ELEC 542: Nanoscale Modeling and Simulations

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:
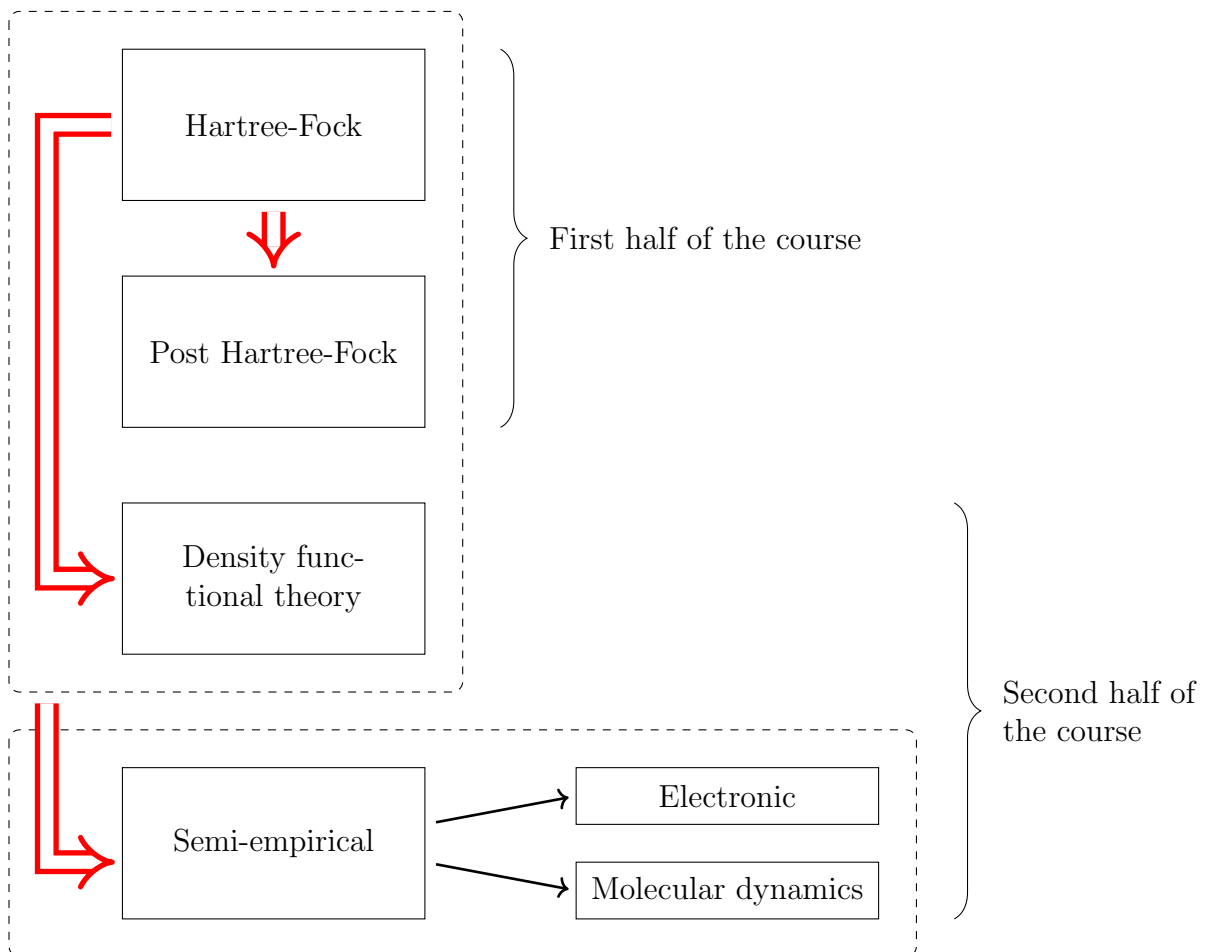
Let us consider the Schrödinger equation for a system.

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

Hamiltonian — wavefunction — energy eigenvalue

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

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

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

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

The Hamiltonian of the system is as follows

$$\mathbf{H} = \underbrace{\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2}_{\text{e}^- \text{ KE}} \;+\; \underbrace{\sum_{A=1}^{M} -\frac{1}{2M_A}\nabla_A^2}_{\text{nuclei KE}} \;+\; \underbrace{\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}}}_{\text{e}^-\text{-e}^- \text{ repulsion}}$$
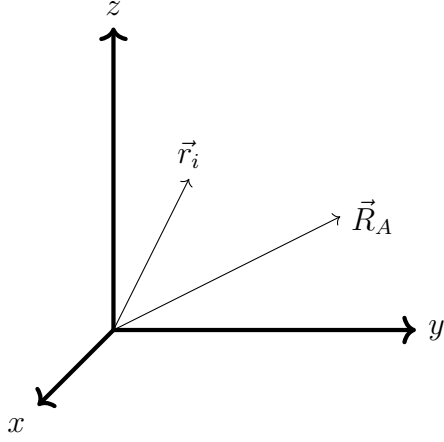$$+\; \underbrace{\sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}}_{\text{nucleus-nucleus repulsion}} \;+\; \underbrace{\sum_{i=1}^{N}\sum_{A=1}^{M}-\frac{Z_A}{r_{iA}}}_{\text{e}^-\text{-nucleus repulsion}} \qquad (1.3)$$

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.

Ground state   Excited state
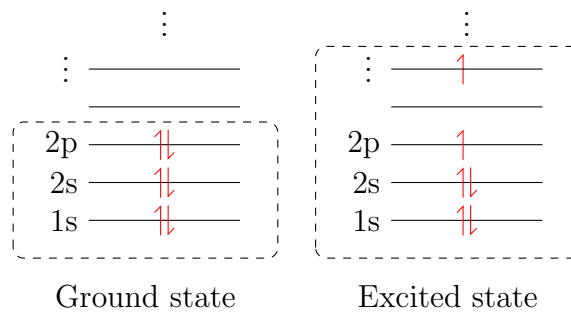
Note that the entire population of electrons in this configuration is the ground state. In a one-electron picture, one may grow accustomed to thinking of the 1s orbital as the ground state, which we will want to move away from.

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

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

neglect   neglect

$$(2.2)$$

Compare the terms $\nabla_A\Psi_{\text{elec}}$ to $\nabla_A\Psi_{\text{nucl}}$.
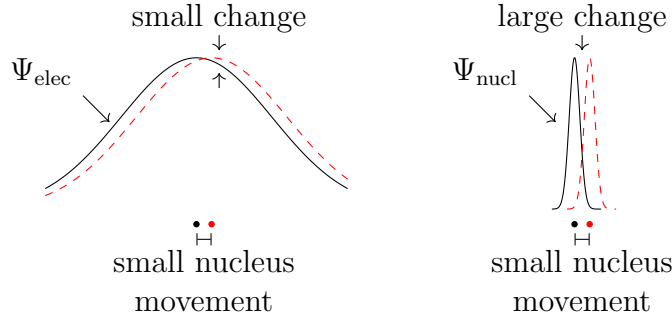
The electron cloud around a nucleus does not change significantly for small movements of the nucleus. In comparison, the nucleus wavefunction changes significantly with movement of the nucleus. Thus, we expect the first and third term on the third line of the equation above to be much smaller than

the second term. Therefore

$$
\mathbf{H}\Psi \approx \left[\left(\overbrace{\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{i=1}^{N}\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}^{E_{\text{elec}}}\right)\Psi_{\text{elec}}\right]\Psi_{\text{nucl}}
$$
$$
+ \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}\Psi_{\text{elec}}\Psi_{\text{nucl}} + \sum_{A=1}^{M} -\frac{1}{2M_A}\left(\nabla_A^2\Psi_{\text{nucl}}\right)\Psi_{\text{elec}}
$$
$$
= E\Psi_{\text{elec}}\Psi_{\text{nucl}}]] \tag{2.3}
$$

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

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.

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

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

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

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

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

The problem is now reduced to

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

Note that these are all the same equation. Thus

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

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.

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

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

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

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

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

Generalizing to N electrons

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

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

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

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

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

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

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

$$\tag{2.25}$$

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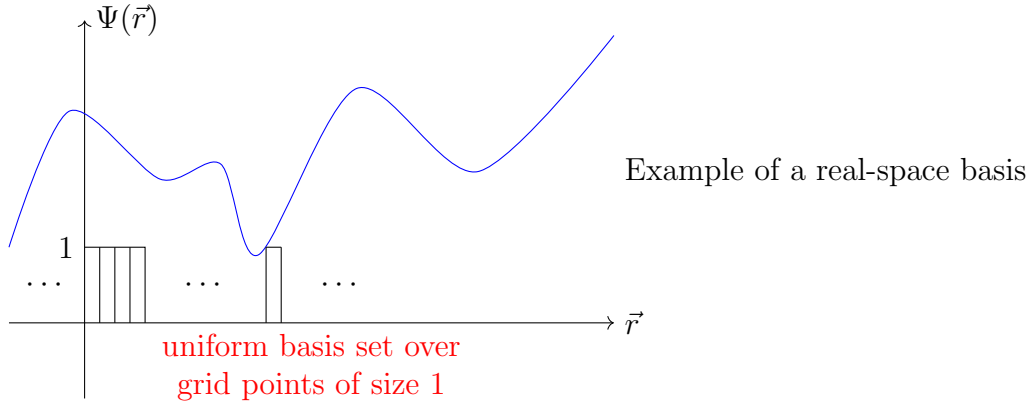
One can then express $\Psi(\vec{r})$ as a vector.

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

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

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

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

Thus

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

The first term is

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

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.

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

The second term is

$$
\begin{aligned}
\langle \mathcal{X}_0 | \mathbf{O}_2 | \mathcal{X}_0 \rangle = {} & \langle \mathcal{X}_0 | \left[\frac{1}{r_{12}} + \frac{1}{r_{13}} + \ldots + \frac{1}{r_{n-1,n}} \right] | \mathcal{X}_0 \rangle \\
= {} & \frac{N!}{2!(N-2)!} \langle \mathcal{X}_0 | \frac{1}{r_{12}} | \mathcal{X}_0 \rangle \\
= {} & \frac{N!}{2!(N-2)!} \int \left[\sum_{n=1}^{N!} \sum_{m=1}^{N!} \left(\frac{1}{\sqrt{N!}} \right) \left(\frac{1}{\sqrt{N!}} \right) (-1)^{n-1}(-1)^{m+1} \right. \\
& \left. \mathrm{P}_n \{ \mathcal{X}_1^*(\vec{x}_1) \ldots \mathcal{X}_N^*(\vec{x}_N) \} \frac{1}{r_{12}} \mathrm{P}_m \{ \mathcal{X}_1(\vec{x}_1) \ldots \mathcal{X}_N(\vec{x}_N) \} \right] \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N
\end{aligned}
\tag{2.46}
$$

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

$$
\begin{aligned}
= {} & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left[\int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m^*(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_n(\vec{x}_1) \mathcal{X}_m(\vec{x}_2) \, \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 \right. \\
& \left. - \int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m^*(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_n(\vec{x}_2) \mathcal{X}_m(\vec{x}_1) \, \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 \right] \\
\equiv {} & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left[\langle nm | nm \rangle - \langle nm | mn \rangle \right] \\
= {} & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left[\int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_n(\vec{x}_1) \frac{1}{r_{12}} \mathcal{X}_m^*(\vec{x}_2) \mathcal{X}_m(\vec{x}_2) \, \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 \right. \\
& \left. - \int \mathcal{X}_n^*(\vec{x}_1) \mathcal{X}_m(\vec{x}_1) \frac{1}{r_{12}} \mathcal{X}_m^*(\vec{x}_2) \mathcal{X}_n(\vec{x}_2) \, \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 \right] \\
\equiv {} & \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left(\underbrace{[nn|mm]}_{\text{Coulomb}} - \underbrace{[nm|mn]}_{\text{exchange}} \right)
\end{aligned}
\tag{2.47}
$$

Thus

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

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 $g(x, y) = C$.
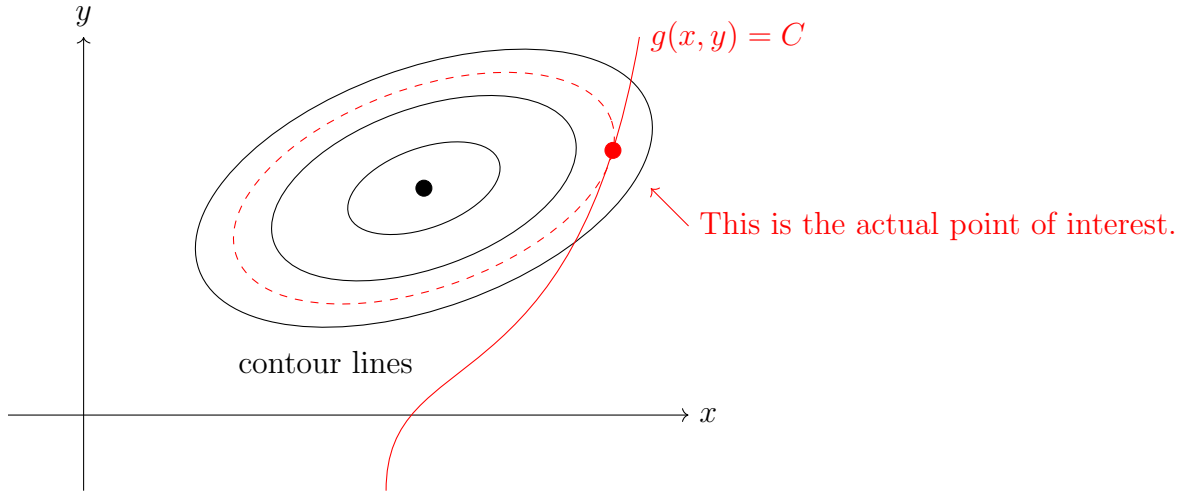
Thus

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

Let us define

$$\mathcal{L} = f(x, y) - \underset{\substack{\text{undetermined} \\ \text{multiplier}}}{\lambda} (g(x, y) - C)$$

$$\implies \delta\mathcal{L} = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y - \lambda \left(\frac{\partial g}{\partial x} \delta x + \frac{\partial g}{\partial y} \delta y \right) - (g(x, y) - C) \delta\lambda$$

$$(2.50)$$

By setting $\delta\mathcal{L} = 0$, since the variations δx, δy, and $\delta\lambda$ are independent, all of their coefficients must vanish.

$$\implies \begin{cases} \left(\dfrac{\partial f}{\partial x} - \lambda\dfrac{\partial g}{\partial x}\right) = 0 & (2.51) \\[2ex] \left(\dfrac{\partial f}{\partial y} - \lambda\dfrac{\partial g}{\partial y}\right) = 0 & (2.52) \\[2ex] (g(x,y) - C) = 0 & (2.53) \end{cases}$$

Note that the first two equations is equivalent to the condition $\nabla f \propto \nabla g$.

