ELEC 542: Nanoscale Modelling

Daniel Wong
University of British Columbia
Instructor: Dr. Alireza Nojeh

Winter 2022

Disclaimer

These notes may be freely used by anyone who comes across them. If any errors are found within these notes, please email me at temp@temp.com. If you are the previous instructor of this course (Dr. Alireza Nojeh) and have concerns about keeping these notes on my website, please contact me at the email provided.

Other course notes may be found on my website at `www.temp.com/notes`.

Contents

1   Introduction

Note the challenge of simulating systems at the nanoscale: too many particles to keep track of but not enough particles to treat statistically and as a continuum.

In the context of this course, the equation of motion for electrons is the Schrödinger equation and the nucleus will be considered as point particles.

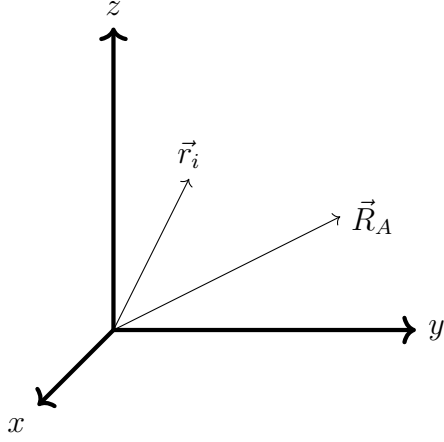The map of the course content is as follows:
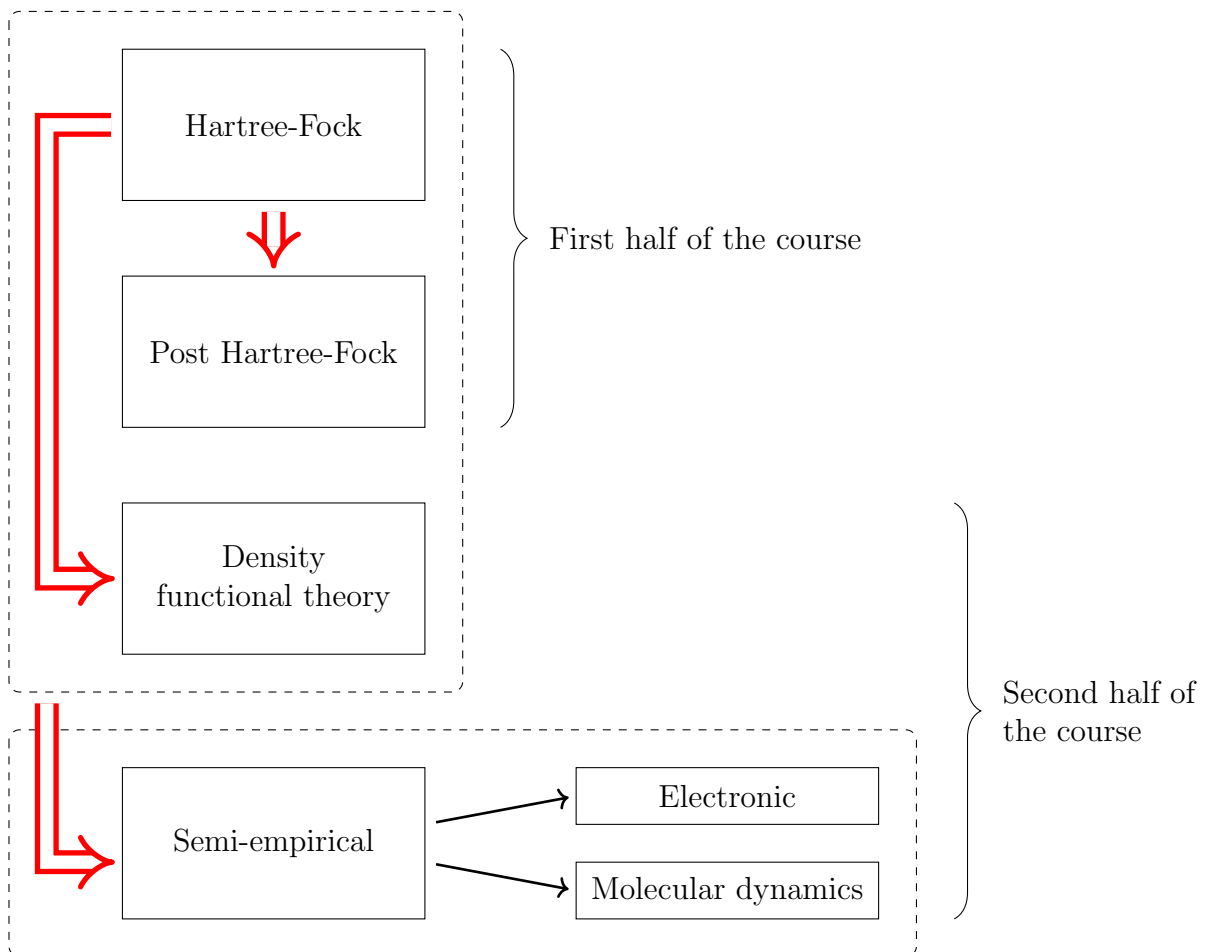
Let us consider the Schrödinger equation for a system.

$$\mathbf{H}\Psi = E\Psi \quad \text{(time-independent)} \tag{1.1}$$

Hamiltonian — wavefunction — energy eigenvalue

The wavefunction is a function of all the electron and nuclei coordinates

$$\Psi = \Psi\left(\vec{r}_1, \ldots, \vec{r}_N, \vec{R}_1, \ldots, \vec{R}_M\right) \quad \text{(many-body wavefunction)} \qquad (1.2)$$

