{r} - \mathbf{r}'|) f(\mathbf{r}') d^3 r'$$

Plugging all this into Eq. (10.2.10):



$$\begin{aligned}
i\hbar \frac{\partial}{\partial t} \psi_k^1(\mathbf{r}, t) \approx & \left( \int \left\{ \frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{XC}^Y(\mathbf{r}, \mathbf{r}') \right\} n_1(\mathbf{r}', t) d^3r' \right. \\
& \left. + v^1(\mathbf{r}, t) \right) \psi_k^0(\mathbf{r}) - \int \rho^1(\mathbf{r}, \mathbf{r}', t) u_\gamma(|\mathbf{r} - \mathbf{r}'|) \psi_k^0(\mathbf{r}') d^3r' \\
& + (\hat{H}[\rho^0] - \epsilon_k) \psi_k^1(\mathbf{r}, t)
\end{aligned} \tag{10.3.3}$$

Plugging in the 1<sup>st</sup> order definitions of  $n^1$  and  $\rho^1$  we find:

$$\begin{aligned}
i\hbar \frac{\partial}{\partial t} \psi_k^1(\mathbf{r}, t) \approx & 2Re \sum_{j=1}^N \int W_{kj}(\mathbf{r}, \mathbf{r}') \psi_j^1(\mathbf{r}', t) d^3r' \\
& + (\hat{H}[\rho^0] - \epsilon_k) \psi_k^1(\mathbf{r}, t) + v^1(\mathbf{r}, t) \psi_k^0(\mathbf{r}) \\
W_{kj}(\mathbf{r}, \mathbf{r}') \equiv & \left[ \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_k^0(\mathbf{r}) \psi_j^0(\mathbf{r}') - u_\gamma(|\mathbf{r} - \mathbf{r}'|) \psi_k^0(\mathbf{r}') \psi_j^0(\mathbf{r}) \right. \\
& \left. + f_{XC}^Y(\mathbf{r}, \mathbf{r}') \psi_k^0(\mathbf{r}) \psi_j^0(\mathbf{r}') \right]
\end{aligned} \tag{10.3.4}$$

We now take a crucial step to get rid of all the integrals and special dependence by expanding in terms of the orthonormal complete eigenfunction of the Hamiltonian  $H[\rho^0]$ . Because the evolution according to the Kohn-Sham equation preserves orthonormality. Thus, the response functions are orthogonal to all occupied MO's. This allows us to write:

$$\psi_k^1(\mathbf{r}, t) = \sum_{s=N+1}^M (c'_{ks}(t) + ic''_{ks}(t)) \psi_s^0(\mathbf{r}) \tag{10.3.5}$$

We use the following index notations:

$$\begin{aligned}
j, k & - \text{occupied orbitals} \\
q, s & - \text{unoccupied orbitals} \\
p, n & - \text{all orbitals}
\end{aligned}$$

The  $W$  operator can be turned into a 4-index matrix:

$$\begin{aligned}
W_{kp,jn} &= \iint W_{kj}(\mathbf{r}, \mathbf{r}') \psi_p^0(\mathbf{r}) \psi_n^0(\mathbf{r}') d^3r d^3r' \\
&= \iint \left[ \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{XC}^Y(\mathbf{r}, \mathbf{r}') \right) \psi_k^0(\mathbf{r}) \psi_j^0(\mathbf{r}') \right. \\
&\quad \left. - u_Y(|\mathbf{r} - \mathbf{r}'|) \psi_k^0(\mathbf{r}') \psi_j^0(\mathbf{r}) \right] \psi_p^0(\mathbf{r}) \psi_n^0(\mathbf{r}') d^3r d^3r' \quad (10.3.6) \\
\rightarrow W_{kj}(\mathbf{r}, \mathbf{r}') &= \sum_{pn} W_{kp,jn} \psi_p^0(\mathbf{r}) \psi_n^0(\mathbf{r}')
\end{aligned}$$