2.9   Hartree-Fock Equation

The ground-state energy is now given by

$$E_0 = \langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle = \sum_{n=1}^{N} \underbrace{[n|\mathbf{h}|n]}_{\text{1-electron}} + \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}\left(\underbrace{[nn|mm]}_{\text{Coulomb}} - \underbrace{[nm|mn]}_{\text{exchange}}\right) \quad (2.54)$$

Let us minimize E_0 subject to the constraint that $\{n\}$ is orthnormal (i.e. $[n|m] = \delta_{nm}$).

$$\mathcal{L} = E_0 - \sum_{n=1}^{N}\sum_{m=1}^{N}\lambda_{nm}([n|m] - \delta_{nm}) \quad (2.55)$$

We want to obtain $\delta\mathcal{L} = 0$. The trial functions will vary from $\{\mathcal{X}_n\} \to \{\mathcal{X}_n + \delta\mathcal{X}_n\}$, which we denote in shorthand as $\{n\} \to \{n + \delta n\}$.
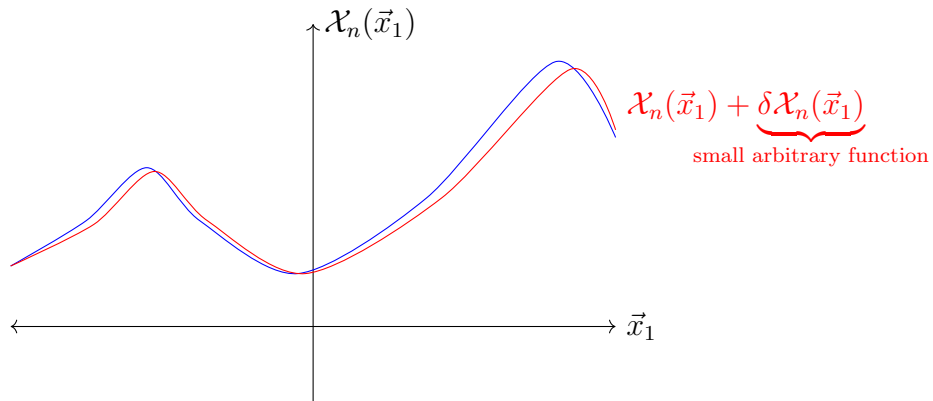
Thus

$$\delta[n|\mathbf{h}|n] = [(n+\delta n)|\mathbf{h}|(n+\delta n)] - [n|\mathbf{h}|n]$$

$$= \int \left(\mathcal{X}_n^*(\vec{x}_1) + \delta\mathcal{X}_n^*(\vec{x}_1)\right)\mathbf{h}(1)(\mathcal{X}_n(\vec{x}_1) + \delta\mathcal{X}_n(\vec{x}_1))\,\mathrm{d}\vec{x}_1$$

$$- \int \mathcal{X}_n^*(\vec{x}_1)\mathbf{h}(1)\mathcal{X}_n(\vec{x}_1)\,\mathrm{d}\vec{x}_1 \tag{2.56}$$

$$= \int \delta\mathcal{X}_n^*(\vec{x}_1)\,\mathbf{h}(1)\mathcal{X}_n(\vec{x}_1)\,\mathrm{d}\vec{x}_1 + \int \mathcal{X}_n^*(\vec{x}_1)\mathbf{h}(1)\,\delta\mathcal{X}_n(\vec{x}_1)\,\mathrm{d}\vec{x}_1$$

$$+ \int \delta Chi_n^*(\vec{x}_1)\,\mathbf{h}(1)\,\delta\mathcal{X}_n(\vec{x}_1)\,\mathrm{d}\vec{x}_1$$

In the limit of $\delta\mathcal{X}_n(\vec{x}_1)$ going to zero, the above expression simplifies to

$$\delta[n|\mathbf{h}|n] = \int \delta\mathcal{X}_n^*(\vec{x}_1)\,\mathbf{h}(1)\mathcal{X}_n(\vec{x}_1)\,\mathrm{d}\vec{x}_1 + \int \mathcal{X}_n^*(\vec{x}_1)\mathbf{h}(1)\,\delta\mathcal{X}_n(\vec{x}_1)\,\mathrm{d}\vec{x}_1$$

$$= [\delta n\,|\mathbf{h}|n] + [n|\mathbf{h}|\,\delta n] \tag{2.57}$$

For the Coulomb portion

$$\delta[nn|mm] = \ldots = [\delta n\,n|mm] + [n\,\delta n\,|mm] + [nn|\,\delta m\,m] + [nn|m\,\delta m] \tag{2.58}$$

The exchange portion is

$$\delta[nm|mn] = \ldots = [\delta n\,m|mn] + [n\delta m|mn] + [nm|\delta mn] + [nm|m\delta n] \tag{2.59}$$

Thus

$$\delta\mathcal{L} = \sum_{n=1}^{N}\left([\delta n\,|\mathbf{h}|n]\right) + \sum_{n=1}^{N}\sum_{m=1}^{N}\left([\delta n\,n|mm] - [\delta nm|mn]\right)$$

$$- \sum_{n=1}^{N}\sum_{m=1}^{N}\lambda_{nm}[\delta n\,|m] + \text{c.c.} \tag{2.60}$$

Setting $\delta L = 0$

$$0 = \sum_{n=1}^{N}\left[\int \underline{\delta\mathcal{X}_n^*(\vec{x}_1)}\mathbf{h}(1)\mathcal{X}_n(\vec{x}_1)\underline{\mathrm{d}\vec{x}_1} + \sum_{m=1}^{N}\left(\int \underline{\delta\mathcal{X}_n^*(\vec{x}_1)}\mathcal{X}_n(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_m(\vec{x}_2)\underline{\mathrm{d}\vec{x}_1}\,\mathrm{d}\vec{x}_2\right.\right.$$

$$\left.- \int \underline{\delta\mathcal{X}_n^*(\vec{x}_1)}\mathcal{X}_m(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_n(\vec{x}_2)\underline{\mathrm{d}\vec{x}_1}\,\mathrm{d}\vec{x}_2\right)$$

$$\left.- \sum_{m=1}^{N}\int \underline{\delta\mathcal{X}_n^*(\vec{x}_1)}\mathcal{X}_m(\vec{x}_1)\underline{\mathrm{d}\vec{x}_1}\right] + \text{c.c.}$$

$$(2.61)$$

Since the underlined factors above are common to all integrals

$$
\begin{aligned}
0 = \sum_{n=1}^{N} \Bigg[\int \delta\mathcal{X}_n^*(\vec{x}_1)\,\mathrm{d}\vec{x}_1 \Bigg[\mathbf{h}(1)\mathcal{X}_n(\vec{x}_1) + \sum_{n=1}^{N} \bigg(\int \mathcal{X}_n(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_m(\vec{x}_2)\,\mathrm{d}\vec{x}_2 \\
- \int \mathcal{X}_m(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_n(\vec{x}_2)\,\mathrm{d}\vec{x}_2 \bigg) - \sum_{m=1}^{N} \lambda_{nm}\mathcal{X}_m(\vec{x}_1) \Bigg]\Bigg]
\end{aligned}
$$

$$(2.62)$$

Since $\delta\mathcal{X}_n(\vec{x}_1)$ are arbitrary functions and $\delta\mathcal{X}_n(\vec{x}_1)$ are completely independent of each other for different n

$$
\begin{aligned}
\sum_{m=1}^{N} \lambda_{nm}\mathcal{X}_m(\vec{x}_1) = \mathbf{h}(1)\mathcal{X}_n(\vec{x}_1) + \sum_{m=1}^{N} \Bigg[\int \mathcal{X}_n(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_m(\vec{x}_2)\,\mathrm{d}\vec{x}_2 \\
- \int \mathcal{X}_m(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_n(\vec{x}_2)\,\mathrm{d}\vec{x}_2 \Bigg]
\end{aligned}
$$

$$(2.63)$$

Let us define

$$
\underbrace{\mathbf{J}_m(\vec{x}_1)}_{\substack{\text{Coulomb operator } m \\ \text{(local)}}}\mathcal{X}_n(\vec{x}_1) = \int \mathcal{X}_n(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_m(\vec{x}_2)\,\mathrm{d}\vec{x}_2
$$

$$
= \left(\int \mathcal{X}_m^*(\vec{x}_2)\frac{1}{r_{12}}\mathcal{X}_m(\vec{x}_2)\,\mathrm{d}\vec{x}_2 \right)\mathcal{X}_n(\vec{x}_1)
$$

$$(2.64)$$

$$
\underbrace{\mathbf{K}_m(\vec{x}_1)}_{\substack{\text{Exchange operator } m \\ \text{(non-local)}}}\mathcal{X}_n(\vec{x}_1) = \int \mathcal{X}_m(\vec{x}_1)\frac{1}{r_{12}}\mathcal{X}_m^*(\vec{x}_2)\mathcal{X}_n(\vec{x}_2)\,\mathrm{d}\vec{x}_2
$$

$$
= \left(\int \mathcal{X}_m^*(\vec{x}_2)\frac{1}{r_{12}}\mathcal{X}_n(\vec{x}_2)\,\mathrm{d}\vec{x}_2 \right)\mathcal{X}_m(\vec{x}_1)
$$

$$(2.65)$$

Thus

$$
\underbrace{\left[\mathbf{h}(1) + \sum_{m=1}^{N} \left(\mathbf{J}_m(1) - \mathbf{K}_m(1) \right) \right]}_{\mathbf{F}(\vec{x}_1)\,:\,\text{Fock operator (non-local operator)}}\mathcal{X}_n(\vec{x}_1) = \sum_{m=1}^{N} \lambda_{nm}\mathcal{X}_m(\vec{x}_1) \qquad (2.66)
$$

$$
\implies \mathbf{F}(\vec{x}_1)\mathcal{X}_n(\vec{x}_1) = \sum_{m=1}^{N} \lambda_{nm}\mathcal{X}_m(\vec{x}_1) \quad \text{(Hartree-Fock equation)} \qquad (2.67)
$$

To recap, we cast

$$\mathbf{H}\mathcal{X} = E\mathcal{X} \rightarrow \text{many-electron wavefunction} \tag{2.68}$$

into

$$\mathbf{F}\mathcal{X}_n = \sum_{m=1}^{N} \lambda_{nm} \mathcal{X}_m \tag{2.69}$$

one-electron wavefunction

where $\mathbf{F}(\vec{x}_1)$ depend on the one-electron functions themselves (which are the solutions that we are after). This equation needs to be done iteratively in a self-consistent manner. We will start with a one-electron wavefunction, build the Coulomb and exchange operators, and then solve for the new one-electron wavefunctions.

2.10   Unitary Transformations on Hartree-Fock Solutions

The Hartree-Fock equation contains an $N \times N$ matrix λ_{nm}. The equation can be greatly simplified by transforming λ_{nm} into a diagonal matrix.

Let us define

$$\mathbf{A} = \begin{bmatrix} \mathcal{X}_1(\vec{x}_1) & \dots & \mathcal{X}_N(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1(\vec{x}_N) & \dots & \mathcal{X}_N(\vec{x}_N) \end{bmatrix} \tag{2.70}$$

such that $\mathcal{X}_0(\vec{x}_1, \cdots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \det(\mathbf{A})$.

One can also define

$$\mathbf{A}' = \begin{bmatrix} \mathcal{X}'_1(\vec{x}_1) & \dots & \mathcal{X}'_N(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}'_1(\vec{x}_N) & \dots & \mathcal{X}'_N(\vec{x}_N) \end{bmatrix} \tag{2.71}$$

such that $\mathcal{X}'_0(\vec{x}_1, \cdots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \det(\mathbf{A}')$.

Assume that $\mathbf{A}'$ is built from $\mathbf{A}$ using a unitary transformation.

$$\mathbf{A}' = \mathbf{A}\mathbf{U}$$

$$\implies \begin{bmatrix} \mathcal{X}_1'(\vec{x}_1) & \dots & \mathcal{X}_N'(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1'(\vec{x}_N) & \dots & \mathcal{X}_N'(\vec{x}_N) \end{bmatrix} = \begin{bmatrix} \mathcal{X}_1(\vec{x}_1) & \dots & \mathcal{X}_N(\vec{x}_1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1(\vec{x}_N) & \dots & \mathcal{X}_N(\vec{x}_N) \end{bmatrix} \begin{bmatrix} U_{11} & \dots & U_{1N} \\ \vdots & \ddots & \vdots \\ U_{N1} & \dots & U_{NN} \end{bmatrix}$$

(2.72)

with $\mathbf{U}$ being unitary (i.e. $\mathbf{U}^\dagger \mathbf{U} = \mathbb{I}$).

Evaluating the probability density

$$\begin{aligned} \langle \mathcal{X}_0' | \mathcal{X}_0' \rangle &= \frac{1}{\sqrt{N!}} (\det \mathbf{A}')^* \frac{1}{\sqrt{N!}} (\det \mathbf{A}') \\ &= \frac{1}{N!} (\det \mathbf{A}'^\dagger)(\det \mathbf{A}') \\ &= \frac{1}{N!} \det (\mathbf{A}'\mathbf{A}'^\dagger) \\ &= \frac{1}{N!} \det (\mathbf{A}\mathbf{U}\mathbf{U}^\dagger\mathbf{A}^\dagger) \\ &= \frac{1}{N!} \det (\mathbf{A}\mathbf{A}^\dagger) \\ &= \langle \mathcal{X}_0 | \mathcal{X}_0 \rangle \end{aligned}$$

(2.73)

Thus, for two sets $\{\mathcal{X}_n\}$ and $\{\mathcal{X}_n'\}$ that are related by a unitary transformation, the physics of the problem remains the same.

2.11 Effect of Unitary Transformation on Fock Operator

Note that $\mathbf{F}(\vec{x}_1)$ contains the solutions $\mathcal{X}_n(\vec{x}_1)$ which will be under a unitary transformation.

$$h'(1) = \mathbf{h}(1) \tag{2.74}$$

$$
\begin{aligned}
\sum_{m=1}^{N} \mathbf{J}'_m(1) &= \sum_{m=1}^{N} \int \mathcal{X}'^*_m(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}'_m(\vec{x}_2) \, \mathrm{d}\vec{x}_2 \\
&= \sum_{m=1}^{N} \int \left(\sum_{i=1}^{N} \mathcal{X}^*_i(\vec{x}_2) U^*_{im} \right) \frac{1}{r_{12}} \left(\sum_{j=1}^{N} \mathcal{X}_j(\vec{x}_2) U_{jm} \right) \mathrm{d}\vec{x}_2 \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{m=1}^{N} U_{jm} U^*_{im} \int \mathcal{X}^*_i(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_j(\vec{x}_2) \, \mathrm{d}\vec{x}_2 \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} \underbrace{\sum_{m=1}^{N} U_{jm} U^*_{mi}}_{(\mathbf{U}\mathbf{U}^\dagger)_{ji} = \delta_{ji}} \int \mathcal{X}^*_i(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_j(\vec{x}_2) \, \mathrm{d}\vec{x}_2 \\
&= \sum_{i=1}^{N} \int \mathcal{X}^*_i(\vec{x}_2) \frac{1}{r_{12}} \mathcal{X}_i(\vec{x}_2) \, \mathrm{d}\vec{x}_2 \\
&= \sum_{i=1}^{N} J_i(1)
\end{aligned}
\tag{2.75}
$$

A similar derivation will show that the exchange portion of the Fock operator does not change under a unitary transformation. Thus, the Fock operator does not change under the unitary transformation.

2.12 Canonical Hartree-Fock Equation

The Hartree-Fock equation is

$$\mathbf{F} \left| \mathcal{X}_n \right\rangle = \sum_{m=1}^{N} \lambda_{nm} \left| \mathcal{X}_m \right\rangle \tag{2.76}$$

By orthonormality

$$\lambda_{nm} = \left\langle \mathcal{X}_m | \mathbf{F} | \mathcal{X}_n \right\rangle \tag{2.77}$$

For the constants under a unitary transformation

$$
\begin{aligned}
\lambda'_{nm} &= \langle \mathcal{X}'_m | \mathbf{F}' | \mathcal{X}'_n \rangle \\
&= \int \mathcal{X}'_m(\vec{x}_1) \mathbf{F}'(\vec{x}_1) \mathcal{X}'_n(\vec{x}_1) \, \mathrm{d}\vec{x}_1 \\
&= \int \left(\sum_{i=1}^{N} \mathcal{X}_i^*(\vec{x}_1) U_{im}^* \right) \mathbf{F}(\vec{x}_1) \left(\sum_{j=1}^{N} \mathcal{X}_j(\vec{x}_1) U_{jn} \right) \mathrm{d}\vec{x}_1 \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} U_{im}^* U_{jn} \underbrace{\int \mathcal{X}_i^*(\vec{x}_1) \mathbf{F}(\vec{x}_1) \mathcal{X}_j(\vec{x}_1) \, \mathrm{d}\vec{x}_1}_{\lambda_{ji}} \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} U_{jn} \lambda_{ji} U_{im}^* \\
&= \left(\mathbf{U}^\dagger \lambda \mathbf{U} \right)_{nm}^*
\end{aligned}
\tag{2.78}
$$

There exists a unitary transformation $\mathbf{U}$ that diagonalizes λ_{nm}. Among all possible sets of solutions $\{\mathcal{X}_n\}$, there exists one for which λ_{nm} is diagonal. Given the physical equivalence of different sets of solutions, we will choose to work with the set for which λ_{nm} is diagonal. Thus

$$
\mathbf{F}(\vec{x}_1) \mathcal{X}_n(\vec{x}_1) = \lambda_n \mathcal{X}_n(\vec{x}_1) \quad (\text{or } \mathbf{F} | \mathcal{X}_n \rangle = \lambda_n | \mathcal{X}_n \rangle) \tag{2.79}
$$

This is the canonical form of the Hartree-Fock equation.