$i = 1, \ldots, N$ (number of e$^-$)

$A = 1, \ldots, M$ (number of nuclei)

The Hamiltonian of the system is as follows

$$\mathbf{H} = \underbrace{\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2}_{\text{e}^- \text{ KE}} + \underbrace{\sum_{A=1}^{M} -\frac{1}{2M_A}\nabla_A^2}_{\text{nuclei KE}} + \underbrace{\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}}_{\text{e}^- \text{-e}^- \text{ repulsion}}$$

$$+ \underbrace{\sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}}_{\text{nucleus-nucleus repulsion}} + \underbrace{\sum_{i=1}^{N}\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}_{\text{e}^- \text{-nucleus repulsion}} \qquad (1.3)$$

This is the (non-relativistic) theory of everything. Note that this is written in atomic coordinates (hence the mass and charge of electrons are 1). Note that spin is also missing from the equation above.

In condensed matter physics, one gets accustomed to imagining quasi-particles. These quasi-particles are models of a higher-abstraction that helps give insight into the behaviour of the system. Rather than appearing to neglect this quasi-particle picture in the Hamiltonian above, the picture comes out from the Hamiltonian equation. At the start of this course, we will want to solve this Hamiltonian without jumping straight away to the usage of quasi-particles or using approximations.

Let us also make a note about states. Consider a carbon atom.
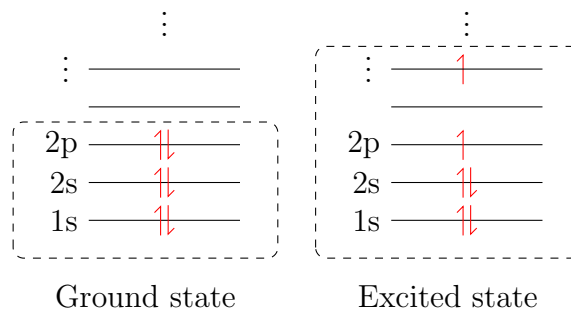
Note that the entire population of electrons in this configuration is the ground state. In a one-electron picture, one may grow accustomed to thinking of the 1s orbital as the ground state, which we will want to move away from.

2   Hartree-Fock

Let us try to decouple this problem into two problems: an electronic portion and a nuclei portion.

2.1   Born-Oppenheimer Approximation

Note that the mass of the nucleus is many orders of magnitude larger than the mass of the electron. One should expect that the nuclei moves much slower than the electrons.

One can make the Born-Oppenheimer approximation. At a given moment in time, we assume that the nuclei are fixed and that the electrons are moving in the background of these fixed nuclei.

Thus, the variables $\vec{R}_1, \ldots, \vec{R}_M$ now represent fixed positions. The second term of the Hamiltonian now drops out and the fourth term just becomes a fixed constant. We now treat the many-body wavefunction as

$$\Psi = \Psi_{\text{elec}}(\underbrace{\vec{r}_1, \ldots, \vec{r}_N}_{\substack{\text{explicit} \\ \text{variables}}}, \underbrace{\vec{R}_1, \ldots, \vec{R}_M}_{\text{parameters}})\Psi_{\text{nucl}}(\underbrace{\vec{R}_1, \ldots, \vec{R}_M}_{\substack{\text{explicit} \\ \text{variables}}}) \tag{2.1}$$

The methodology for solving the motion of the system thus starts with solving the many-body electronic problem in the background of fixed nuclei. From this solution, one can calculate the forces acting on the nuclei from the electronic cloud and other nuclei and evolve the nuclei by one time-step. One then again makes the Born-Oppenheimer approximation and recalculates the electronic solution in the new background of fixed nuclei. These steps are then iterated to find the motion of the system in a "quasi-static" manner.

Note for the nuclei portion, one is not looking at the effects of fixed electrons in space acting on the nuclei. Rather one is looking at the average effects by the electron cloud on the nuclei. Thus, Ψ_{nucl} does not parametrically depend on the electron coordinates.

Thus, in more detail

$$
\begin{aligned}
\mathbf{H}\Psi &= \left(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 \Psi_{\text{elec}}\right)\Psi_{\text{nucl}} + \left(\sum_{A=1}^{M} -\frac{1}{2M_A}\nabla_A^2(\Psi_{\text{elec}}\Psi_{\text{nucl}})\right) \\
&\quad + \left(\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}\right)\Psi_{\text{elec}}\Psi_{\text{nucl}} + \left(\sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}\right)\Psi_{\text{elec}}\Psi_{\text{nucl}} \\
&\quad + \left(\sum_{i=1}^{N}\sum_{A=1}^{M}-\frac{Z_A}{r_{iA}}\right)\Psi_{\text{elec}}\Psi_{\text{nucl}} \\
&= \left[\left(\sum_{i=1}^{N}-\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{i=1}^{N}\sum_{A=1}^{M}-\frac{Z_A}{r_{iA}}\right)\Psi_{\text{elec}}\right]\Psi_{\text{nucl}} \\
&\quad + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}\Psi_{\text{elec}}\Psi_{\text{nucl}} \\
&\quad + \sum_{A=1}^{M}-\frac{1}{2M_A}\left[\left(\nabla_A^2\Psi_{\text{elec}}\right)\Psi_{\text{nucl}} + \left(\nabla_A^2\Psi_{\text{nucl}}\right)\Psi_{\text{elec}} + 2(\nabla_A\Psi_{\text{elec}})(\nabla_A\Psi_{\text{nucl}})\right]
\end{aligned}
$$

neglect (pointing to $\left(\nabla_A^2\Psi_{\text{elec}}\right)\Psi_{\text{nucl}}$)

neglect (pointing to $2(\nabla_A\Psi_{\text{elec}})(\nabla_A\Psi_{\text{nucl}})$)

$$(2.2)$$

Compare the terms $\nabla_A\Psi_{\text{elec}}$ to $\nabla_A\Psi_{\text{nucl}}$.