Using these expansions we find:

$$\begin{aligned}
\sum_s i\hbar(\dot{c}'_{ks}(t) + ic''_{ks}(t)) \psi_s^0(\mathbf{r}) \\
&= 2 \sum_{s,j,p} c'_{js}(t) W_{kp,js} \psi_p^0(\mathbf{r}) \quad (10.3.7) \\
&\quad + \sum_s (c'_{ks}(t) + ic''_{ks}(t)) \omega_{sk} \psi_s^0(\mathbf{r}) + v^1(\mathbf{r}, t) \psi_k^0(\mathbf{r})
\end{aligned}$$

with:  $\hbar\omega_{sk} = \epsilon_s - \epsilon_k$ . Now multiply by  $\psi_q(\mathbf{r})$  and integrate over  $\mathbf{r}$  and obtain a time-dependent algebraic equation:

$$\begin{aligned}
i\hbar(\dot{c}'_{kq}(t) + ic''_{kq}(t)) \\
&= 2 \sum_s \sum_j W_{kq,js} c'_{js}(t) + (c'_{kq}(t) + ic''_{kq}(t)) \hbar\omega_{qk} \quad (10.3.8) \\
&\quad + v_{kq}^1(t)
\end{aligned}$$

Where:

$$v_{kq}^1(t) = \int \psi_q^0(\mathbf{r}) v^1(\mathbf{r}, t) \psi_k^0(\mathbf{r}) d^3r \quad (10.3.9)$$

We can separate to two real equations by the real and imaginary parts:

$$-\hbar\dot{c}_{kq}''(t) = 2 \sum_s \sum_j W_{kq,js} c_{js}'(t) + c_{kq}'(t) \hbar\omega_{qk} + v_{kq}^1(t) \quad (10.3.10)$$

$$\dot{c}_{kq}'(t) = c_{kq}''(t) \omega_{qk}$$

This equation can be used to get a time-dependent response. It will be useful for describing the response to a complicated weak pulse  $v_{kq}^1(t)$ .

### ***C. Frequency-domain response***

Instead of viewing the response in time we can view it in frequency. This will allow us to address questions of response to sinusoidal fields, such as laser absorption. We Fourier-transform the coefficients:

$$d_{kq}(\omega) = \int_0^\infty c_{kq}(t) e^{i\omega t} dt \quad (10.3.11)$$

Note that  $c_{kq}(0) = 0$  and so:  $-i\omega d_{kq}(\omega) = \int_0^\infty \dot{c}_{kq}(t) e^{i\omega t} dt$ .

Note further that  $d'$  and  $d''$  are not real anymore. In fact, because  $c$ 's are real we now have:  $d_{kq}(\omega)^* = d_{kq}(-\omega)$ .

Applying a Fourier transform to the time-dependent equations gives:

$$i\hbar\omega d_{kq}''(\omega) = 2 \sum_{js} W_{kq,js} d_{js}'(\omega) + d_{kq}'(\omega) \hbar\omega_{qk} + v_{kq}^1(\omega) \quad (10.3.12)$$

$$-i\omega d_{kq}'(\omega) = d_{kq}''(\omega) \omega_{qk}$$

Combining the equations we rid ourselves of  $d'$  and obtain an equation for  $d''$ :

$$\sum_{js} [2\hbar^{-1}\omega_{qk} W_{kq,js} + (\omega_{qk}^2 - \omega^2) \delta_{(qk)(sj)}] d_{kq}''(\omega) = i\hbar^{-1}\omega v_{kq}^1(\omega), \quad (10.3.13)$$

The number of these coupled algebraic equations is infinite. However, in any practical application we use a finite number of states and so we have a finite number of equations. If  $M$  is the total number of states used we then for any

given perturbation one can get an approximate response by solving one of these two equations for  $d'_{js}(\omega)$ . This involves solving a linear  $NM \times NM$  set of equations and obtain the response to the perturbation. We can write it in matrix form:

$$(A - \omega^2 I)d''(\omega) = i\hbar^{-1}\omega v^1(\omega). \quad (10.3.14)$$

With:

$$A_{qkjs} = 2\hbar^{-1}\omega_{qk}W_{kq,js} + \omega_{qk}^2\delta_{(qk)(sj)} \quad (10.3.15)$$