2.13   Koopmans' Theorem

After obtaining the solutions for $| \mathcal{X}_n \rangle$, we obtain the n one-electron functions with the lowest energy to construct the ground-state Slater determinant.

Consider λ_n

$$
\lambda_n = \langle \mathcal{X}_n | \mathbf{F} | \mathcal{X}_n \rangle = \ldots = [n|\mathbf{h}|n] + \sum_{m=1}^{N} \left([nn|mm] - [nm|mn] \right)
$$

$$
\implies \sum_{n=1}^{N} \lambda_n = \sum_{n=1}^{N} [n|\mathbf{h}|n] + \sum_{n=1}^{N} \sum_{m=1}^{N} \left([nn|mm] - [nm|mn] \right)
$$

$$
\tag{2.80}
$$

Let us also recall the ground-state energy.

$$
\begin{aligned}
E_0 &= \langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle \\
&= \sum_{n=1}^{N} [n|\mathbf{h}|n] + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \left([nn|mm] - [nm|mn] \right)
\end{aligned}
\tag{2.81}
$$

While similar, $\sum_{n=1}^{N} \lambda_n \neq E_0$. One should not immediately associate λ_n as energies. Moreover, one should be aware that $\mathcal{X}_n$ are not physical objects by themselves.

Koopmans' theorem states that considering a system of N electrons, if one takes one of the filled one-electron levels (λ_f), then λ_f is the negative of the ionization potential corresponding to an electron in that level. This however assumes that the other levels and functions do not change as a result of such an ionization event.

$\lambda_f \uparrow$, assuming that removing this electron does not change the other levels and functions
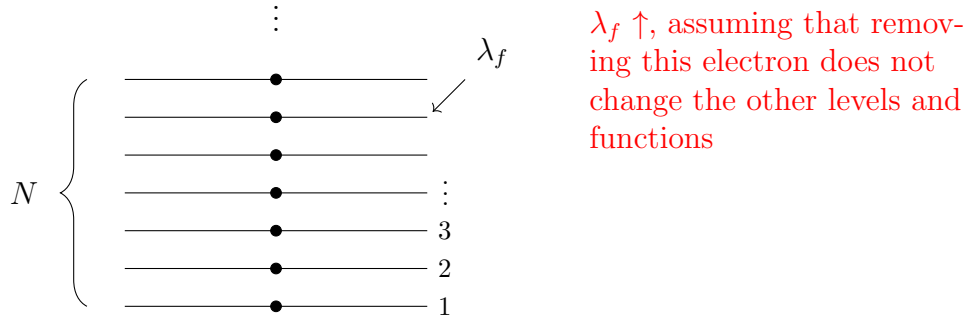
Proof: Consider a system of N electrons with wavefunction ${}^{N}|\mathcal{X}\rangle$ and energy ${}^{N}E_0$. Let us say that the system with $N-1$ electrons is described by ${}^{N-1}|\mathcal{X}\rangle$ and ${}^{N-1}E_0$.

The ionization potential is ${}^{N-1}E_0 - {}^{N}E_0$. Thus

$$
\begin{aligned}
{}^{N-1}E_0 - {}^{N}E_0 &= \sum_{\substack{n=1 \\ n\neq f}}^{N}[n|\mathbf{h}|n] + \frac{1}{2}\sum_{\substack{n=1 \\ n\neq f}}^{N}\sum_{\substack{m=1 \\ m\neq f}}^{N}\left([nn|mm]-[nm|mn]\right) \\
&\quad - \sum_{n=1}^{N}[n|\mathbf{h}|n] - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}\left([nn|mm]-[nm|mn]\right) \\
&= -[f|\mathbf{h}|f] - \frac{1}{2}\sum_{n=1}^{N}\left([nn|ff]-[nf|fn]\right) \\
&\quad - \frac{1}{2}\sum_{m=1}^{N}\left([ff|mm]-[fm|mf]\right) \\
&= -[f|\mathbf{h}|f] - \sum_{m=1}^{N}\left([ff|mm]-[fm|mf]\right) \\
&= -\lambda_f
\end{aligned}
$$

$$(2.82)$$

Similarly, if λ_e corresponds to an empty level, one can show that it represents the negative of an electron affinity (within the same assumption as before).

$$
{}^{N+1}E_0 - {}^{N}E_0 = \ldots = \lambda_e \tag{2.83}
$$

This theorem gives (approximate) physical legitimacy to $\{\lambda_n\}$ and $\{\mathcal{X}_n\}$. From now on, we call $\{\mathcal{X}_n\}$ <u>spin orbitals</u>. We will also use $\{\varepsilon_n\}$ instead of $\{\lambda_n\}$.

2.14   Spatial Form of the Hartree-Fock Equation

The Hartree-Fock equation is

$$
\mathbf{F}(\vec{x}_1)\mathcal{X}_n(\vec{x}_1) = \varepsilon_n\mathcal{X}_n(\vec{x}_1) \tag{2.84}
$$

where

$$
\mathcal{X}_n(\vec{x}_1) = \begin{cases} \Psi_{n\alpha}(\vec{r}_1)\alpha(\omega_1) & \text{or} \\ \Psi_{n\beta}(\vec{r}_1)\beta(\omega_1) \end{cases} \tag{2.85}
$$

In order to come up with an implementation of Hartree-Fock, we divide this equation into two possibilities.

For restricted Hartree-Fock (RHF), one assumes that spin orbitals come in pairs and that the spatial part remains the same (i.e. $\Psi_{n\alpha}(\vec{r}_1) = \Psi_{n\beta}(\vec{r}_1)$). This may work well for systems with paired electrons. Alternatively, for unrestricted Hartree-Fock (UHF), one does not have the above assumption. One can use either RHF or UHF for both close-shell (even N) or open-shell (odd N) systems.

Let us consider RHF for even N.

$$\mathcal{X}_1(\vec{x}_1) = \Psi_1(\vec{r}_1)\alpha(\omega_1) \quad , \quad \mathcal{X}_2(\vec{x}_1) = \Psi_1(\vec{r}_1)\beta(\omega_1)$$
$$\mathcal{X}_3(\vec{x}_1) = \Psi_2(\vec{r}_1)\alpha(\omega_1) \quad , \quad \mathcal{X}_4(\vec{x}_1) = \Psi_2(\vec{r}_1)\beta(\omega_1) \tag{2.86}$$
$$\vdots$$

$$\mathbf{F}(\vec{x}_1) = \mathbf{h}(1) + \sum_{m=1}^{N} (\mathbf{J}_m(1) - \mathbf{K}_m(1)) \tag{2.87}$$

$$\mathbf{F}(\vec{x}_1)\mathcal{X}_n(\vec{x}_1) = \varepsilon_n \mathcal{X}_n(\vec{x}_1) \tag{2.88}$$

After integrating over all spin parts

$$\mathbf{F}(\vec{r}_1)\Psi_n(\vec{r}_1) = \varepsilon_n \Psi_n(\vec{r}_1) \tag{2.89}$$

with

half vanish since $\int \alpha^*(\omega_1)\beta(\omega_1)\,\mathrm{d}\omega_1 = \int \beta^*(\omega_1)\alpha(\omega_1)\,\mathrm{d}\omega_1 = 0$

$$\mathbf{F}(\vec{r}_1) = \mathbf{h}(\vec{r}_1) + \sum_{m=1}^{N/2} (2\mathbf{J}_m(\vec{r}_1) - \mathbf{K}_m(\vec{r}_1)) \tag{2.90}$$

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

One then obtains $\frac{N}{2}$ orbitals each containing a pair of electrons with opposite spin.

2.15   Roothaan Equation

To implement a Hartree-Fock solver, one must choose a basis set $\{\phi_j(\vec{r}_1)\}$. Ideally, this should be a complete basis. In practice, one must choose a limited basis set with K functions ($j = 1, \cdots, K$). In this basis, $\Psi_n(\vec{r}_1)$ is represented as

$$\Psi_n(\vec{r}_1) = \sum_{j=1}^{K} c_{jn}\phi_j(\vec{r}_1) \tag{2.91}$$

The Hartree-Fock equation becomes

$$\mathbf{F}(\vec{r}_1)\left[\sum_{j=1}^{K} c_{jn}\phi_j(\vec{r}_1)\right] = \varepsilon_n\left[\sum_{j=1}^{K} c_{jn}\phi_j(\vec{r}_1)\right] \qquad (2.92)$$

Multiplying by $\phi_i^*(\vec{r}_1)$ and integrating

$$\int \phi_i^*(\vec{r}_1)\mathbf{F}(\vec{r}_1)\left[\sum_{j=1}^{K} c_{jn}\phi_j(\vec{r}_1)\right]\mathrm{d}\vec{r}_1 = \int \phi_i^*(\vec{r}_1)\varepsilon_n\left[\sum_{j=1}^{K} c_{jn}\phi_j(\vec{r}_1)\right]\mathrm{d}\vec{r}_1$$

$$\implies \sum_{j=1}^{K} c_{jn}\underbrace{\int \phi_i^*(\vec{r}_1)\mathbf{F}(\vec{r}_1)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1}_{\color{red}{F_{ij}\,:\,\text{Fock matrix element}}} = \sum_{j=1}^{K} \varepsilon_n c_{jn}\underbrace{\int \phi_i^*(\vec{r}_1)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1}_{\color{red}{S_{ij}\,:\,\text{Overlap matrix element}}}$$

$$(2.93)$$

where $i, \cdots, K$. In matrix form

$$\text{Fock matrix} \longrightarrow \mathbf{FC} = \mathbf{SC}\varepsilon \leftarrow \text{eigenvalue matrix} \qquad (2.94)$$

(with annotations: "overlap matrix" pointing to $\mathbf{S}$, "coefficient matrix" pointing to $\mathbf{C}$)

$$\begin{bmatrix} F_{11} & \dots & F_{1K} \\ \vdots & \ddots & \vdots \\ F_{K1} & \cdots & F_{KK} \end{bmatrix}\begin{bmatrix} c_{11} & \dots & c_{1K} \\ \vdots & \ddots & \vdots \\ c_{K1} & \cdots & c_{KK} \end{bmatrix} = \begin{bmatrix} S_{11} & \dots & S_{1K} \\ \vdots & \ddots & \vdots \\ S_{K1} & \cdots & S_{KK} \end{bmatrix}\begin{bmatrix} c_{11} & \dots & c_{1K} \\ \vdots & \ddots & \vdots \\ c_{K1} & \cdots & c_{KK} \end{bmatrix}\begin{bmatrix} \varepsilon_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \epsilon_K \end{bmatrix}$$

$$(2.95)$$

Note that this method only gives K solutions, namely

$$\Psi_1(\vec{r}_1) = \sum_{j=1}^{K} c_{j1}\phi_j(\vec{r}_1)$$

$$\Psi_2(\vec{r}_1) = \sum_{j=1}^{K} c_{j2}\phi_j(\vec{r}_1)$$

$$\vdots \qquad\qquad\qquad (2.96)$$

$$\Psi_K(\vec{r}_1) = \sum_{j=1}^{K} c_{jK}\phi_j(\vec{r}_1)$$

Calculating the Fock matrix

$$
\begin{aligned}
F_{ij} &= \int \phi_i^*(\vec{r}_1)\left[\mathbf{h}(\vec{r}_1) + \sum_{m=1}^{N/2}\left(2\mathbf{J}_m(\vec{r}_1) - \mathbf{K}_m(\vec{r}_1)\right)\right]\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1 \\
&= \int \phi_i^*(\vec{r}_1)\mathbf{h}(\vec{r}_1)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1 + \sum_{m=2}^{N/2} 2\left[\int \phi_i^*(\vec{r}_1)\Psi_m^*(\vec{r}_2)\frac{1}{r_{12}}\Psi_m(\vec{r}_2)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1\,\mathrm{d}\vec{r}_2 \right. \\
&\quad \left. - \frac{1}{2}\int \phi_i^*(\vec{r}_1)\Psi_m^*(\vec{r}_2)\frac{1}{r_{12}}\Psi_m(\vec{r}_1)\phi_j(\vec{r}_2)\,\mathrm{d}\vec{r}_1\,\mathrm{d}\vec{r}_2\right] \\
&= \int \phi_i^*(\vec{r}_1)\mathbf{h}(\vec{r}_1)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1 + \sum_{m=1}^{N/2} 2\left[\int \phi_i^*(\vec{r}_1)\sum_{\ell=1}^{K} c_{\ell m}^*\phi_\ell^*(\vec{r}_2)\frac{1}{r_{12}}\sum_{q=1}^{K} c_{qm}\phi_q(\vec{r}_2)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1\,\mathrm{d}\vec{r}_2 \right. \\
&\quad \left. - \frac{1}{2}\int \phi_i^*(\vec{r}_1)\sum_{\ell=1}^{K} c_{\ell m}^*\phi_\ell^*(\vec{r}_2)\frac{1}{r_{12}}\sum_{q=1}^{K} c_{qm}\phi_q(\vec{r}_1)\phi_j(\vec{r}_2)\,\mathrm{d}\vec{r}_1\,\mathrm{d}\vec{r}_2\right] \\
&= \underbrace{\int \phi_i^*(\vec{r}_1)\mathbf{h}(\vec{r}_1)\phi_j(\vec{r}_1)\,\mathrm{d}\vec{r}_1}_{h_{ij}} + \sum_{\ell=1}^{K}\sum_{q=1}^{K} 2\underbrace{\sum_{m=1}^{N/2} c_{\ell m}^* c_{qm}}_{P_{q\ell}}\left[\int \phi_i^*(\vec{r}_1)\phi_j(\vec{r}_1)\frac{1}{r_{12}}\phi_\ell^*(\vec{r}_2)\phi_q(\vec{r}_2)\,\mathrm{d}\vec{r}_1\,\mathrm{d}\vec{r}_2 \right. \\
&\quad \left. - \frac{1}{2}\int \phi_i^*(\vec{r}_1)\phi_q(\vec{r}_1)\frac{1}{r_{12}}\phi_\ell^*(\vec{r}_2)\phi_j(\vec{r}_2)\,\mathrm{d}\vec{r}_1\,\mathrm{d}\vec{r}_2\right] \\
&= h_{ij} + \sum_{\ell=1}^{K}\sum_{q=1}^{K} P_{q\ell}\left([ij|\ell q] - \frac{1}{2}[iq|\ell j]\right)
\end{aligned}
$$

$$(2.97)$$

The procedure for the self-consistent field (SCF) method thus becomes

1. Choose a basis set $\{\phi_j\}$

2. Calculate S_{ij} and h_{ij}

3. Construct a guess for c_{ij}

4. Calculate $P_{q\ell}$

5. Calculate F_{ij}

6. Solve $\mathbf{FC} = \mathbf{SC}\varepsilon$

   NO

7. Check convergence   YES   → Done

Note that one does not need to recalculate h_{ij}, S_{ij}, $[ij|\ell q]$, and $[iq|\ell j]$ in each iteration. Moreover, note that i, j, ℓ, and q go from 1 to K. Thus, this form of the Hartree-Fock equation scales to the fourth power of the number of basis functions.