small change

Ψ_{elec}

large change

Ψ_{nucl}

small nucleus
movement

small nucleus
movement
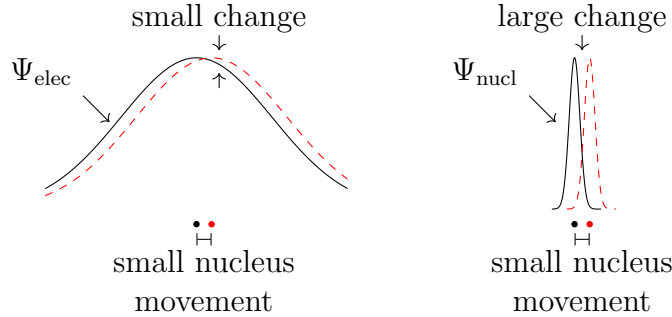
The electron cloud around a nucleus does not change significantly for small movements of the nucleus. In comparison, the nucleus wavefunction changes significantly with movement of the nucleus. Thus, we expect the first and third term on the third line of the equation above to be much smaller than

the second term. Therefore

$$\mathbf{H}\Psi \approx \left[\left(\overbrace{\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{i=1}^{N}\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}^{E_{\text{elec}}} \right)\Psi_{\text{elec}} \right]\Psi_{\text{nucl}}$$

$$+ \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}}\Psi_{\text{elec}}\Psi_{\text{nucl}} + \sum_{A=1}^{M} -\frac{1}{2M_A}\left(\nabla_A^2 \Psi_{\text{nucl}}\right)\Psi_{\text{elec}} \tag{2.3}$$

$$= E\Psi_{\text{elec}}\Psi_{\text{nucl}}]]$$

$$\implies \left[\sum_{A=1}^{M} -\frac{1}{2M_A}\nabla_A^2 + \underbrace{\sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} + E_{\text{elec}}}_{\substack{\text{potential energy landscape} \\ \text{for nuclear motion}}} \right]\Psi = E\Psi \tag{2.4}$$

To calculate the force on the nuclei, one would want to find the derivative of this nuclear energy with respect to these nuclei coordinates. This can be done easily using the Hellman-Feynman theorem. Thus, one can do dynamics of the nuclei.

2.2 Pauli Exclusion Principle and Hartree Products

Back to the electronic problem:

$$\mathbf{H}_{\text{elec}}\Psi_{\text{elec}} = E_{\text{elec}}\Psi_{\text{elec}} \quad \text{(Schrödinger equation)} \tag{2.5}$$

where

$$\mathbf{H}_{\text{elec}} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{2.6}$$

One half of the story is still missing: the Pauli exclusion principle. One must now talk about electron spin.

Consider the one-electron problem.

$$\mathcal{X}(\overbrace{\vec{r}_1, \omega_1}^{\vec{x}_1}) = \begin{cases} \Psi(\vec{r}_1)\alpha(\omega_1) \\ \quad\text{or} \\ \underbrace{\Psi(\vec{r}_1)}_{\substack{\text{spatial} \\ \text{orbitals}}}\underbrace{\beta(\omega_1)}_{\substack{\text{spin} \\ \text{orbitals}}} \end{cases} \quad \text{(one-electron wavefunction)} \tag{2.7}$$

where

$$\int \alpha^*(\omega_1)\beta(\omega_1)\,\mathrm{d}\omega_1 = 0 \tag{2.8}$$

and

$$\int \alpha^*(\omega_1)\alpha(\omega_1)\,\mathrm{d}\omega_1 = \int \beta^*(\omega_1)\beta(\omega_1)\,\mathrm{d}\omega_1 = 1 \tag{2.9}$$

Generalizing, the many-electron wavefunction is

$$\underbrace{\mathcal{X}(\vec{x}_1,\ldots,\vec{x}_N)}_{4N\text{ variables}} \tag{2.10}$$

By Pauli's principle, the wavefunction is anti-symmetric by exchange.

$$\mathcal{X}(\vec{x}_1,\cdots,\vec{x}_i,\cdots,\vec{x}_j,\cdots,\vec{x}_N) = -\mathcal{X}(\vec{x}_1,\cdots,\vec{x}_j,\cdots,\vec{x}_i,\cdots,\vec{x}_N) \tag{2.11}$$

swapped

One must find the solution of Schrödinger's equation in the form of an anti-symmetric wavefunction. Let us try a separation of variables (which may not always be justified) for each electron. The first two terms of the electronic Hamiltonian lends itself easily to separating each electron. The third term in $\mathbf{H}_{\text{elec}}$ couples the electrons together. As a first step, let us neglect this troublesome third term.