We now prove that as long as  $\omega_{qk} \neq 0$  for all possible  $q$  and  $k$  the matrix  $A$  has only real eigenvalues. In fact,  $\omega_{qk}$  is always positive unless there is a degeneracy at the HOMO and the level is not full. The eigenvalue equation is:

$$(2\hbar^{-1}\omega_{qk}W_{kq,js} + \omega_{qk}^2\delta_{(qk)(sj)})y_{sj} = \Omega^2 y_{qk} \quad (10.3.16)$$

where  $y_{sj}$  are the eigenvector components. Under the assumption we can bring this equation to the following equivalent form:

$$\sum_{js} (2\hbar^{-1}W_{kq,js}\sqrt{\omega_{qk}\omega_{sj}} + \omega_{qk}^2\delta_{(qk)(sj)})x_{sj} = \Omega^2 x_{qk} \quad (10.3.17)$$

Where  $x_{qk} = \frac{y_{qk}}{\sqrt{\omega_{qk}}}$ . These new equations involve the eigenvalues of a symmetric matrix. Thus, the eigenvalues  $\Omega^2$  are all real. Eq. (10.3.17) is called the Casida equations.

#### ***D. Excitation energies from the algebraic treatment***

The algebraic equation for linear response derived in Eq. (10.3.14) allows calculation of a specific response problem. One incredible feature of this equation is that the matrix  $A$  on the left is independent of the perturbation  $v^1$ . We can thus analyze  $A$  to get general response properties of the system,

without reference to any specific perturbation. We can define a “response matrix”:

$$R(\omega) = i\hbar^{-1}\omega(A - \omega^2I)^{-1}. \quad (10.3.18)$$

Where  $A$  is defined in (10.3.15) We can define right and left eigenvectors:

$$Ay_\alpha = \Omega_\alpha^2 y_\alpha, \quad A^T z_\alpha = \Omega_\alpha^2 z_\alpha \quad (10.3.19)$$

It is easy to see from Eq. (10.3.16) that:

$$\omega_{qk}(z_\alpha)_{qk} = (y_\alpha)_{qk} \quad (10.3.20)$$

We choose the normalizations so that:

$$z_\alpha^T y_\beta = \delta_{\alpha\beta} = \sum_{js} (z_\alpha)_{js} (y_\beta)_{js} = \sum_{js} \frac{1}{\omega_{sj}} (y_\alpha)_{js} (y_\beta)_{js} \quad (10.3.21)$$

Collecting the vectors in a matrix and the eigenvalues in a diagonal matrix  $\Omega^2$  we can write:

$$A = y^T \Omega^2 z \quad (10.3.22)$$

And:

$$R(\omega) = y^T \frac{i\hbar^{-1}\omega}{(\Omega^2 - \omega^2 I)} z \quad (10.3.23)$$

We can now take  $\omega$  near an eigenvalue  $\Omega_\alpha$ . This has an infinite response. Thus we can then write:

$$R(\omega) = y_\alpha^T \frac{i\hbar^{-1}\omega}{(\Omega_\alpha^2 - \omega^2)} z_\alpha \quad \omega \rightarrow \Omega_\alpha \quad (10.3.24)$$

A general technique to study the resonance is to insert an infinitesimal imaginary part which we eventually take to zero:  $\omega = \Omega_\alpha + i\eta$ . Then, after a few algebraic manipulations  $(R(\omega) = y_\alpha^T \frac{i\omega}{(\Omega_\alpha + \omega)(\Omega_\alpha - \omega)} z_\alpha \rightarrow R(\Omega_\alpha + i\eta) = y_\alpha^T \frac{i(\Omega_\alpha + i\eta)}{(2\Omega_\alpha + i\eta)(-i\eta)} z_\alpha)$ :

$$R(\Omega_\alpha + i\eta) = \hbar^{-1} y_\alpha^T \frac{\Omega_\alpha (2\Omega_\alpha + i\eta)}{-\eta(4\Omega_\alpha^2 + \eta^2)} z_\alpha \quad (10.3.25)$$

Thus:

$$\begin{aligned} \lim_{\eta \rightarrow \infty} \Re[R(\Omega_\alpha + i\eta)] &= \infty \\ \lim_{\eta \rightarrow \infty} \Im[R(\Omega_\alpha + i\eta)] &= y_\alpha^T \frac{-1}{4\hbar\Omega_\alpha} z_\alpha \end{aligned} \quad (10.3.26)$$

The imaginary part of the response is finite at resonances.