2.16   Density Matrix and Mulliken Charge Population

Let us now regard the matrix P. Note that if N is the total number of electrons in the system and $\rho(\vec{r})$ is the electron density, then

$$N = \int \rho(\vec{r})\, \mathrm{d}\vec{r} \tag{2.98}$$

In terms of the 1-electron orbitals, one may say

$$\rho(\vec{r}) = \sum_{m=1}^{N/2} 2|\Psi_m(\vec{r})|^2 \quad \text{(for RHF)} \tag{2.99}$$

Thus

$$
\begin{aligned}
N &= \int \sum_{n=1}^{N/2} 2|\Psi_n(\vec{r})|^2\, \mathrm{d}\vec{r} \\
&= \int \sum_{n=1}^{N/2} 2 \sum_{\ell=1}^{K} c_{\ell n}^* \phi_\ell^*(\vec{r}) \sum_{q=1}^{K} c_{qn} \phi_q(\vec{r})\, \mathrm{d}\vec{r} \\
&= \sum_{\ell=1}^{K} \sum_{q=1}^{K} \underbrace{\sum_{n=1}^{N/2} 2 c_{\ell n}^* c_{qn}}_{P_{ql}} \underbrace{\int \phi_\ell^*(\vec{r}) \phi_q(\vec{r})\, \mathrm{d}\vec{r}}_{S_{\ell q}} \\
&= \sum_{\ell=1}^{K} \sum_{q=1}^{K} P_{q\ell} S_{\ell q} \\
&= \sum_{q=1}^{K} (\mathbf{PS})_{qq} \\
&= \mathrm{Tr}(\mathbf{PS})
\end{aligned}
\tag{2.100}
$$

The object $(\mathbf{PS})_{qq}$ can be associated with the number of electrons (which may be a fractional number) corresponding with basis function labeled by q. This is called <u>Mullikan population</u>. Note that P is also a density matrix.

2.17   Basis Functions

A given solution of the problem $\Psi_n(\vec{r})$ is expanded over K basis functions.

$$\Psi_n(\vec{r}) = \sum_{j=1}^{K} c_{jn}\phi_j(\vec{r}) \tag{2.101}$$

If the basis set is infinitely large, then the choice of basis should have no bearing on $\Psi_n(\vec{r})$. However, in practice, one wants a small basis set that is as "complete" as possible. In the context of molecules, one can think of molecular orbitals as being built up from combinations of atomic orbitals. As such, it is convenient to choose basis functions that resemble atomic orbitals (1s, 2s, 2p, etc.) centred around the different nuclei.

In obtaining a small basis set, the Slater type orbitals have been found to work well. The equations that describe these orbitals are

$$e^{-\xi|\vec{r}-\vec{R}_A|} \tag{2.102}$$

where $\vec{R}_A$ is the location of nucleus A.

For the purpose of computational efficiency, the Gaussian type orbitals simplify the integrals shown earlier in the Roothaan equation and are described by

$$\begin{aligned}
g_{\mathrm{s}j}(\vec{r}) &\equiv e^{-\alpha_j|\vec{r}-\vec{R}_A|^2} \quad \text{(s-type Gaussian)} \\
g_{\mathrm{p}j}(\vec{r}) &\equiv xe^{-\alpha_j|\vec{r}-\vec{R}_A|^2} \quad \text{(p-type Gaussian along } x\text{)} \\
g_{\mathrm{d}j}(\vec{r}) &\equiv xye^{-\alpha_j|\vec{r}-\vec{R}_A|^2} \quad \text{(d-type Gaussian along } xy\text{)}
\end{aligned} \tag{2.103}$$

The above are called primitive Gaussians.

To obtain the best features of both orbitals, one introduces contracted Gaussians. They are a weighted sum of a number of primitive Gaussians

$$\sum_j A_j g_j(\vec{r}) \tag{2.104}$$

A_j and α_j are typically determined by running Hartree-Fock on individual atoms in a molecule and matching these variables as closely to the solution given.

Note that these are not the only basis functions one can use. It may be desirable for the basis functions to have the same properties as the system being studied. For example, crystals have translational symmetry and long-range distribution. Thus, one could use plane-wave basis functions instead.

3   Post Hartree-Fock

3.1   Configuration Interaction

Let us return to the many-electron Schrödinger equation.

$$\to \mathbf{H}\mathcal{X} = E\mathcal{X} \leftarrow \tag{3.1}$$

many-electron Hamiltonian   many-electron wavefunction   ↑ energy

In Hartree-Fock, $\mathcal{X}$ was assumed to take the form of a Slater determinant. Let us not make this approximation. To solve this problem, we need a basis set to be able to expand $\mathcal{X}(\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_N)$ over.

Let us recall the Hartree-Fock equation.

$$\mathbf{F}\mathcal{X}_n(\vec{x}_1) = \varepsilon_n \mathcal{X}_n(\vec{x}_1) \quad , \quad n = 1, \cdots, \infty \tag{3.2}$$

Since $\mathbf{F}$ is Hermitian, the eigenfunctions form a complete basis set. If one takes an arbitrary function $\Psi(\vec{x}_1)$ of one electron, it can be expanded as

$$\Psi(\vec{x}_1) = \sum_{i=1}^{\infty} a_i \mathcal{X}_i(\vec{x}_1) \tag{3.3}$$

Now consider an arbitrary two-electron function $\Psi(\vec{x}_1, \vec{x}_2)$. If $\vec{x}_2$ is held fixed, the function depends on $\vec{x}_1$ only and can be expanded as done above. For a different value of fixed $\vec{x}_2$, the expansion above can still be done, but the values of the coefficients a_i must change. Thus

$$\Psi(\vec{x}_1, \vec{x}_2) = \sum_{i=1}^{\infty} a_i(\vec{x}_2)\mathcal{X}_i(\vec{x}_1) \tag{3.4}$$

However, if a_i is a function of $\vec{x}_2$, the coefficient itself can be expanded over the same basis set. Thus

$$a_i(\vec{x}_2) = \sum_{j=1}^{\infty} b_{ji}\mathcal{X}_j(\vec{x}_2)$$

$$\implies \Psi(\vec{x}_1, \vec{x}_2) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} b_{ji}\mathcal{X}_j(\vec{x}_2)\mathcal{X}_i(\vec{x}_1) \tag{3.5}$$

However, $\Psi(\vec{x}_1, \vec{x}_2)$ should be anti-symmetric.

$$\Psi(\vec{x}_1, \vec{x}_2) = -\Psi(\vec{x}_2, \vec{x}_1) \tag{3.6}$$

which leads to the constraint

$$b_{ji} = -b_{ij} \tag{3.7}$$

Thus

$$\Psi(\vec{x}_1, \vec{x}_2) = \sum_{i,j=1}^{\infty} b_{ji}[\mathcal{X}_i(\vec{x}_1)\mathcal{X}_j(\vec{x}_2) - \mathcal{X}_i(\vec{x}_2)\mathcal{X}_j(\vec{x}_1)] \tag{3.8}$$

with constants being absorbed into b_{ji}. Thus, an arbitrary two-electron wavefunction can be expanded over two-electron Slater determinants that are themselves constructed from the one-electron basis set found from Hartree-Fock. These sets of two-electron Slater determinants are themselves a complete basis set for all anti-symmetric two-electron wavefunctions. This argument generalizes to N-electron wavefunctions. Note that the one-electron basis set needs not to be obtained from Hartree-Fock for this argument to apply.

Thus, for configuration interaction (CI), one first solves Hartree-Fock with K basis functions. One then constructs all possible N-electron Slater determinants out of the Hartree-Fock solutions. There are $\binom{K}{N}$ of these Slater determinants. One then writes the original Schrödinger equation in matrix form using these Slater determinants as basis functions and solves for the eigenvalues and eigenfunctions. As K approaches infinity, these solutions become exact.
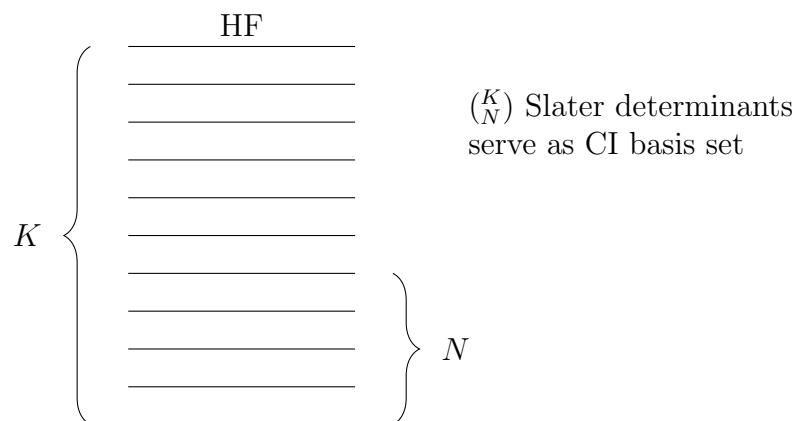
In practice, CI becomes extremely computationally intensive. However, there are variants one can use to truncate this set. For instance, using only the singly excited determinants or using only the singly and doubly excited determinants (CISD) drastically reduces the computational resources required.
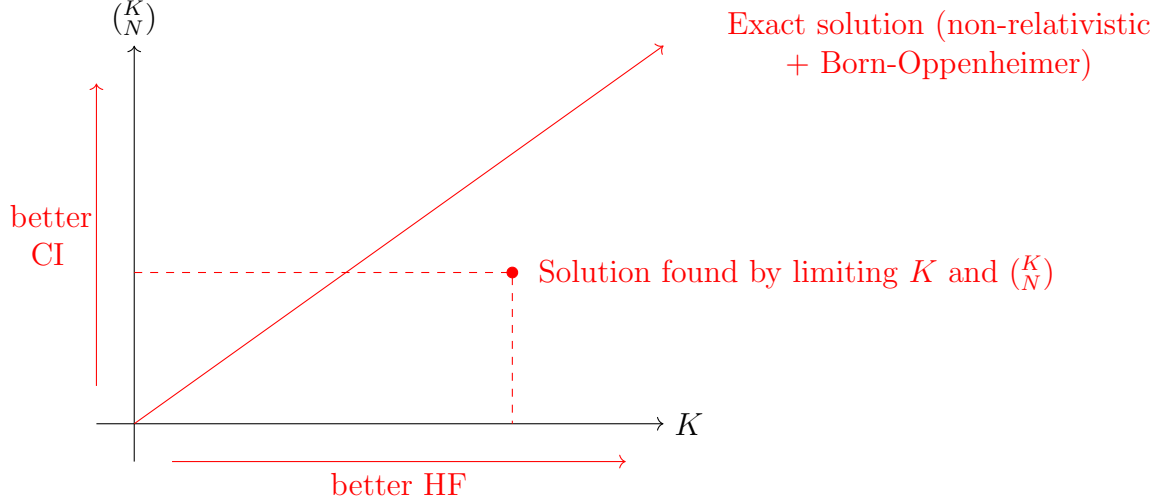
One can define the correlation energy as

$$E_{\text{correlation}} = E_{\text{exact}} - E_{\text{HF}} \tag{3.9}$$

3.2   Møller-Plesset Perturbation Theory

Let us write the many-body Hamiltonian as

$$\mathbf{H} = \mathbf{H}_0 + \lambda \mathbf{V} \longleftarrow \text{perturbation} \tag{3.10}$$

$\uparrow$

reference Hamiltonian

In Møller-Plesset perturbation theory (MPPT), $\mathbf{H}_0$ is the Hartree-Fock Hamiltonian (not the Fock operator). It is given by

$$\mathbf{H}_0 = \sum_{i=1}^{N} \mathbf{F}(i) \tag{3.11}$$

The energy for a given state will be written as

$$\begin{aligned}
E_i &= \sum_{j=0}^{\infty} \lambda^j E_i^{(j)} \\
&= E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots
\end{aligned} \tag{3.12}$$

The many-body wavefunction will be written as

$$\begin{aligned}
|\mathcal{X}_i\rangle &= \sum_{j=0}^{\infty} \lambda^j \left|\mathcal{X}_i^{(j)}\right\rangle \\
&= \left|\mathcal{X}_i^{(0)}\right\rangle + \lambda \left|\mathcal{X}_i^{(1)}\right\rangle + \lambda^2 \left|\mathcal{X}_i^{(2)}\right\rangle + \dots
\end{aligned} \tag{3.13}$$

Substituting into the many-electron Schrödinger equation

$$(\mathbf{H}_0 + \lambda\mathbf{V})\left(\left|\mathcal{X}_i^{(0)}\right\rangle + \lambda\left|\mathcal{X}_i^{(1)}\right\rangle + \dots\right) = \left(E_i^{(0)} + \lambda E_i^{(1)} + \dots\right)\left(\left|\mathcal{X}_i^{(0)}\right\rangle + \lambda\left|\mathcal{X}_i^{(1)}\right\rangle + \dots\right)$$

(3.14)

Matching orders of λ

$$\mathbf{H}_0\left|\mathcal{X}_i^{(0)}\right\rangle = E_i^{(0)}\left|\mathcal{X}_i^{(0)}\right\rangle \quad \text{(Zeroth order)}$$

$$\mathbf{H}_0\left|\mathcal{X}_i^{(1)}\right\rangle + \mathbf{V}\left|\mathcal{X}_i^{(0)}\right\rangle = E_i^{(0)}\left|\mathcal{X}_i^{(1)}\right\rangle + E_i^{(1)}\left|\mathcal{X}_i^{(0)}\right\rangle \quad \text{(First order)}$$

$$\mathbf{H}_0\left|\mathcal{X}_i^{(2)}\right\rangle + \mathbf{V}\left|\mathcal{X}_i^{(1)}\right\rangle = E_i^{(0)}\left|\mathcal{X}_i^{(2)}\right\rangle + E_i^{(1)}\left|\mathcal{X}_i^{(1)}\right\rangle + E_i^{(2)}\left|\mathcal{X}_i^{(0)}\right\rangle \quad \text{(Second order)}$$

$$\vdots$$

(3.15)

One can introduce normalization conditions

$$\left\langle\mathcal{X}_i^{(0)}\middle|\mathcal{X}_i^{(0)}\right\rangle = 1 \quad \text{(can always be done)}$$

(3.16)

Additionally, one can then choose

$$\left\langle\mathcal{X}_i^{(0)}\middle|\mathcal{X}_i\right\rangle = 1 \quad \text{(chosen to help find } |\mathcal{X}_i\rangle\text{)}$$

(3.17)

Thus

$$\left\langle\mathcal{X}_i^{(0)}\middle|\mathcal{X}_i^{(1)}\right\rangle = \left\langle\mathcal{X}_i^{(0)}\middle|\mathcal{X}_i^{(2)}\right\rangle = \dots = 0$$

(3.18)

Multiplying on the left by $\left\langle\mathcal{X}_i^{(0)}\right|$ for the first order equation

$$\overbrace{\left\langle\mathcal{X}_i^{(0)}\middle|\mathbf{H}\middle|\mathcal{X}_i^{(1)}\right\rangle}^{\text{0, since }\mathbf{H}\left|\mathcal{X}_i^{(0)}\right\rangle = E_i^{(0)}\left|\mathcal{X}_i^{(0)}\right\rangle} + \left\langle\mathcal{X}_i^{(0)}\middle|\mathbf{V}\middle|\mathcal{X}_i^{(0)}\right\rangle$$

$$= \overbrace{\left\langle\mathcal{X}_i^{(0)}\middle|E_i^{(0)}\middle|\mathcal{X}_i^{(1)}\right\rangle}^{0} + \left\langle\mathcal{X}_i^{(0)}\middle|E_i^{(1)}\middle|\mathcal{X}_i^{(0)}\right\rangle$$

$$\implies E_I^{(1)} = \left\langle\mathcal{X}_i^{(0)}\middle|\mathbf{V}\middle|\mathcal{X}_i^{(0)}\right\rangle$$

(3.19)

Solving for $\left|\mathcal{X}_i^{(1)}\right\rangle$

$$\left(E_i^{(0)} - \mathbf{H}_0\right)\left|\mathcal{X}_i^{(1)}\right\rangle = \left(\mathbf{V} - E_i^{(1)}\right)\left|\mathcal{X}_i^{(0)}\right\rangle$$

(3.20)