$$\mathbf{H}_{\text{elec}} \approx \sum_{i=1}^{N}\bigg(\underbrace{-\frac{1}{2}\nabla_i^2 + \sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}_{h(i)\,:\,1\text{-e}^-\text{ Hamiltonian}}\bigg) \tag{2.12}$$

$$= h(1) + h(2) + \ldots + h(N) \qquad \text{(Hartree approximation)}$$

Let us now separate the wavefunction $\mathcal{X}(\vec{x}_1,\cdots,\vec{x}_N)$ as

$$\mathcal{X}(\vec{x}_1,\cdots,\vec{x}_N) = \mathcal{X}_1(\vec{x}_1)\mathcal{X}_2(\vec{x}_2)\ldots\mathcal{X}_N(\vec{x}_N) \quad \text{(Hartree product)} \tag{2.13}$$

Thus, the Schrödinger equation becomes

$$(h(1) + \ldots + h(N))(\mathcal{X}_1(\vec{x}_1)\cdots\mathcal{X}_N(\vec{x}_N)) = E(\mathcal{X}_1(\vec{x}_1)\cdots\mathcal{X}_N(\vec{x}_N)) \tag{2.14}$$

with $E = E_1 + \ldots + E_N$.

The problem is now reduced to

$$\begin{cases} h(1)\mathcal{X}_1(\vec{x}_1) = E_1\mathcal{X}_1(\vec{x}_1) & (2.15) \\ \qquad\qquad \vdots & (2.16) \\ h(N)\mathcal{X}_N(\vec{x}_N) = E_N\mathcal{X}_N(\vec{x}_N) & (2.17) \end{cases}$$

Note that these are all the same equation. Thus

$$h(1)\mathcal{X}_i(\vec{x}_1) = E_i\mathcal{X}_i(\vec{x}_1), \quad i = 1, \cdots, \infty \qquad (2.18)$$

Each Hartree product is a solution to the Schrödinger equation with energy eigenvalue E which is a sum of all E_i. Note that are infinitely many solutions.

2.3   Slater Determinants

Note that the solution of the full electron Hamiltonian is a function of $3N$ variables (where N is the number of electrons). This is an incredibly difficult problem to solve numerically.

Consider the Hartree product.

$$\mathcal{X}(\vec{x}_1, \cdot, \vec{x}_N) = \mathcal{X}_1(\vec{x}_1)\dots\mathcal{X}_N(\vec{x}_N) \qquad (2.19)$$

Note that this form of the many-electron wavefunction is not anti-symmetric under exchange. Let us introduce the Slater determinant to solve this issue. Consider a system with two electrons. The Slater determinant is

$$\mathcal{X}(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}}[\mathcal{X}_1(\vec{x}_1)\mathcal{X}_2(\vec{x}_2) - \mathcal{X}_1(\vec{x}_2)\mathcal{X}(\vec{x}_1)] \qquad (2.20)$$

One can see that $\mathcal{X}(\vec{x}_2, \vec{x}_1) = -\mathcal{X}(\vec{x}_1, \vec{x}_2)$. Moreover, $\mathcal{X}(\vec{x}_1, \vec{x}_1) = \mathcal{X}(\vec{x}_2, \vec{x}_2) = 0$, which matches the description of the Pauli exclusion principle.

In determinant form, the Slater determinant is

$$\mathcal{X}(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}}\begin{vmatrix} \mathcal{X}_1(\vec{x}_1) & \mathcal{X}_2(\vec{x}_1) \\ \mathcal{X}_1(\vec{x}_2) & \mathcal{X}_2(\vec{x}_2) \end{vmatrix} \qquad (2.21)$$

Generalizing to N electrons

$$\mathcal{X}(\vec{x}_1, \cdots, \vec{x}_N) = \underbrace{\frac{1}{\sqrt{N!}}}_{\text{for normalization}}\begin{vmatrix} \mathcal{X}_1(\vec{x}_1) & \dots & \mathcal{X}_N(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1(\vec{x}_N) & \dots & \mathcal{X}_N(\vec{x}_N) \end{vmatrix} \qquad (2.22)$$

One can ultimately show that this is anti-symmetric. Its normalization states

$$\int \mathcal{X}^*(\vec{x}_1, \cdots, \vec{x}_N)\mathcal{X}(\vec{x}_1, \cdots, \vec{x}_N)\,d\vec{x}_1 \ldots d\vec{x}_N = 1 \tag{2.23}$$

Ex:　　Consider a two-electron Slater determinant. What occurs if the two electrons have opposite spin?

$$\mathcal{X}(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}}[\mathcal{X}_1(\vec{x}_1)\mathcal{X}_2(\vec{x}_2) - \mathcal{X}_2(\vec{x}_1)\mathcal{X}_1(\vec{x}_2)]$$
$$\mathcal{X}_1(\vec{x}_1) = \Psi_1(\vec{r}_1)\alpha(\omega_1) \tag{2.24}$$
$$\mathcal{X}_2(\vec{x}_2) = \Psi_2(\vec{r}_2)\beta(\omega_2)$$