## References

- 1 E. Livshits and R. Baer, *J. Phys. Chem. A* **112**, 12789 (2008).
- 2 D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **100**, 2975 (1994); R. M. Dreizler and E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many Body Problem* (Springer, Berlin, 1990); R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989); W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley, Heidelberg, 2001); C. Fiolhais, F. Nogueira, and M. Marques, *A primer in density functional theory* (Springer, Berlin ; New York, 2003).
- 3 T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- 4 K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure, Vol 4: Constants of Diatomic Molecules* (Prentice-Hall, New York, 1979); A. Dalgarno, *Adv. Phys.* **11**, 281 (1962); F. Furche and R. Ahlrichs, *J. Chem. Phys.* **117**, 7433 (2002); F. Neese, *Journal of Biological Inorganic Chemistry* **11**, 702 (2006); M. Marques, A. Rubio, C. A. Ullrich, et al., in *Lecture notes in physics*, 706 (Springer, Berlin, 2006).
- 5 H. Goldstein, C. P. Poole, and J. L. Safko, *Classical mechanics* (Addison Wesley, San Francisco, 2002).
- 6 J. C. Slater, *J. Chem. Phys.* **1**, 687 (1933).
- 7 B. N. Finkelstein and G. E. Horowitz, *Z. Phys.* **48**, 118 (1928).
- 8 I. N. Levine, *Quantum chemistry* (Prentice Hall, Upper Saddle River, N.J., 2000); F. Rioux, *The Chemical Educator* **2**, 1 (1997).
- 9 R. Peierls, *Phys. Rev* **54**, 918 (1938).
- 10 R. Granot and R. Baer, *J. Chem. Phys.* **128**, 184111 (2008).
- 11 E. Teller, *Rev. Mod. Phys.* **34**, 627 (1962).
- 12 U. Salzner and R. Baer, *J. Chem. Phys.* **in press** (2010).
- 13 J. Janak, *Phys. Rev. B* **18**, 7165 (1978).
- 14 J. C. Slater and J. H. Wood, *Int. J. Quant. Chem.* **S-4**, 3 (1971).
- 15 R. Baer and M. Head-Gordon, *Phys. Rev. Lett.* **79**, 3962 (1997).
- 16 D. C. Langreth and J. P. Perdew, *Sol. Stat. Comm.* **17**, 1425 (1975).
- 17 W. Kohn and L. J. Sham, *Phys. Rev* **140**, A1133 (1965).
- 18 M. Gellmann and K. A. Brueckner, *Phys. Rev* **106**, 364 (1957).
- 19 E. Wigner, *Phys. Rev* **46**, 1002 (1934).
- 20 O. Gunnarsson and B. I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976).
- 21 A. D. Becke, *J. Chem. Phys.* **98**, 1372 (1993).
- 22 A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- 23 A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).