Since $\left\{ \left| \mathcal{X}_i^{(0)} \right\rangle \right\}$ forms a complete basis set, one can write

$$
\begin{aligned}
\left| \mathcal{X}_i^{(1)} \right\rangle &= \sum_j c_{ji} \left| \mathcal{X}_i^{(0)} \right\rangle \\
c_{ji} &= \left\langle \mathcal{X}_j^{(0)} \middle| \mathcal{X}_i^{(1)} \right\rangle
\end{aligned}
\tag{3.21}
$$

Thus

$$
\left\langle \mathcal{X}_j^{(0)} \middle| \left(E_i^{(0)} - \mathbf{H}_0 \right) \middle| \mathcal{X}_i^{(1)} \right\rangle = \left\langle \mathcal{X}_j^{(0)} \middle| \left(\mathbf{V} - E_i^{(1)} \right) \middle| \mathcal{X}_i^{(0)} \right\rangle
$$
$$
\implies E_i^{(0)} \left\langle \mathcal{X}_j^{(0)} \middle| \mathcal{X}_i^{(1)} \right\rangle - E_j^{(0)} \left\langle \mathcal{X}_j^{(0)} \middle| \mathcal{X}_i^{(1)} \right\rangle = \left\langle \mathcal{X}_j^{(0)} \middle| \mathbf{V} \middle| \mathcal{X}_i^{(0)} \right\rangle - E_i^{(1)} \left\langle \mathcal{X}_j^{(0)} \middle| \mathcal{X}_i^{(0)} \right\rangle
\tag{3.22}
$$

For $i \neq j$

$$
\begin{aligned}
c_{ji} &= \frac{\left\langle \mathcal{X}_j^{(0)} \middle| \mathbf{V} \middle| \mathcal{X}_i^{(0)} \right\rangle}{E_i^{(0)} - E_j^{(0)}} \\
\implies \left| \mathcal{X}_i^{(1)} \right\rangle &= \sum_{j \neq i} \frac{\left\langle \mathcal{X}_j^{(0)} \middle| \mathbf{V} \middle| \mathcal{X}_i^{(0)} \right\rangle}{E_i^{(0)} - E_j^{(0)}} \left| \mathcal{X}_j^{(0)} \right\rangle
\end{aligned}
\tag{3.23}
$$

4   Density Functional Theory

4.1   Functionals and Functional Differentiation

A <u>functional</u> maps a function to a numeric value. In the case of density functional theory, the electron density gets mapped to different values by various functionals.
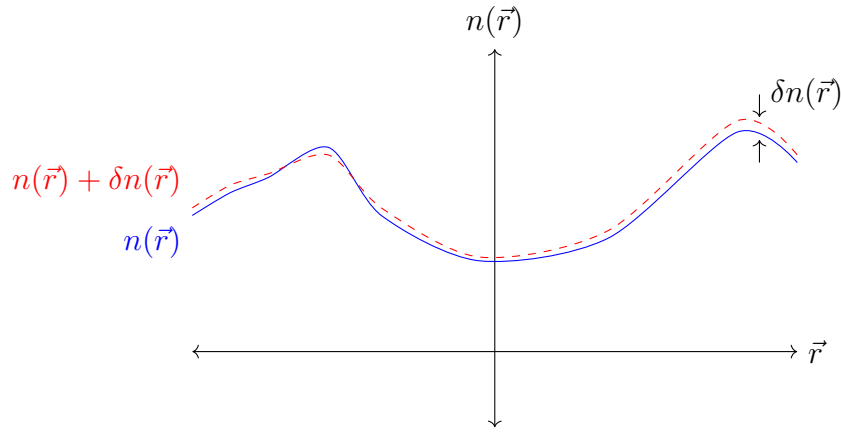
$$n(\vec{r}) \to F[n(\vec{r})]$$

For a difference in function $n(\vec{r})$, how does the functional $F[n(\vec{r})]$ change? This is given by

$$\frac{\delta F[n(\vec{r})]}{\delta n(\vec{r})} \quad \text{(functional differentation)}$$

where

$$\int \frac{\delta F[n(\vec{r})]}{\delta n(\vec{r})} g(\vec{r})\, \mathrm{d}\vec{r} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big[F[n(\vec{r}) + \varepsilon g(\vec{r})] - F[n(\vec{r})] \Big] \tag{4.1}$$

<u>Ex</u>:   Consider

$$F[n(\vec{r})] = \int v(\vec{r}) n(\vec{r})\, \mathrm{d}\vec{r} \tag{4.2}$$

By definition

$$\int \frac{\delta F[n(\vec{r})]}{\delta n(\vec{r})} g(\vec{r})\, \mathrm{d}\vec{r} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int v(\vec{r})[n(\vec{r}) + \varepsilon g(\vec{r})]\, \mathrm{d}\vec{r} - \int v(\vec{r}) n(\vec{r})\, \mathrm{d}\vec{r} \right]$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int v(\vec{r}) \varepsilon g(\vec{r})\, \mathrm{d}\vec{r}$$

$$= \int v(\vec{r}) g(\vec{r})\, \mathrm{d}\vec{r}$$

$$(4.3)$$

Thus, since $g(\vec{r})$ is arbitrary

$$\frac{\delta F[n(\vec{r})]}{\delta n(\vec{r})} = v(\vec{r}) \tag{4.4}$$

4.2   Density Functional Theory

The original many-electron problem had the form of

$$\mathbf{H}\ket{\mathcal{X}_i} = E_i \ket{\mathcal{X}_i} \quad , \quad i = 1, \cdots, \infty \quad , \quad \ket{\mathcal{X}_i} \equiv \mathcal{X}_i(\vec{x}_1, \cdots, \vec{x}_N) \tag{4.5}$$

Density functional theory (DFT) transforms this problem into

$$\mathbf{H}_{\text{eff}}\Psi_i(\vec{r}_1) = \varepsilon_i \Psi_i(\vec{r}_1) \tag{4.6}$$

The history of density functional theory can be traced back to two seminal papers: Hohenberg and Kohn (1964) and Kohn and Sham (1965). In essence, it recasts the problem from regarding the many-electron wavefunction to regarding the electron density $n(\vec{r})$.

4.3   Hohenberg-Kohn Theorems

Recall the many-electron Hamiltonian is

$$\mathbf{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} + \sum_{i=1}^{N}\overbrace{\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}^{\text{external potential } v(\vec{r}_i)} \tag{4.7}$$

$$\underbrace{\phantom{\sum_{i=1}^{N}\sum_{A=1}^{M} -\frac{Z_A}{r_{iA}}}}_{\substack{\text{total external potential acting} \\ \text{on the electronic system}}}$$

<u>Theorem 1</u>: The external potential $v(\vec{r})$ acting as a fully-interacting system of N electrons in a ground state is determined, within an additive constant, by the electron density $n(\vec{r})$.

<u>Proof:</u> Imagine two external potentials, $v(\vec{r})$ and $v'(\vec{r})$, which differ by more than an additive constant, which lead to the same electron density $n(\vec{r})$ in the ground state.

Let $\mathbf{H}$ be the Hamiltonian associated with $v(\vec{r})$, with ground state $|\mathcal{X}_0\rangle$ and ground state energy E_0. Let $\mathbf{H}'$ be the Hamiltonian associated with $v'(\vec{r})$, with ground state $|\mathcal{X}_0'\rangle$ and ground state energy E_0'.

By the variational principle

$$
\begin{aligned}
\langle \mathcal{X}_0'|\mathbf{H}|\mathcal{X}_0'\rangle &> E_0 \\
\langle \mathcal{X}_0|\mathbf{H}|\mathcal{X}_0\rangle &> E_0'
\end{aligned}
\tag{4.8}
$$

The terms evaluated explicitly are

$$
\begin{aligned}
\langle \mathcal{X}_0'|\mathbf{H}|\mathcal{X}_0'\rangle = & \int \mathcal{X}_0'^*(\vec{x}_1, \cdots, \vec{x}_N) \left(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 \right) \mathcal{X}_0'(\vec{x}_1, \cdots, \vec{x}_N)\, \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N \\
& + \int \mathcal{X}_0'^*(\vec{x}_1, \cdots, \vec{x}_N) \left(\sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} \right) \mathcal{X}_0'(\vec{x}_1, \cdots, \vec{x}_N)\, \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N \\
& + \int \mathcal{X}_0'^*(\vec{x}_1, \cdots, \vec{x}_N) \left(\sum_{i=1}^{N} v(\vec{r}_i) \right) \mathcal{X}_0'(\vec{x}_1, \cdots, \vec{x}_N)\, \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N \\
= & \text{ Kinetic term} + \text{e}^-\text{-e}^- \text{ interaction term} \\
& + \int \mathcal{X}_0'^*(\vec{x}_1, \cdots, \vec{x}_N) \left(\sum_{i=1}^{N} \int v(\vec{r})\delta(\vec{r}_i - \vec{r})\, \mathrm{d}\vec{r} \right) \mathcal{X}_0'(\vec{x}_1, \cdots, \vec{x}_N)\, \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N
\end{aligned}
\tag{4.9}
$$

The third term becomes

$$
\int v(\vec{r})\, \mathrm{d}\vec{r} \overbrace{\int \mathcal{X}_0'^*(\vec{x}_1, \cdots, \vec{x}_N) \underbrace{\left(\sum_{i=1}^{N} \delta(\vec{r}_i - \vec{r}) \right)}_{\text{density operator}} \mathcal{X}_0'(\vec{x}_1, \cdots, \vec{x}_N)\, \mathrm{d}\vec{x}_1 \ldots \mathrm{d}\vec{x}_N}^{n(\vec{r})}
\tag{4.10}
$$

Thus

$$
\langle \mathcal{X}_0'|\mathbf{H}|\mathcal{X}_0'\rangle = \text{Kinetic term} + \text{e}^-\text{-e}^- \text{ interaction term} + \int v(\vec{r})n(\vec{r})\, \mathrm{d}\vec{r}
$$

$$(4.11)$$

By a similar argument

$$\langle \mathcal{X}_0 | \mathbf{H}' | \mathcal{X}_0 \rangle = \text{Kinetic term} + \text{e}^- \text{-e}^- \text{ interaction term} + \int v'(\vec{r}) n(\vec{r}) \, \mathrm{d}\vec{r}$$

$$(4.12)$$

Note that

$$\begin{aligned}
\langle \mathcal{X}_0' | \mathbf{H} | \mathcal{X}_0' \rangle &= \langle \mathcal{X}_0' | (\mathbf{H} - \mathbf{H}') | \mathcal{X}_0' \rangle + \langle \mathcal{X}_0' | \mathbf{H}' | \mathcal{X}_0' \rangle \\
&= \int (v(\vec{r}) - v'(\vec{r})) n(\vec{r}) \, \mathrm{d}\vec{r} + E_0' \\
\implies E_0 &< \int (v(\vec{r}) - v'(\vec{r})) n(\vec{r}) \, \mathrm{d}\vec{r} + E_0'
\end{aligned}$$

$$(4.13)$$

Similarly

$$\begin{aligned}
\langle \mathcal{X}_0 | \mathbf{H}' | \mathcal{X}_0 \rangle &= \langle \mathcal{X}_0 | (\mathbf{H}' - \mathbf{H}) | \mathcal{X}_0 \rangle + \langle \mathcal{X}_0 | \mathbf{H} | \mathcal{X}_0 \rangle \\
&= \int (v'(\vec{r}) - v(\vec{r})) n(\vec{r}) \, \mathrm{d}\vec{r} + E_0 \\
\implies E_0' &< \int (v'(\vec{r}) - v(\vec{r})) n(\vec{r}) \, \mathrm{d}\vec{r} + E_0
\end{aligned}$$

$$(4.14)$$

Thus, adding the two inequalities together

$$E_0 + E_0' < E_0 + E_0' \qquad (4.15)$$

which is a contradiction.

Thus, $n(\vec{r})$ uniquely determines $v(\vec{r})$, up to an additive constant, which will give the Hamiltonian $\mathbf{H}$. This subsequently gives everything else. Therefore, one can write the various energies as a functional of the electron density $n(\vec{r})$.

$$E_0[n(\vec{r})] = \underbrace{T[n(\vec{r})] + U_{\text{ee}}[n(\vec{r})]}_{F_{\text{HK}}[n(\vec{r})] \,:\, \text{Hohenberg-Kohn functional}} + \int v(\vec{r}) n(\vec{r}) \, \mathrm{d}\vec{r} \qquad (4.16)$$

Thus, one can see that the problem involving $3N$ spatial coordinates has been converted into a problem involving only 3 spatial coordinates.

<u>Theorem 2</u>: This restates the variational principle in terms of electron density.

$$E_0[n'(\vec{r})] > E_0 \tag{4.17}$$

<u>Proof</u>: Note that $n'(\vec{r})$ determines the Hamiltonian, which ultimately gives the state $|\mathcal{X}_0'\rangle$. From the previous form of the variational principle

$$\underbrace{\langle \mathcal{X}_0'|\mathbf{H}|\mathcal{X}_0'\rangle}_{= F_{\mathrm{HK}}[n'(\vec{r})] + \int v(\vec{r})n'(\vec{r})\mathrm{d}\vec{r} = E_0[n'(\vec{r})]} > E_0 \tag{4.18}$$

4.4   Kohn-Sham Method

The functional describing the electron-electron interaction can be broken up into a classical Coulomb term and a non-classical term

$$U_{\mathrm{ee}}[n(\vec{r})] = \underbrace{\frac{1}{2}\int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r} - \vec{r}'|}\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{r}'}_{J[n(\vec{r})]} + \text{non-classical term} \tag{4.19}$$

In this method, imagine a reference system of non-interacting electrons

$$\mathbf{H}_{\mathrm{ref}} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N} v_{\mathrm{ref}}(\vec{r}_i) \quad \overset{\text{\color{red}acts only on one electron}}{\color{red}\text{because they are non-interacting}} \tag{4.20}$$

with $\mathbf{H}_{\mathrm{ref}}\Psi_i(\vec{r}) = \varepsilon_i \Psi_i(\vec{r})$ and $\Psi_i(\vec{r})$ is some one-electron function.

The key feature is that the external potential of the reference system v_{ref} is chosen such that the electron density of this reference system is equal to the electron density of the actual system of interest (where the electrons are interesting). The electron density is given by

$$n(\vec{r}) = \sum_{i=1}^{N} \Psi_i^*(\vec{r})\Psi_i(\vec{r}) \tag{4.21}$$

The kinetic energy of the reference system is given by

$$T_{\mathrm{ref}}[n(\vec{r})] = \sum_{i=1}^{N} \langle \Psi_i|-\frac{1}{2}\nabla_i^2|\Psi_i\rangle \neq T[n(\vec{r})] \tag{4.22}$$

$$\color{red}\text{do not assume they are the same}$$

For the actual system

$$E_0[n(\vec{r})] = \underbrace{T_{\text{ref}}[n(\vec{r})] + \text{kinetic energy missing from reference system}}_{T[n(\vec{r})]}$$

$$+ \underbrace{J[n(\vec{r})] + \text{non-classical term}}_{U_{\text{ee}}[n(\vec{r})]} + \int v(\vec{r}) n(\vec{r}) \, d\vec{r} \qquad (4.23)$$

The sum of these two unknown terms is the exchange-correlation energy $E_{xc}[n(\vec{r})]$.

$$\implies E_0[n(\vec{r})] = T_{\text{ref}}[n(\vec{r})] + J[n(\vec{r})] + \int v(\vec{r}) n(\vec{r}) \, d\vec{r} + E_{\text{xc}}[n(\vec{r})] \quad (4.24)$$

Since the reference system was chosen to match electron density, $n(\vec{r})$ can be calculated using the definition above using the reference $\Psi_i(\vec{r})$.