Its probability density as a function of $\vec{r}_1$ and $\vec{r}_2$ is

$$\begin{aligned}
P(\vec{r}_1, \vec{r}_2) &= \int \mathcal{X}^*(\vec{x}_1, \vec{x}_2)\mathcal{X}(\vec{x}_1, \vec{x}_2)\,d\omega_1\,d\omega_2 \\
&= \frac{1}{2}\int [\Psi_1(\vec{r}_1)\alpha^*(\omega_1)\Psi_2^*(\vec{r}_2)\beta^*(\omega_2) - \Psi_1^*(\vec{r}_2)\alpha^*(\omega_2)\Psi_2^*(\vec{r}_1)\beta^*(\omega_1)] \\
&\quad [\Psi_1(\vec{r}_1)\alpha(\omega_1)\Psi_2(\vec{r}_2)\beta(\omega_2) - \Psi_1(\vec{r}_2)\alpha(\omega_2)\Psi_2(\vec{r}_1)\beta(\omega_1)]\,d\omega_1\,d\omega_2 \\
&= \frac{1}{2}\int \Big[|\Psi_1(\vec{r}_1)|^2|\Psi_2(\vec{r}_2)|^2|\alpha(\omega_1)|^2|\beta(\omega_2)|^2 \\
&\quad + |\Psi_1(\vec{r}_2)|^2|\Psi_2(\vec{r}_1)|^2|\alpha(\omega_2)|^2|\beta(\omega_1)|^2 \\
&\quad - \underline{\Psi_1^*(\vec{r}_1)\Psi_1(\vec{r}_2)\Psi_2^*(\vec{r}_2)\Psi_2(\vec{r}_1)\alpha^*(\omega_1)\beta(\omega_1)\beta^*(\omega_2)\alpha(\omega_2)} \overset{0}{\longrightarrow} \\
&\quad - \underline{\Psi_1^*(\vec{r}_2)\Psi_1(\vec{r}_1)\Psi_2^*(\vec{r}_1)\Psi_2(\vec{r}_2)\alpha^*(\omega_2)\beta(\omega_2)\beta^*(\omega_1)\alpha(\omega_1)} \overset{0}{\longrightarrow} \Big]\,d\omega_1\,d\omega_2 \\
&= \frac{1}{2}\Big[\underbrace{|\Psi_1(\vec{r}_1)|^2}_{P_1(\vec{r}_1)}\underbrace{|\Psi_2(\vec{r}_2)|^2}_{P_2(\vec{r}_2)} + |\Psi_1(\vec{r}_2)|^2|\Psi_2(\vec{r}_1)|^2 \Big]
\end{aligned}$$

$$\tag{2.25}$$

This can be interpreted as the average of finding electron 1 at $\vec{r}_1$ and electron 2 at $\vec{r}_2$ and of finding the same electrons in swapped positions. This form comes from the indistinguishability of electrons.

The probability overall came out to be the product of these two individual probabilities. Thus, the motion of the two electrons of opposite spin is uncorrelated.

What occurs if the two electrons have the same spin? Following the same algebra, but with replacing β with α in $\mathcal{X}_2$

$$P(\vec{r}_1, \vec{r}_2) = \frac{1}{2} \left[|\Psi_1(\vec{r}_1)|^2 |\Psi_2(\vec{r}_2)|^2 + |\Psi_1(\vec{r}_2)|^2 |\Psi_2(\vec{r}_1)|^2 \right]$$
$$- \frac{1}{2} [\Psi_1^*(\vec{r}_1)\Psi_1(\vec{r}_2)\Psi_2^*(\vec{r}_2)\Psi_2(\vec{r}_1) + \Psi_1^*(\vec{r}_2)\Psi_1(\vec{r}_1)\Psi_2^*(\vec{r}_1)\Psi_2(\vec{r}_2)]$$

$$(2.26)$$

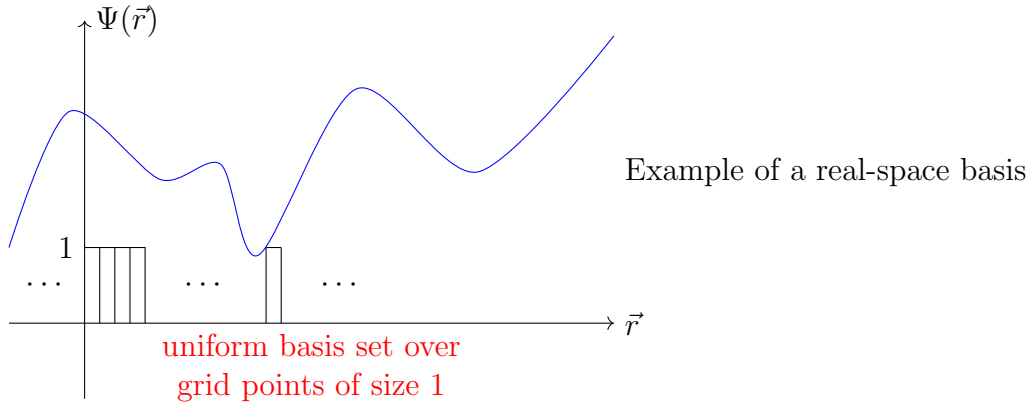Now the motion of the two electrons with the same spin is correlated.

2.4   Representation of Functions

In general, we will expand our functions over a complete basis set.

$$\Psi(\vec{r}) = \sum_{\alpha=1}^{\infty} c_\alpha \phi_\alpha(\vec{r}) \tag{2.27}$$

where $\{\phi_\alpha(\vec{r})\}$ is a complete basis set.

Consider a spatial basis set as follows:

Example of a real-space basis

uniform basis set over
grid points of size 1

One can then express $\Psi(\vec{r})$ as a vector.

$$\Psi(\vec{r}) \equiv \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_\alpha \\ \vdots \end{bmatrix} \tag{2.28}$$

$$\equiv |\Psi\rangle \quad \text{(Bra-ket notation)}$$

Note that these vectors will be infinitely long. In practice, these vectors will be truncated at some point.

The conjugate transpose is

$$\Psi^*(\vec{r}) \equiv \begin{bmatrix} c_1 & c_2 & \cdots & c_\alpha & \cdots \end{bmatrix} \equiv \langle\Psi| \tag{2.29}$$

Operators will be represented as matrices. Given that they correspond with real physical quantities, these matrices will be square Hermitian matrices ($\mathbf{O} = \mathbf{O}^\dagger$).