- 24 J. P. Perdew, in *Electronic structure of solids '91 : proceedings of the 75. WE-Heraeus-Seminar and 21st Annual International Symposium on Electronic Structure of Solids held in Gaussig (Germany), March 11-15, 1991*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), Vol. 17, p. 11; J. P. Perdew, J. A. Chevary, S. H. Vosko, et al., *Phys. Rev. B* **48**, 4978 (1993); J. P. Perdew, J. A. Chevary, S. H. Vosko, et al., *Phys. Rev. B* **46**, 6671 (1992).
- 25 P. Hohenberg and W. Kohn, *Phys. Rev* **136**, B864 (1964); S. H. Vosko and L. Wilk, *J. Phys. B* **16**, 3687 (1983); J. P. Perdew and W. Yue, *Phys. Rev. B* **33**, 8800 (1986); C. T. Lee, W. T. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988); J. P. Perdew and Y. Wang, *Phys. Rev. B* **45**, 13244 (1992); J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996); J. P. Perdew, S. Kurth, A. Zupan, et al., *Phys. Rev. Lett.* **82**, 2544 (1999); J. M. Tao, J. P. Perdew, V. N. Staroverov, et al., *Phys. Rev. Lett.* **91**, 146401 (2003).
- 26 T. Leininger, H. Stoll, H.-J. Werner, et al., *Chem. Phys. Lett.* **275**, 151 (1997).
- 27 H. Iikura, T. Tsuneda, T. Yanai, et al., *J. Chem. Phys.* **115**, 3540 (2001).
- 28 R. Baer and D. Neuhauser, *Phys. Rev. Lett.* **94**, 043002 (2005).
- 29 D. J. Tozer, R. D. Amos, N. C. Handy, et al., *Mol. Phys.* **97**, 859 (1999); J. Fabian, *Th. Chem. Acc.* **106**, 199 (2001).
- 30 M. S. Liao, Y. Lu, V. D. Parker, et al., *J. Phys. Chem. A* **107**, 8939 (2003).
- 31 M. S. Liao, Y. Lu, and S. Scheiner, *J. Comp. Chem.* **24**, 623 (2003).
- 32 A. Dreuw and M. Head-Gordon, *J. Am. Chem. Soc.* **126**, 4007 (2004).
- 33 D. J. Tozer, *J. Chem. Phys.* **119**, 12697 (2003).
- 34 S. Kuemmel and L. Kronik, *Rev. Mod. Phys.* **80**, 3 (2008).
- 35 A. J. Cohen, P. Mori-Sanchez, and W. T. Yang, *Science* **321**, 792 (2008); J. P. Perdew, *Adv. Chem. Phys.* **21**, 113 (1990).
- 36 T. Yanai, D. P. Tew, and N. C. Handy, *Chem. Phys. Lett.* **393**, 51 (2004).
- 37 M. J. G. Peach, T. Helgaker, P. Salek, et al., *Phys. Chem. Chem. Phys.* **8**, 558 (2005); O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.* **125**, 234109 (2006); Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A* **110**, 13126 (2006); J. D. Chai and M. Head-Gordon, *J. Chem. Phys.* **128**, 084106 (2008).
- 38 E. Livshits and R. Baer, *Phys. Chem. Chem. Phys.* **9**, 2932 (2007).
- 39 A. Dreuw, J. L. Weisman, and M. Head-Gordon, *J. Chem. Phys.* **119**, 2943 (2003).
- 40 Y. Tawada, T. Tsuneda, S. Yanagisawa, et al., *J. Chem. Phys.* **120**, 8425 (2004).
- 41 A. Savin, in *Recent Advances in Density Functional Methods Part I*, edited by D. P. Chong (World Scientific, Singapore, 1995), p. 129.



- 42 R. Baer, E. Livshits, and D. Neuhauser, *Chem. Phys.* **329**, 266 (2006).
- 43 R. Merkle, A. Savin, and H. Preuss, *J. Chem. Phys.* **97**, 9216 (1992); T. Bally and G. N. Sastry, *J. Phys. Chem. A* **101**, 7923 (1997); M. Lundberg and P. E. M. Siegbahn, *J. Chem. Phys.* **122** (2005); J. P. Perdew, A. Ruzsinszky, G. I. Csonka, et al., *Phys. Rev. A* **76**, 040501 (2007); O. A. Vydrov, G. E. Scuseria, and J. P. Perdew, *J. Chem. Phys.* **126**, 154109 (2007); A. Ruzsinszky, J. P. Perdew, G. I. Csonka, et al., *J. Chem. Phys.* **126**, 104102 (2007); J. Grafenstein, E. Kraka, and D. Cremer, *Phys. Chem. Chem. Phys.* **6**, 1096 (2004).
- 44 J. P. Perdew, R. G. Parr, M. Levy, et al., *Phys. Rev. Lett.* **49**, 1691 (1982); C.-O. Almbladh and U. von-Barth, *Phys. Rev. B* **31**, 3231 (1985).
- 45 T. Stein, R. Baer, and L. Kronik, submitted (2008).
- 46 I. Hanazaki, *J. Phys. Chem.* **76**, 1982 (1972).
- 47 Y. Shao, L. F. Molnar, Y. Jung, et al., *Phys. Chem. Chem. Phys.* **8**, 3172 (2006).
- 48 B. Uno, N. Okumura, and K. Seto, *J. Phys. Chem. A* **104**, 3064 (2000).
- 49 K. E. Wise and R. A. Wheeler, *J. Phys. Chem. A* **103**, 8279 (1999).
- 50 R. S. Mulliken, *J. Am. Chem. Soc.* **72**, 600 (1950).
- 51 R. Baer and D. Neuhauser, *J. Chem. Phys.* **121**, 9803 (2004); D. Neuhauser and R. Baer, *J. Chem. Phys.* **123**, 204105 (2005).