We now minimize $E_0[n(\vec{r})]$ subject to the constraint that

$$\int \Psi_i^*(\vec{r}) \Psi_j(\vec{r}) \, d\vec{r} = \delta_{ij} \qquad (4.25)$$

Define a Lagrange function

$$\mathcal{L}[n(\vec{r})] = E_0[n(\vec{r})] - \sum_{i=1}^{N} \sum_{j \neq i}^{N} \varepsilon_{ij} \left(\int \Psi_i^*(\vec{r}) \Psi_j(\vec{r}) \, d\vec{r} - \delta_{ij} \right) \qquad (4.26)$$

Set $\delta\mathcal{L} = 0$ when the set $\{\Psi_i\}$ is changed to $\{\Psi_i + \delta\Psi_i\}$. To first order, the variation in $n(\vec{r})$ is

$$\delta n(\vec{r}) = \sum_{i=1}^{N} [\delta\Psi_i^*(\vec{r}) \, \Psi_i(\vec{r}) + \Psi_i^*(\vec{r}) \, \delta\Psi_i(\vec{r})] \qquad (4.27)$$

For the terms in $E_0[n(\vec{r})]$, the variations are

$$\delta T_{\text{ref}}[n(\vec{r})] = \sum_{i=1}^{N} \int \delta\Psi_i^*(\vec{r}) \left(-\frac{1}{2}\nabla^2 \right) \Psi_i(\vec{r}) \, d\vec{r} + \text{c.c.} \qquad (4.28)$$

$$\delta J[n(\vec{r})] = \int \frac{n(\vec{r}') \, \delta n(\vec{r})}{|\vec{r} - \vec{r}'|} \, d\vec{r} \, d\vec{r}'$$

$$= \int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|} \, d\vec{r}' \left(\sum_{i=1}^{N} \delta\Psi_i^*(\vec{r}) \, \Psi_i(\vec{r}) + \text{c.c.} \right) d\vec{r} \qquad (4.29)$$

$$\delta\left(\int v(\vec{r})n(\vec{r})\,\mathrm{d}\vec{r}\right) = \int v(\vec{r})\,\delta n(\vec{r})\,\mathrm{d}\vec{r}$$

$$= \int v(\vec{r})\left(\sum_{i=1}^{N}\delta\Psi_i^*(\vec{r})\,\Psi_i(\vec{r}) + \text{c.c.}\right)\mathrm{d}\vec{r} \tag{4.30}$$

The variation in the exchange-correlation energy is

$$\delta E_{\text{xc}}[n(\vec{r})] = \int \frac{\delta E_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})}\,\delta n(\vec{r})\,\mathrm{d}\vec{r}$$

$$= \int \frac{\delta E_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})}\left(\sum_{i=1}^{N}\delta\Psi_i^*(\vec{r})\,\Psi_i(\vec{r}) + \text{c.c.}\right)\mathrm{d}\vec{r} \tag{4.31}$$

Thus

$$\delta\mathcal{L}[n(\vec{r})] = \sum_{i=1}^{N}\int \mathrm{d}\vec{r}\,\delta\Psi_i^*(\vec{r})\left\{\left[-\frac{1}{2}\nabla^2 + \int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}' + v(\vec{r}) + \frac{\delta E_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})}\right]\Psi_i(\vec{r})\right.$$

$$\left.- \sum_{j=1}^{N}\varepsilon_{ij}\Psi_j(\vec{r})\right\} + \text{c.c.}$$

$$\tag{4.32}$$

For $\delta\mathcal{L} = 0$ where $\{\delta\Psi_i\}$ are arbitrary

$$\left[-\frac{1}{2}\nabla^2 + \int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}' + v(\vec{r}) + \frac{\delta E_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})}\right]\Psi_i(\vec{r}) = \sum_{j=1}^{N}\varepsilon_{ij}\Psi_j(\vec{r}) \tag{4.33}$$

Without loss of generality, let us work with diagonal matrix elements ε_i. We shall also define an effective potential given by

$$v_{\text{ref}}(\vec{r}) = v(\vec{r}) + \int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}' + \frac{\delta E_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})} \tag{4.34}$$

Thus, we recast the above equation as

$$\mathbf{H}_{\text{eff}}\Psi_i(\vec{r}) = \varepsilon_i\Psi_i(\vec{r}) \quad , \quad i = 1,\cdots,\infty \tag{4.35}$$

where

$$\mathbf{H}_{\text{eff}}(\vec{r}) = -\frac{1}{2}\nabla^2 + v_{\text{eff}}(\vec{r}) \tag{4.36}$$

and

$$n(\vec{r}) = \sum_{i=1}^{N} \Psi_i^*(\vec{r}) \Psi_i(\vec{r}) \tag{4.37}$$

These are the <u>Kohn-Sham equations</u>. Thus, the steps of the Kohn-Sham method are

1. Guess initial $\{\Psi_i\}$

2. Calculate $n(\vec{r})$

3. Calculate $v_{\text{eff}}(\vec{r})$

4. Calculate $\mathbf{H}_{\text{eff}}$ NO

5. Solve $\mathbf{H}_{\text{eff}}\Psi_i = \varepsilon_i \Psi_i$

6. Test for convergence ——————— YES ——————→ Done

4.5   Total Energy and Koopmans' Theorem in DFT

Recall the total energy functional is

$$E_0[n(\vec{r})] = T_{\text{ref}}[n(\vec{r})] + \int v(\vec{r})n(\vec{r})\,\mathrm{d}\vec{r} + \frac{1}{2}\int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{r}' + E_{\text{xc}}[n(\vec{r})] \tag{4.38}$$

We shall relate ε_i to this value. Note that

$$\sum_{i=1}^{N}\varepsilon_i = \sum_{i=1}^{N}\langle\Psi_i|\mathbf{H}_{\text{eff}}|\Psi_i\rangle$$

$$= \sum_{i=1}^{N}\langle\Psi_i|\left(-\frac{1}{2}\nabla_i^2 + v(\vec{r}) + \int\frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}' + \overbrace{\frac{\delta E_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})}}^{V_{\text{xc}}[n(\vec{r})]}\right)|\Psi_i\rangle \tag{4.39}$$

$$= T_{\text{ref}}[n(\vec{r})] + \int v(\vec{r})n(\vec{r})\,\mathrm{d}\vec{r} + \int\frac{n(\vec{r}')n(\vec{r})}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}'\,\mathrm{d}\vec{r} + E_{\text{xc}}[n(\vec{r})]$$

$$+ \int V_{\text{xc}}[n(\vec{r})]n(\vec{r})\,\mathrm{d}\vec{r}$$

$$\implies E_0[n(\vec{r})] = \sum_{i=1}^{N}\varepsilon_i - \frac{1}{2}\int\frac{n(\vec{r}')n(\vec{r})}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}'\,\mathrm{d}\vec{r} + E_{\text{xc}}[n(\vec{r})] - \int V_{\text{xc}}[n(\vec{r})]n(\vec{r})\,\mathrm{d}\vec{r}$$

$$(4.40)$$

Thus

$$E_0[n(\vec{r})] \neq \sum_{i=1}^{N} \varepsilon_i \tag{4.41}$$

In DFT, Koopmans' theorem states that the first ionization energy of this system is equal to the negative of the energy of the highest occupied molecular orbital. A proof of this theorem follows a similar vein of the proof done for Hartree-Fock.

Let us now consider the chemical potential μ using the following Lagrange function.

$$\mathcal{L}[n(\vec{r})] = E_0[n(\vec{r})] - \mu\left(\int n(\vec{r})\,\mathrm{d}\vec{r} - N\right) \tag{4.42}$$

Setting $\delta\mathcal{L} = 0$

$$\mu = \frac{\delta E_0[n(\vec{r})]}{\delta n(\vec{r})} = v(\vec{r}) + \delta F_{\mathrm{HK}}[n(\vec{r})] \tag{4.43}$$

4.6   Exchange Correlation Functionals

What is $E_{\mathrm{xc}}[n(\vec{r})]$?

With local density approximation (LDA), one assumes that the exchange-correlation potential at each point in space is equal to the exchange-correlation potential of a uniform electron gas whose density is equal to the density of the actual system at that point.

$$E_{\mathrm{xc}}[n(\vec{r})] = \int \varepsilon_{\mathrm{xc}}[n(\vec{r})]n(\vec{r})\,\mathrm{d}\vec{r} \tag{4.44}$$

where

$$\varepsilon_{\mathrm{xc}}[n(\vec{r})] = \underbrace{\varepsilon_{\mathrm{x}}[n(\vec{r})]}_{\substack{\text{exchange} \\ \text{term}}} + \underbrace{\varepsilon_{\mathrm{c}}[n(\vec{r})]}_{\substack{\text{correlation} \\ \text{term}}} \tag{4.45}$$

One possible exchange term is the Dirac exchange.

$$\varepsilon_{\mathrm{x}}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n^{\frac{1}{3}} \quad \text{(for a uniform electron gas)} \tag{4.46}$$

Another is the Slater exchange

$$\varepsilon_{\mathrm{x}}[n] = -\frac{3}{2}\alpha\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n^{\frac{1}{3}} \tag{4.47}$$

↑ adjustable parameter, often around 0.75

Some common examples of exchange-correlation functionals include:

1. von Barth and Hedin

2. Gunnarsson and Lundqvist

3. Vosko, Wilk, and Nusair (VWN)

Another category invokes the generalized gradient approximation (GGA). These exchange-correlation functionals go beyond LDA and involves a dependence on $\nabla n(\vec{r})$

$$E_{\mathrm{xc}}[n(\vec{r})] = \int \varepsilon_{xc}[n(\vec{r})]n(\vec{r})\,\mathrm{d}\vec{r} + \int g[\nabla n(\vec{r})]\,\mathrm{d}\vec{r} \tag{4.48}$$

where $g[\nabla n(\vec{r})]$ is some functional acting on $\nabla n(\vec{r})$. Some examples include:

1. Becke, Lee, Yang, Parr (BLYP), which uses the Becke 88 exchange which tries to approximate the Hartree-Fock exchange by adding corrections to LDA exchange and LYP correlation term (which tries to reproduce the correlation of the helium atom)

2. Perdew, Burke, Ernzerhot (PBE)

There are also hybrid methods, combining DFT with Hartree-Fock. Some examples include:

1. B3LYP, whose exchange term combines the Hartree-Fock exchange with the LDA exchange and Becke correction and whose correlation term combines the LYP correlation with the VWN correlation

When comparing computational time for different methods, in general

BLYP < HF < B3LYP < CI

In comparison, the accuracy in general

HF < BLYP < B3LYP < CI

Note in the context of modeling systems, the choice of exchange-correlation functional will depend highly on the system as well as the property of interest.

4.7   Connection to Thomas-Fermi-Dirac Model

The Thomas-Fermi model assumes uniformly distributed electrons in the phase space and that the effective potential comes due to the nuclear charge and the electron distribution itself. To obtain the kinetic energy, one uses a 3D particle-in-a-box problem. This gives the density of states

$$g(\varepsilon) \sim \sqrt{\varepsilon} \tag{4.49}$$

The electron density then is

$$n \sim \int f_{\text{FD}}(\varepsilon)g(\varepsilon)\,\mathrm{d}\varepsilon \sim \varepsilon_F^{\frac{3}{4}} \quad (\text{at 0 K}) \tag{4.50}$$

Fermi-Dirac statistics

The kinetic energy density is

$$\rho_{\text{KE}} \sim \int \varepsilon f_{\text{FD}}(\varepsilon)g(\varepsilon)\,\mathrm{d}\varepsilon \sim \varepsilon_F^{\frac{5}{2}} \sim n^{\frac{5}{3}} \tag{4.51}$$

Thus, kinetic energy is

$$T_{\text{TF}}[n(\vec{r})] = c_F \int n^{\frac{5}{3}}(\vec{r})\,\mathrm{d}\vec{r} \tag{4.52}$$

and the total energy is

$$E_{\text{TF}}[n(\vec{r})] = c_F \int n^{\frac{5}{3}}(\vec{r})\,\mathrm{d}\vec{r} + \int -\frac{Z}{r}n(\vec{r})\,\mathrm{d}\vec{r} + \frac{1}{2}\int \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}\mathrm{d}\vec{r}' \tag{4.53}$$

One assumes that this is minimized for the ground state density, subject to the constraint

$$N = \int n(\vec{r})\,\mathrm{d}\vec{r} \tag{4.54}$$

Thus, the Lagrange problem after setting the variation to 0 is

$$\delta\left\{ E_{\text{TF}}[n(\vec{r})] - \mu_{\text{TF}}\left(\int n(\vec{r})\,\mathrm{d}\vec{r} - N\right) \right\} = 0 \tag{4.55}$$

One gets

$$\mu_{\text{TF}} = \frac{\delta E_{\text{TF}}[n(\vec{r})]}{\delta n(\vec{r})} = \frac{5}{3}c_F n^{\frac{2}{3}}(\vec{r}) + \underbrace{\int \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|}\,\mathrm{d}\vec{r}' - \frac{Z}{r}}_{V_{\text{eff}}(\vec{r})} \tag{4.56}$$

In the context of DFT

$$E_0[n(\vec{r})] = T[n(\vec{r})] + U_{\text{ee}}[n(\vec{r})] + \int v(\vec{r})n(\vec{r})\,\mathrm{d}\vec{r} \tag{4.57}$$

In the Thomas-Fermi model, $T[n(\vec{r})]$ is that of a uniform, non-interacting electron gas. This can be calculated using Hartree-Fock and the states of a particle-in-a-box problem with periodic boundary conditions, giving $c_F \int n^{\frac{5}{3}}(\vec{r})\,\mathrm{d}\vec{r}$. Likewise, in this model, $U_{\text{ee}}[n(\vec{r})]$ is simply the Coulomb term $J[n(\vec{r})]$.

Adding the Dirac exchange gives the Thomas-Fermi-Dirac model. Using the Hartree-Fock exchange formula with the same particle-in-a-box states as above gives

$$K[n(\vec{r})] = c_{\text{x}} \int n^{\frac{4}{3}}(\vec{r})\,\mathrm{d}\vec{r} \tag{4.58}$$

Thus, the total energy is

$$E_{\text{TFD}}[n(\vec{r})] = c_F \int n^{\frac{5}{3}}(\vec{r})\,\mathrm{d}\vec{r} + \int v(\vec{r})n(\vec{r})\,\mathrm{d}\vec{r} + J[n(\vec{r})] - c_{\text{x}} \int n^{\frac{4}{3}}(\vec{r})\,\mathrm{d}\vec{r} \tag{4.59}$$

The Lagrange method gives

$$\mu_{\text{TFD}} = \frac{5}{3}c_F n^{\frac{2}{3}}(\vec{r}) - \frac{4}{3}c_{\text{x}} n^{\frac{1}{3}}(\vec{r}) + v_{\text{eff}}(\vec{r}) \tag{4.60}$$

This can also be improved with the Thomas-Fermi-Dirac-Weizsacker model, which also adds a second-order gradient correction.

The problem currently is formulated in terms of electron density $n(\vec{r})$, but it can be connected to wavefunctions $\Psi_i(\vec{r})$ using the Hartree-Fock-Slater formalism

$$\left[-\frac{1}{2}\nabla^2 + v(\vec{r}) + J[n(\vec{r})] - K[n(\vec{r})] \right] \Psi_i(\vec{r}) = \varepsilon_i \Psi_i(\vec{r}) \tag{4.61}$$

which resembles the Kohn-Sham equations.

4.8   Bloch's Theorem and Band Structures

Bloch's theorem states that wavefunctions in a system with periodicity can be written as

$$\Psi_{nk}(\vec{r}) = \mathcal{U}_{nk}(\vec{r})e^{i\vec{k}\cdot\vec{r}} \tag{4.62}$$

where $\mathcal{U}_{nk}(\vec{r})$ has the same periodicity as that of a crystal, namely $\mathcal{U}_{nk}(\vec{r} + \vec{T}) = \mathcal{U}_{nk}(\vec{r})$ with $\vec{T}$ describing the period. Thus

$$\Psi_{nk}(\vec{r} + \vec{T}) = \mathcal{U}_{nk}(\vec{r})e^{i\vec{k}\cdot\vec{r}}e^{i\vec{k}\cdot\vec{T}} = \Psi_{nk}(\vec{r})e^{i\vec{k}\cdot\vec{T}} \tag{4.63}$$

Let us cast $\mathcal{U}_{nk}(\vec{r})$ in the form of its Fourier series

$$\mathcal{U}_{nk}(\vec{r}) = \sum_{\vec{G}} c_{nk,\vec{G}}e^{i\vec{G}\cdot\vec{r}} \tag{4.64}$$

where $\vec{G}$ is called the <u>reciprocal lattice vector</u>.