2.5 Hartree-Fock Method

The idea is to find the best approximation to the ground state of the system in the form of a single Slater determinant.

$$\mathbf{H}\mathcal{X}_0 = E\mathcal{X}_0 \tag{2.30}$$

where $\mathcal{X}_0 \equiv \mathcal{X}_0(\vec{x}_1, \cdots, \vec{x}_N)$ is the ground state and is given by the equation

$$\mathcal{X}_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \mathcal{X}_1(\vec{x}_1) & \dots & \mathcal{X}_N(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1(\vec{x}_N) & \dots & \mathcal{X}_N(\vec{x}_N) \end{vmatrix} \tag{2.31}$$

Note that the Slater determinant does not need to be constructed using single electron wavefunctions 1 to N. They simply need to be made from <u>any</u> N functions. In our case, we will assume all the $\mathcal{X}_i$ have been ordered such that we have a good starting point for our ground state. An "excited" Slater determinant may be constructed using one-electron wavefunctions with higher energy.

$$\mathcal{X} = \frac{1}{\sqrt{N!}} \underbrace{\begin{vmatrix} \mathcal{X}_i(\vec{x}_1) & \mathcal{X}_j(\vec{x}_1) & \dots & \mathcal{X}_k(\vec{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}_i(\vec{x}_N) & \mathcal{X}_j(\vec{x}_N) & \dots & \mathcal{X}_k(\vec{x}_N) \end{vmatrix}}_{\text{determinant of } N \times N \text{ matrix}} \tag{2.32}$$

Hartree-Fock will transform this problem into

$$F\,\mathcal{X}_i(\vec{x}_1) = \varepsilon_i \mathcal{X}_i(\vec{x}_1) \tag{2.33}$$

Fock operator

This Fock operator itself will depend on $\{\mathcal{X}_i(\vec{x}_1)\}$ that needs to be solved in a self-consistent manner. It will still contain a kinetic term for the electrons and electron-nuclei attraction, but it will transform the electron-electron repulsion into some kind of average electron-electron interaction.

2.6 Variational Principle

The variational principle will serve as a guideline for finding the ground-state Slater determinant. The principle states that the expectation value of the Hamiltonian for the system in a trial wave function $\mathcal{X}$ that is normalized will be an upper bound to the ground state energy of the N electron system. This transforms our problem into a minimization problem.

$$\langle \mathcal{X} | \mathbf{H} | \mathcal{X} \rangle >= E_0 \tag{2.34}$$

Proof: Assume that $\{\phi_\alpha(\vec{x}_1, \cdots, \vec{x}_N)\}$ is the actual exact solution of the N-electron problem.

$$\mathbf{H} | \phi_\alpha \rangle = E_\alpha | \phi_\alpha \rangle, \quad \alpha = 0, 1, \cdots, N, \cdots \tag{2.35}$$

such that $E_0 < E_1 < \cdots < E_N < \cdots$.

Since $\mathbf{H}$ is Hermitian, its solutions form a complete orthonormal set. Thus

$$| \mathcal{X} \rangle = \sum_{\alpha=0}^{\infty} c_\alpha | \phi_\alpha \rangle \tag{2.36}$$

Multiplying on the left by $\langle \phi_\beta |$

$$\langle \phi_\beta | \mathcal{X} \rangle = \sum_{\alpha=0}^{\infty} c_\alpha \langle \phi_\beta | \phi_\alpha \rangle = \sum_{\alpha=0}^{\infty} c_\alpha \delta_{\alpha\beta} = c_\beta \tag{2.37}$$

Thus

$$| \mathcal{X} \rangle = \sum_{\alpha=0}^{\infty} \langle \phi_\alpha | \mathcal{X} \rangle | \phi_\alpha \rangle = \underbrace{\sum_{\alpha=0}^{\infty} | \phi_\alpha \rangle\langle \phi_\alpha |}_{\mathbb{I}} | \mathcal{X} \rangle \tag{2.38}$$

Calculating the expectation value

$$\langle \mathcal{X} | \mathbf{H} | \mathcal{X} \rangle = \left(\sum_{\alpha=0}^{\infty} \langle \mathcal{X} | \phi_\alpha \rangle \langle \phi_\alpha | \right) \mathbf{H} \left(\sum_{\beta=0}^{\infty} | \phi_\beta \rangle \langle \phi_\beta | \mathcal{X} \rangle \right)$$

$$= \sum_{\alpha=0}^{\infty} \sum_{\beta=0}^{\infty} \langle \mathcal{X} | \phi_\alpha \rangle \langle \phi_\alpha | \mathbf{H} | \phi_\beta \rangle \langle \phi_\beta | \mathcal{X} \rangle$$

$$= \sum_{\alpha=0}^{\infty} \sum_{\beta=0}^{\infty} E_\beta \langle \mathcal{X} | \phi_\alpha \rangle \underbrace{\langle \phi_\alpha | \phi_\beta \rangle}_{\delta_{\alpha\beta}} \langle \phi_\beta | \mathcal{X} \rangle \qquad (2.39)$$

$$= \sum_{\alpha=0}^{\infty} E_\alpha \langle \mathcal{X} | \phi_\alpha \rangle \langle \phi_\alpha | \mathcal{X} \rangle$$

$$= \sum_{\alpha=0}^{\infty} E_\alpha |\langle \mathcal{X} | \phi_\alpha \rangle|^2$$

$$\geq E_0 \sum_{\alpha=0}^{\infty} |\langle \mathcal{X} | \phi_\alpha \rangle|^2$$