Recall the Schrödinger equation for the Kohn-Sham method is

$$\underbrace{\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\vec{r}) \right]}_{\mathbf{H}_{\text{eff}}} \Psi_{nk} = \varepsilon_{nk}\Psi_{nk} \tag{4.65}$$

Thus

$$\mathbf{H}_{\text{eff}}\left(\sum_{\vec{G}} c_{nk,\vec{G}}e^{i(\vec{k}+\vec{G})\cdot\vec{r}} \right) = \varepsilon_{nk}\left(\sum_{\vec{G}} c_{nk,\vec{G}}e^{i(\vec{k}+\vec{G})\cdot\vec{r}} \right)$$

$$\implies \sum_{\vec{G}} \left[\frac{1}{2}\left| \vec{k}+\vec{G} \right|^2 + v_{\text{eff}}(\vec{r}) \right] c_{nk,\vec{G}}e^{i(\vec{k}+\vec{G})\cdot\vec{r}} = \varepsilon_{nk}\left(\sum_{\vec{G}} c_{nk,\vec{G}}e^{i(\vec{k}+\vec{G})\cdot\vec{r}} \right) \tag{4.66}$$

Multiplying both sides by $e^{-i(\vec{k}+\vec{G}')\cdot\vec{r}}$ and integrate over the space of $\vec{r}$

$$\sum_{\vec{G}} \int \left[\frac{1}{2}\left| \vec{k}+\vec{G} \right|^2 + v_{\text{eff}}(\vec{r}) \right] e^{i(\vec{G}-\vec{G}')\cdot\vec{r}}\,\mathrm{d}\vec{r}\, c_{nk,\vec{G}} = \varepsilon_{nk}\left[\sum_{\vec{G}} \int c_{nk,\vec{G}}e^{i(\vec{G}-\vec{G}')\cdot\vec{r}} \right]$$

$$\implies \sum_{\vec{G}} \left[\frac{1}{2}\left| \vec{k}+\vec{G} \right|^2 \delta(\vec{G}-\vec{G}') + \underbrace{V_{\text{eff}}(\vec{G}-\vec{G}')}_{\text{Fourier transform of } v_{\text{eff}}} \right] c_{nk,\vec{G}} = \varepsilon_{nk}c_{nk,\vec{G}} \tag{4.67}$$

Assume a basis set (of plane waves) containing M different vectors $\vec{G}$. Thus, the above equation can be written for all $\vec{G}'$ that are part of this set. This gives the matrix equation

$$\mathbf{A}\vec{c} = \varepsilon\vec{c} \tag{4.68}$$

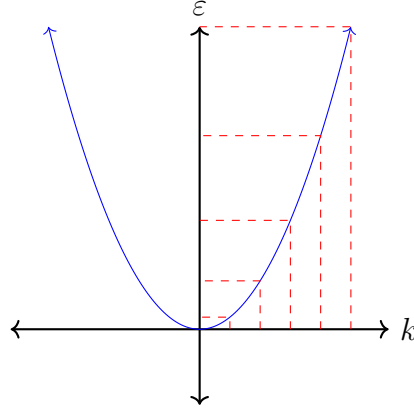where the matrix elements A_{ij} are given by

$$A_{ij} = \frac{1}{2}\left|\vec{k} + \vec{G}_i\right|^2 \delta(\vec{G}_i - \vec{G}) + V_{\text{eff}}(\vec{G}_i - \vec{G}) \quad (4.69)$$

and the vector elements are $c_{nk,\vec{G}_i}$.

Note that k must be chosen for the coefficients $c_{nk,\vec{G}}$. This becomes a choice of k-point sampling. One choice is the Monkhorst-Pack scheme which uses uniform sampling. One can also implement a cut-off energy

$$\frac{1}{2}\left|\vec{k} + \vec{G}\right|^2 < E_{\text{cutoff}} \quad (4.70)$$

4.9   Finite Temperature DFT

At finite temperature, entropy can no longer be ignored. One now has to look at the Helmholtz free energy.

The grand canonical potential is defined as

$$\Omega[n(\vec{r})] = T[n(\vec{r})] + U_{\text{ee}}[n(\vec{r})] - \Theta S[n(\vec{r})] + \int (v(\vec{r}) - \mu)n(\vec{r})\,\mathrm{d}\vec{r} \quad (4.71)$$

where Θ is the temperature, S is the entropy, and μ is the chemical potential.

The Helmholtz free energy is

$$A[n(\vec{r})] = T[n(\vec{r})] - \Theta S[n(\vec{r})] \quad (4.72)$$

Let us define a reference system similar as before.

$$A_{\text{ref}}[n(\vec{r})] = T_{\text{ref}}[n(\vec{r})] - \Theta S_{\text{ref}}[n(\vec{r})] \quad (4.73)$$

Thus

$$\Omega[n(\vec{r})] = A_{\text{ref}}[n(\vec{r})] + J[n(\vec{r})] + \int (v(\vec{r}) - \mu)n(\vec{r}) \, \mathrm{d}\vec{r} + \underbrace{F_{\text{xc}}[n(\vec{r})]}_{\substack{\text{exchange correlation} \\ \text{functional}}} \tag{4.74}$$

where

$$F_{\text{xc}}[n(\vec{r})] = (A[n(\vec{r})] - A_{\text{ref}}[n(\vec{r})]) + (U_{\text{ee}}[n(\vec{r})] - J[n(\vec{r})]) \tag{4.75}$$

Minimizing $\Omega[n(\vec{r})]$ gives the finite temperature Kohn-Sham equations. One gets

$$\frac{\delta A_{\text{ref}}[n(\vec{r})]}{\delta n(\vec{r})} + \int \frac{n(\vec{r'})}{|\vec{r} - \vec{r'}|} \, \mathrm{d}\vec{r'} + (v(\vec{r}) - \mu) + \frac{\delta F_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})} = 0 \tag{4.76}$$

Thus

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\vec{r}) \right)\Psi_i(\vec{r}) = \varepsilon_i \Psi_i(\vec{r}) \tag{4.77}$$

where

$$v_{\text{eff}}(\vec{r}) = v(\vec{r}) + \int \frac{n(\vec{r'})}{|\vec{r} - \vec{r'}|} \, \mathrm{d}\vec{r'} + \frac{\delta F_{\text{xc}}[n(\vec{r})]}{\delta n(\vec{r})} \tag{4.78}$$

and

$$n(\vec{r}) = \sum_{i=1}^{\infty} f_i \Psi_i^*(\vec{r})\Psi_i(\vec{r}) \tag{4.79}$$

with f_i being the Fermi-Dirac distribution given by

$$f_i = \frac{1}{1 + e^{\frac{\varepsilon_i - \mu}{k_B T}}} \tag{4.80}$$

The exchange-correlation functional can be dealt with much in a similar manner as was done at zero temperature. For example, one can do LDA, where one assumes the potential at each point in space equals the free energy of a uniform electron gas at temperature Θ whose density is equal to the density of the actual system at that point.

4.10   Time-Dependent DFT

Recall in time-independent DFT, the electron density $n(\vec{r})$ uniquely determines the external potential $v(\vec{r})$ to an additive constant. A similar relation holds in the time-dependent case and was shown in 1984 by Runge and Gross, but the additive constant is now time-dependent, though it remains spatially constant. Note that this assume $v(\vec{r}, t)$ can be Taylor expanded about some initial time and both $v(\vec{r}, t)$ and $n(\vec{r}, t)$ vanish at the boundary of the system.

Functionals now depend on $[n, \Psi_0, \phi_0]$, where Ψ_0 is the initial many-body state at time t_0 (starting point) and ϕ_0 is the initial Kohn-Sham state.

Moreover

$$v_{\text{eff}}(\vec{r}, t) = v(\vec{r}, t) + \int \frac{n(\vec{r}', t)}{|\vec{r} - \vec{r}'|} \, \mathrm{d}\vec{r}' + v_{\text{xc}}[n, \Psi_0, \phi_0](\vec{r}, t) \tag{4.81}$$

The effective Hamiltonian equation is

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\vec{r}, t) \right) \Psi_j(\vec{r}, t) = i\frac{\partial}{\partial t}\Psi_j(\vec{r}, t) \tag{4.82}$$

with

$$\Psi_j(\vec{r}, t) = \Psi_{\text{amp},j}(\vec{r}, t) \exp\left(i\Psi_{\text{phase},j}(\vec{r}, t) \right) \tag{4.83}$$

and

$$n(\vec{r}, t) = \sum_{j=1}^{N} \Psi_j^*(\vec{r}, t)\Psi_j(\vec{r}, t) \tag{4.84}$$

$v_{\text{xc}}[n, \Psi_0, \phi_0](\vec{r}, t)$ can be treated with an adiabatic approximation, where the functional uses LDA given the electron density at that given time.

5   Semi-empirical Electronic Structure Modeling

At the start of the course, topics focused on ab-initio (first-principles) methods, such as Hartree-Fock, which did not bring any parameters from empirical observations, experiments, or other methods. This was followed by post Hartree-Fock methods, which include configuration interaction and perturbation theory. Afterwards, we considered density functional theory.

Let us now consider semi-empirical methods. These will include linear combinations of atomic orbitals, tight-binding theory, Hückel theory, PPP methods, etc.

We are trying to solve problems of the form $\mathbf{FC} = \mathbf{SC}\varepsilon$. In practice, semi-empirical methods approximate $\mathbf{F}$ and $\mathbf{S}$ in order to reduce the computations required to solve this eigenvalue problem.

5.1   Hückel Theory

Consider Hückel π-orbital theory. In this theory, the core electrons and their effects are lumped in together with the nuclei. Slater type orbitals are also used as basis functions to describe π-orbitals. Furthermore, the basis functions are assumed to be orthogonal (i.e. $\mathbf{S} = \mathbb{I}$).

Let us now attempt to determine the matrix elements of the effective Hamiltonian. Note that the same atom, put in different molecular systems, would lead to different Hamiltonian matrix elements. This can be illustrated in the system of pyridine and pyrrole.

pyridine (protonated)

pyrrole
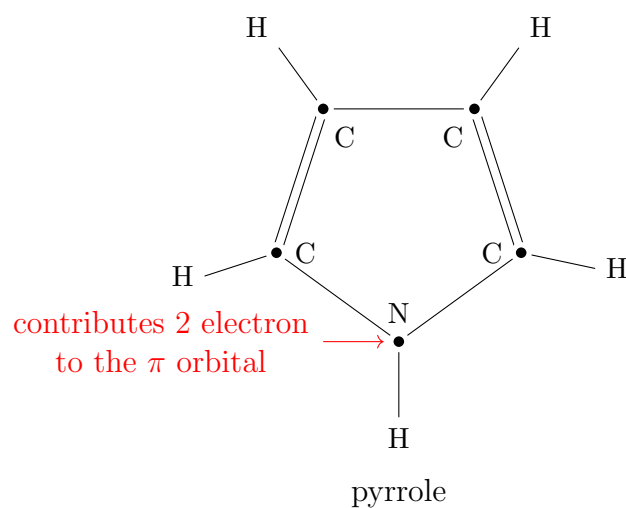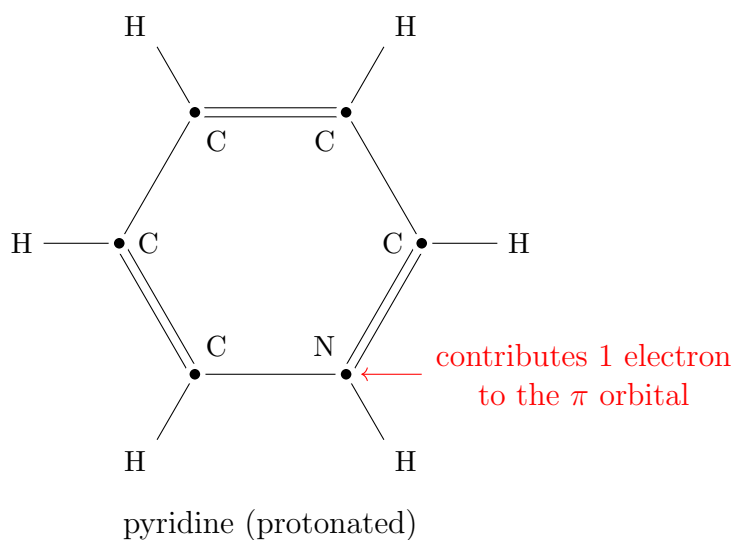
The Hückel parameter would be different for the nitrogen atoms in the two cases above. The effective Hamiltonian is written as follows:

- Diagonal elements for carbon atoms are α_C

- Off-diagonal elements for carbon-carbon are β_{CC}

- For other atoms: $\begin{cases} \alpha_x = \alpha_c + h_x\beta_{CC} \\ \beta_{xy} = k_{xy}\beta_{CC} \end{cases}$

For example, for N contributing one π electron yields $h_x = 0.5$ and $k_{xy} = 0.84$ (for connection to C). For example, in pyridine
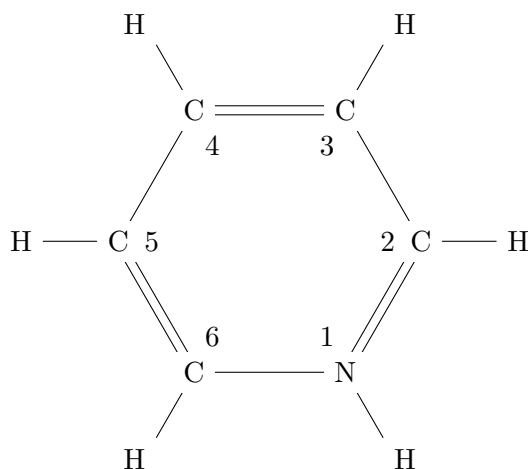
$$\mathbf{F} = \begin{bmatrix} \alpha_\mathrm{C} + 0.5\beta_\mathrm{CC} & 0.8\beta_\mathrm{CC} & 0 & 0 & 0 & 0.8\beta_\mathrm{CC} \\ 0.8\beta_\mathrm{CC} & \alpha_\mathrm{C} & \beta_\mathrm{CC} & 0 & 0 & 0 \\ 0 & \beta_\mathrm{CC} & \alpha_\mathrm{C} & \beta_\mathrm{CC} & 0 & 0 \\ 0 & 0 & \beta_\mathrm{CC} & \alpha_\mathrm{C} & \beta_\mathrm{CC} & 0 \\ 0 & 0 & 0 & \beta_\mathrm{CC} & \alpha_\mathrm{C} & \beta_\mathrm{CC} \\ 0.8\beta_\mathrm{CC} & 0 & 0 & 0 & \beta_\mathrm{CC} & \alpha_\mathrm{C} \end{bmatrix} \tag{5.1}$$

$\mathbf{S}$ is taken to be unity ($\mathbb{I}$). There is also no iterative process as was done in Hartree-Fock.

This describes basic Hückel theory. Consider now extended Hückel theory. This treats all valence electrons explicitly and uses Slater type orbital basis functions. Depending on which atoms are being treated, additional basis functions may also be included (e.g. Li to Ne includes 2s and 2p orbitals). The diagonal elements of the Hamiltonian are the negative of the valence shell ionization energies (which are obtained from atomic spectra). The off-diagonal elements are written as

$$h_{ij} = kS_{ij}\frac{h_{ii} + h_{jj}}{2} \tag{5.2}$$

where k is an empirical parameter to fit experimental data and S_{ij} is the overlap integral that is calculated exactly. Note that for the purpose of solving the eigenvalue problem, the basis functions are considered to be orthogonal.