Moreover, following a similar procedure

$$\langle \mathcal{X} | \mathcal{X} \rangle = \sum_{\alpha=0}^{\infty} |\langle \mathcal{X} | \phi_\alpha \rangle|^2 = 1 \qquad (2.40)$$

Thus

$$\langle \mathcal{X} | \mathbf{H} | \mathcal{X} \rangle \geq E_0 \qquad (2.41)$$

2.7   Expectation Value of Hamiltonian for Single Slater Determinant

Let us now work out $\langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle$ for $|\mathcal{X}_0\rangle$ being a single Slater determinant. Recall that

$$\mathbf{H} = \underbrace{\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}_{\mathbf{O}_1 = \sum_{i=1}^{N} h(1)} + \underbrace{\sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}}}_{\mathbf{O}_2} \qquad (2.42)$$

Thus

$$\langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle = \langle \mathcal{X}_0 | \mathbf{O}_1 | \mathcal{X}_0 \rangle + \langle \mathcal{X}_0 | \mathbf{O}_2 | \mathcal{X}_0 \rangle \tag{2.43}$$

The first term is

$$
\begin{aligned}
\langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle &= \langle \mathcal{X}_0 | [h(1) + h(2) + \ldots + h(N)] | \mathcal{X}_0 \rangle \\
&= N \langle \mathcal{X}_0 | h(1) | \mathcal{X}_0 \rangle \quad \text{(since } 1, \cdots, N \text{ are just dummy variables)} \\
&= \frac{N}{\sqrt{N!}} \int \left[\sum_{n-1}^{N!} (-1)^{n+1} \mathrm{P}_n \{ \mathcal{X}_1^*(\vec{x}_1) \mathcal{X}_2^*(\vec{x}_2) \ldots \mathcal{X}_N^*(\vec{x}_N) \} h(1) \right. \\
&\qquad \left. \frac{1}{\sqrt{N!}} \sum_{m=1}^{N!} (-1)^{m+1} \mathrm{P}_m \{ \mathcal{X}_1(\vec{x}_1) \mathcal{X}_2(\vec{x}_2) \ldots \mathcal{X}_N(\vec{x}_N) \} \right] \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N \\
&= \frac{1}{(N-1)!} \int \left[\sum_{n=1}^{N!} \sum_{m=1}^{N!} (-1)^{m+n+2} \mathrm{P}_n \{ \mathcal{X}_1^*(\vec{x}_1) \mathcal{X}_2^*(\vec{x}_2) \ldots \mathcal{X}_N^*(\vec{x}_N) \} h(1) \right. \\
&\qquad \left. \mathrm{P}_m \{ \mathcal{X}_1(\vec{x}_1) \mathcal{X}_2(\vec{x}_2) \ldots \mathcal{X}_N(\vec{x}_N) \} \right] \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N
\end{aligned}
\tag{2.44}
$$

where P_n denotes the n-th permutation. By orthonormality of $\{ \mathcal{X}_i(\vec{x}_i) \}$, in the above terms, $\vec{x}_2, \cdots, \vec{x}_N$ must be in the same one-electron function in both permutations for that term to be non-zero. Thus, $\vec{x}_1$ will also have one choice left. Moreover, P_n and P_m must be the same permutation.

$$
\begin{aligned}
\langle \mathcal{X}_0 | \mathbf{O}_1 | \mathcal{X}_0 \rangle &= \frac{1}{(N-1)!} \sum_{n=1}^{N} \int \mathcal{X}_n^*(\vec{x}_1) h(1) \mathcal{X}_n(\vec{x}_1) \, \mathrm{d}\vec{x}_1 \, (N-1)! \\
&= \sum_{n=1}^{N} \int \mathcal{X}_n^*(\vec{x}_1) h(1) \mathcal{X}_n(\vec{x}_1) \, \mathrm{d}\vec{x}_1 \\
&= \sum_{n=1}^{N} \langle \mathcal{X}_n | h | \mathcal{X}_n \rangle \quad \text{or} \quad \underbrace{\sum_{n=1}^{N} [n|h|n]}_{\text{chemistry notation}}
\end{aligned}
\tag{2.45}
$$

The second term is

$$
\begin{aligned}
\langle \mathcal{X}_0 | \mathbf{O}_2 | \mathcal{X}_0 \rangle =\ & \langle \mathcal{X}_0 | \left[\frac{1}{r_{12}} + \frac{1}{r_{13}} + \ldots + \frac{1}{r_{n-1,n}} \right] | \mathcal{X}_0 \rangle \\
=\ & \frac{N!}{2!(N-2)!} \langle \mathcal{X}_0 | \frac{1}{r_{12}} | \mathcal{X}_0 \rangle \\
=\ & \frac{N!}{2!(N-2)!} \int \left[\sum_{n=1}^{N!} \sum_{m=1}^{N!} \left(\frac{1}{\sqrt{N1}} \right) \left(\frac{1}{\sqrt{N!}} \right) (-1)^{n-1} (-1)^{m+1} \right. \\
& \left. \mathrm{P}_n \{ \mathcal{X}_1^*(\vec{x}_1) \ldots \mathcal{X}_N^*(\vec{x}_N) \} \frac{1}{r_{12}} \mathrm{P}_m \{ \mathcal{X}_1(\vec{x}_1) \ldots \mathcal{X}_N(\vec{x}_N) \} \right] \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N
\end{aligned}
$$

$$(2.46)$$

For a given term to be non-zero, $\vec{x}_j$ for $j \neq 1, 2$ has to be in the same $\mathcal{X}_i$ in both permutations n and m due to orthonormality $\left(\int \mathcal{X}_k^*(\vec{x}_j) \mathcal{X}_\ell(\vec{x}_j) \mathrm{d}\vec{x}_j = \delta_{k\ell} \right)$. Thus

even since they are in the same permutation

$$
\begin{aligned}
\langle \mathcal{X}_0 | \mathbf{O}_2 | \mathcal{X}_0 \rangle =\ & \frac{1}{2(N-2)!} \sum_{n=1}^{N} \sum_{m \neq n}^{N} \left[\int (-1)^{n+1+m+1} \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m^*(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_n(\vec{x}_1) \mathcal{X}_m(\vec{x}_2) \mathrm{d}\vec{x}_1 \mathrm{d}\vec{x}_2 \right. \\
& \left. + \int (-1)^{n+1+m+1} \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m^*(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_n(\vec{x}_2) \mathcal{X}_m(\vec{x}_1) \mathrm{d}\vec{x}_1 \mathrm{d}\vec{x}_2 \right]
\end{aligned}
$$

$(N-2)! \leftarrow$ arises from the possibility of putting $\vec{x}_j$ in the remaining $\mathcal{X}_i$ functions

$$
\begin{aligned}
=\ & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left[\int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m^*(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_n(\vec{x}_1) \mathcal{X}_m(\vec{x}_2) \mathrm{d}\vec{x}_1 \mathrm{d}\vec{x}_2 \right. \\
& \left. - \int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m^*(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_n(\vec{x}_2) \mathcal{X}_m(\vec{x}_1) \mathrm{d}\vec{x}_1 \mathrm{d}\vec{x}_2 \right] \\
\equiv\ & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left[\langle nm | nm \rangle - \langle nm | mn \rangle \right] \\
=\ & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left[\int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_n(\vec{x}_1) \frac{1}{r_{12}} \mathcal{X}_m^*(\vec{x}_2) \mathcal{X}_m(\vec{x}_2) \mathrm{d}\vec{x}_1 \mathrm{d}\vec{x}_2 \right. \\
& \left. - \int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m(\vec{x}_1) \frac{1}{r_{12}} \mathcal{X}_m^*(\vec{x}_2) \mathcal{X}_n(\vec{x}_2) \mathrm{d}\vec{x}_1 \mathrm{d}\vec{x}_2 \right] \\
\equiv\ & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left(\underbrace{[nn|mm]}_{\text{Coulomb}} - \underbrace{[nm|mn]}_{\text{exchange}} \right)
\end{aligned}
$$

$$(2.47)$$

Thus

minimize this!
$$\langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle = \sum_{n=1}^{N} [n|h|n] + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left([nn|mm] - [nm|mn] \right) \qquad (2.48)$$

2.8   Minimization and Lagrange Method of Undetermined Multipliers

Note that the previous result for $\langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle$ requires the orthonormality of the trial function used to build the Slater determinant $|\mathcal{X}_0\rangle$. Thus, varying the trial functions while following the variational principle becomes a constrained minimization problem. This can be tackled using the Lagrange method of undetermined multipliers.

Let us consider an example. Imagine a function $f(x, y)$ that is to be the subject of minimization following the constraint $g(x, y) = C$.
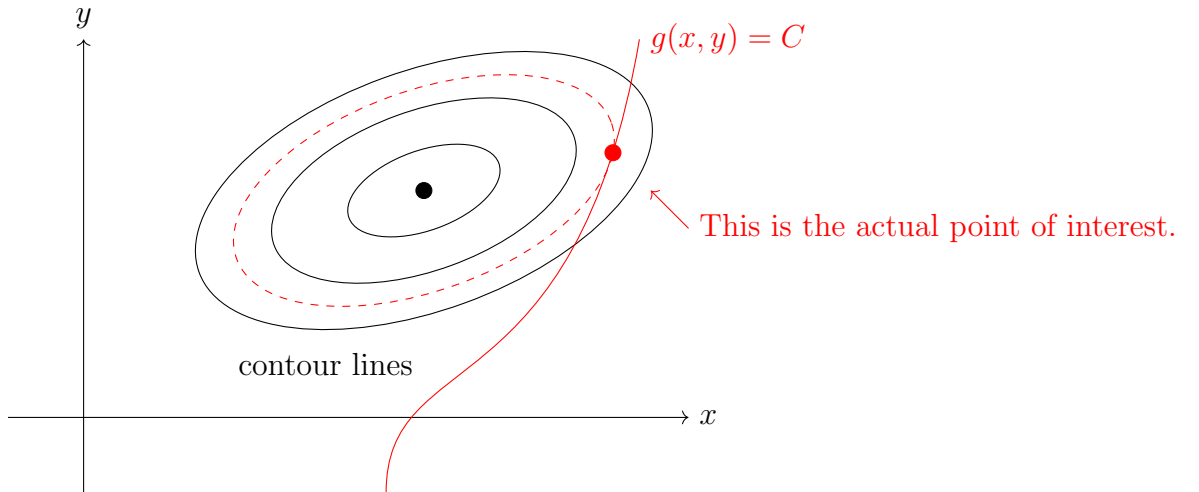
Thus

$$\begin{cases} g(x, y) - C = 0 \\ \nabla f \propto \nabla g \quad \text{(since they're tangential to the contour)} \end{cases} \qquad (2.49)$$