5.2 Pariser, Parr, Pople (PPP) Method

Recall in Hartree-Fock, one was required to compute two-electron integrals $[ij|k\ell]$. The number of integrals scaled with the fourth power of the number

of electrons in the problem. Instead of computing each integral, let us discard some integrals and parametrize some of the other ones. The central idea to this is <u>zero-differential overlap (ZDO)</u>. This states that the integral $[ij|k\ell]$ is zero if i and j or k and ℓ are two basis functions centred on different atoms. We are still solving the Roothaan equation

$$\mathbf{FC} = \mathbf{SC}\varepsilon \tag{5.3}$$

where

$$[ij|k\ell] = \int \phi_i^*(\vec{r_1})\phi_j(\vec{r_1})\frac{1}{r_{12}}\phi_k^*(\vec{r_2})\phi_\ell(\vec{r_2}) \, d\vec{r_1} \, d\vec{r_2} \tag{5.4}$$

This reduces the scaling from $\mathcal{O}(N^4)$ to $\mathcal{O}(N^2)$.

The Fock matrix element for first raw atoms is

$$F_{ii} = \underbrace{W_i}_{\substack{\text{valence ionization} \\ \text{energy}}} + \frac{1}{2}\overbrace{P_{ii}}^{\substack{\text{density matrix} \\ \text{elements}}}\gamma_{ii} + \sum_{k \neq i}(P_{kk} - Z_k)\gamma_{ik} \tag{5.5}$$

where

$$\gamma_{ik} = \frac{1}{R_{ik} + \alpha_{ik}} \leftarrow \text{determined using spectral data} \tag{5.6}$$

$\nwarrow$ separation between two centres

and

$$F_{ij} = \beta_{ij} - \frac{1}{2}P_{ij}\gamma_{ij} \tag{5.7}$$

One can also include <u>neglect of diatomic differential overlap (NDDO)</u>, where the one-electron integrals that involve three centres are neglected. In other words, for $[i|h|j]$, terms that come from nuclei that do not host i or j are neglected.

Consider <u>intermediate neglect of differential overlap (INDO)</u>. Compared to ZDO, $[ij|k\ell]$ vanishes if $i \neq j$ or $k \neq \ell$ (i.e. must be the same basis function), unless if i, j, k, and ℓ all lie on the same centre. For <u>complete neglect of differential overlap (CNDO)</u>, $[ij|k\ell] = 0$ if $i \neq j$ or $k \neq \ell$ in all cases.

For complete neglect of differential overlap, some parameterizations include:

- All remaining two-electron integrals involving atoms A and B are assumed to be equal (regardless of basis functions involved)

$$F_{ii} = \underbrace{U_{ii}}_{=\int \phi_i^*(\vec{r}_1)\left(-\frac{1}{2}\nabla^2 - \frac{Z_A}{R_A}\right)\phi_i(\vec{r}_1)\mathrm{d}\vec{r}_1} + \left(P_{AA} - \frac{1}{2}P_{ii}\right)\gamma_{AA} + \sum_{B \neq A}(P_{BB}\gamma_{AB} - \underbrace{V_{AB}}_{\substack{\text{contributions from} \\ \text{other nuclei}}}) \quad (5.8)$$

with P_{AA} being the sum of diagonals for all the basis functions on atom A. Additionally

$$F_{ij} = \underbrace{\beta_{AB}}_{\substack{\text{parameter chosen from comparison} \\ \text{with ab-initio calculations}}} S_{ij} - \frac{1}{2}P_{ij}\gamma_{AB} \qquad (5.9)$$

- Alternatively, U_{ii} is calculated as

$$U_{ii} = -\frac{1}{2}(\underbrace{I_i}_{\text{ionization energy}} + \underbrace{E_i}_{\text{electron affinity}}) - \left(Z_A - \frac{1}{2}\right)\gamma_{AA} \quad , \quad V_{AB} = -Z_B\gamma_{AB} \quad (5.10)$$

The list continues and includes modified INDO (MINDO), modified neglect of differential order (MNDO), Austin model 1 (AM1), third parameterization of MNDO (PM3), semi ab-initio model (SAM1), and so on.

6   Molecular Mechanics and Molecular Dynamics

The force on each particle will be due to interactions with all other particles. A force field that describes these interactions can be calculated by quantum mechanical ab-initio calculations, but can also come from semi-empirical approaches.
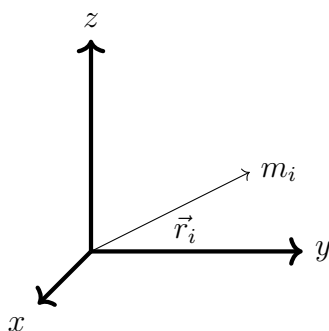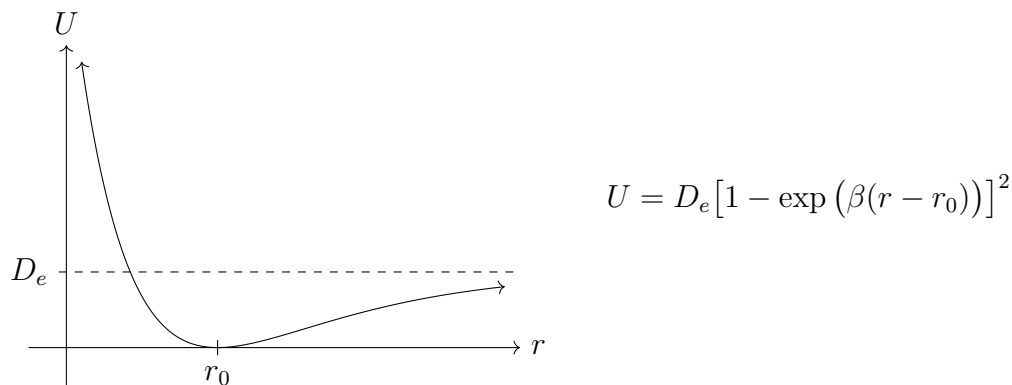

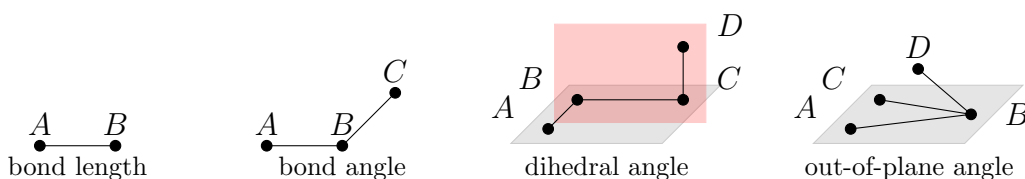Consider some basic force fields. One example is Hooke's law.

$$\vec{F} = -\vec{\nabla}U$$
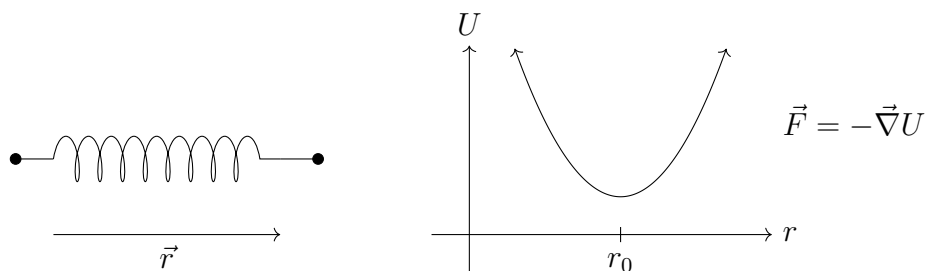
Regardless of the distance r, the two atoms will always want to come back together. As such, no bond dissociation is captured.

One improvement to this force field is the Morse potential.

$$U = D_e\left[1 - \exp\left(\beta(r - r_0)\right)\right]^2$$

A more general form for the force field looks like

$$U_{\text{tot}} = \sum_{A,B} U_{AB} + \sum_{A,B,C} U_{ABC} + \sum_{A,B,C,D} U_{ABCD,\text{dih}}$$
$$+ \sum_{A,B,C,D} U_{ABCD,\text{out of plane}} + \sum U_{\text{non-bonded}} + \sum U_{\text{Coulomb}} \tag{6.1}$$

bond length   bond angle   dihedral angle   out-of-plane angle

One example is

$$U_{\text{tot}} = \sum_{\text{bonds}} \frac{1}{2} k_{AB} (R_{AB} - R_{\text{eq},AB})^2 + \sum_{\text{bonds}} \frac{1}{2} k_{ABC} (\theta_{ABC} - \theta_{\text{eq},ABC})^2$$
$$+ \sum_{\text{dihedrals}} \frac{U_0}{2} \left[1 - \cos\left(n(x - x_0)\right) \right] + \sum_{\substack{\text{out-of-}\\\text{plane}}} \frac{k}{2\sin^2 \varphi_{\text{eq}}} (\cos\psi - \cos\psi_0)^2$$
$$+ \sum_{\substack{\text{non-}\\\text{bonded}}} \left(\frac{c_{AB}^{12}}{R_{AB}^{12}} - \frac{c_{AB}^6}{R_{AB}^6} \right) + \frac{1}{4\pi\epsilon_0} \sum_{\text{charges}} \frac{Q_A Q_B}{R_{AB}}$$

$$\tag{6.2}$$

Some examples of available force fields include

- Dreiding
  - Works for H, C, N, O, P, S ,F, Cl, Br, I
  - Does not have a Coulomb term
- MM1
  - Works for hydrocarbons
  - Does not have a Coulomb term
  - Contains exponential-6 Lennard-Jones term
  - Includes some mixed (bond length-bond angle) terms
- MM2
  - Improves MM1
  - Includes more extensive dihedral term
  - Includes more extensive bending term
  - Drops mixed terms
- AMBER (Assisted Model Building and Energy Refinement)
  - Works for nucleic acids and proteins

- Includes distance dependent relative permittivity to include solvent effects
- Hydrogen atoms bonded to carbon are not treated explicitly
- Hydrogen bonds include using a 12-10 Lennard-Jones potential
- OPLS (Optimized Potential for Liquid Simulations)
  - Works for amino acids and proteins
  - CH$_n$ groups are treated as unit atoms centred on carbon
- R.A. Johnson
  - Works for Fe, W, V

One could do molecular statics with the force fields above. Moreover, one can also do molecular dynamics (MD). Accelerations are calculated from the force field on each atom and each atom is propagated according to the accelerations. Note that a force field that is suitable for statics may not necessarily be suitable for dynamics.

6.1 Classical Molecular Dynamics

The atoms in a system follow Newton's law

$$\vec{F}_i = m_i \vec{a}_i \quad , \quad i = 1, \cdots, N \tag{6.3}$$

which is known from a force field, based on the coordinates of all the atoms in the system. Let us solve these equations of motion. We will start from an initial set of coordinates and propagate in time.

There are different time propagations. Consider the <u>leapfrog algorithm</u>.

$$\begin{cases} v_i\left(t + \dfrac{\Delta t}{2}\right) = v_i(t) + \dfrac{dv_i}{dt}\Big|_t \left(\dfrac{\Delta t}{2}\right) + \dfrac{1}{2}\dfrac{d^2 v_i}{dt^2}\Big|_t \left(\dfrac{\Delta t}{2}\right)^2 + \dots \\[2ex] v_i\left(t - \dfrac{\Delta t}{2}\right) = v_i(t) + \dfrac{dv_i}{dt}\Big|_t \left(-\dfrac{\Delta t}{2}\right) + \dfrac{1}{2}\dfrac{d^2 v_i}{dt^2}\Big|_t \left(-\dfrac{\Delta t}{2}\right)^2 + \dots \end{cases}$$

$$\implies v_i\left(t + \frac{\Delta t}{2}\right) = v_i\left(t - \frac{\Delta t}{2}\right) + \underbrace{\frac{dv_i}{dt}\Big|_t}_{a_i(t) = \frac{F_i(t)}{m}}(\Delta t) \quad \text{(neglecting higher orders)}$$

$$\tag{6.4}$$

Similarly

$$r_i(t + \Delta t) = r_i(t) + \underbrace{\frac{dr_i}{dt}\Big|_{t + \frac{\Delta t}{2}}}_{v_i\left(t + \frac{\Delta t}{2}\right)}(\Delta t) \tag{6.5}$$
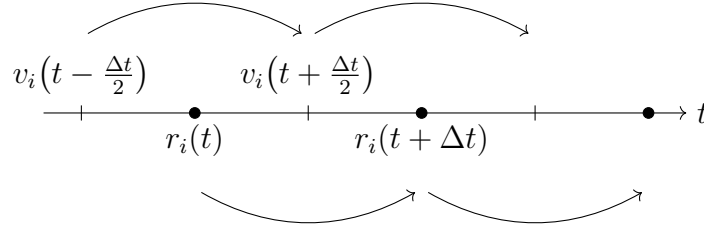
Consider the <u>Verlet algorithm</u>.

$$\begin{cases} r_i(t + \Delta t) = r_i(t) + \left.\dfrac{\mathrm{d}r_i}{\mathrm{d}t}\right|_t (\Delta t) + \dfrac{1}{2}\left.\dfrac{\mathrm{d}^2 r_i}{\mathrm{d}t^2}\right|_t (\Delta t)^2 + \dots \\[2ex] r_i(t - \Delta t) = r_i(t) + \left.\dfrac{\mathrm{d}r_i}{\mathrm{d}t}\right|_t (-\Delta t) + \dfrac{1}{2}\left.\dfrac{\mathrm{d}^2 r_i}{\mathrm{d}t^2}\right|_t (-\Delta t)^2 + \dots \end{cases}$$

$$\implies r_i(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) + \underbrace{\left.\dfrac{\mathrm{d}^2 r_i}{\mathrm{d}t^2}\right|_t}_{a_i(t) = \frac{F_i(t)}{m}} (\Delta t)^2 \quad \textcolor{red}{\text{(neglecting higher orders)}}$$

$$(6.6)$$

If the velocities of the atoms are desired

$$v_i(t) = \frac{r_i(t + \Delta t) - r_i(t - \Delta t)}{2\Delta t} \tag{6.7}$$

Consider the <u>velocity Verlet algorithm</u>.

$$r_i(t + \Delta t) = r_i(t) + \underbrace{\left.\dfrac{\mathrm{d}r_i}{\mathrm{d}t}\right|_t}_{v_i(t)} + \dfrac{1}{2}\underbrace{\left.\dfrac{\mathrm{d}^2 r_i}{\mathrm{d}t^2}\right|_t}_{a_i(t)} (\Delta t)^2$$

$$v_i(t + \Delta t) = v_i(t) + \frac{a_i(t) + a_i(t + \Delta t)}{2}\Delta t$$

$$(6.8)$$

6.2   Physics of Molecular Dynamics Considerations

There are a number of considerations one should account for before performing a molecular dynamics simulation

- Choice of force field (interatomic potential)
- Time propagation algorithm
- Time step (should be smaller than or comparable to the time scale of atomic motion, $\sim$ fs order)
- Number of time steps required

- Boundary conditions (including periodicity provided they are larger than the length scales of importance)
- Key features/parameters that describe the system (fixed or changing) which may include particle number, volume, pressure, temperature
  - May require different statistical ensembles such as nVE (micro-canonical), nVT (canonical), or μVE (grand canonical)

Temperature in the system is treated as

$$
\sum_{i=1}^{N} \frac{1}{2} m_i v_i^2 = ND\left(\frac{1}{2} k_B T\right)
$$

$$
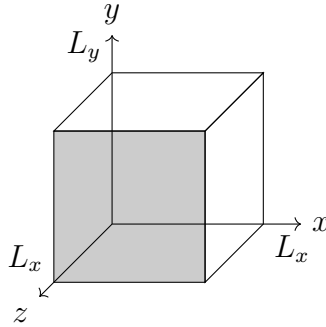\implies T = \frac{1}{NDk_B} \sum_{i=1}^{N} m_i v_i^2
$$

(6.9)

where D is the dimensionality of the problem.

Consider pressure. Starting with the virial theorem

$$
2E_{\text{kin}} = -\sum_{i=1}^{N} \left\langle \vec{r}_i \cdot \vec{F}_i \right\rangle
$$

$$
= -\sum_{i=1}^{N} \left\langle \vec{r}_i \cdot \vec{F}_{i,\text{int}} \right\rangle - \sum_{i=1}^{N} \left\langle \vec{r}_i \cdot \vec{F}_{i,\text{ext}} \right\rangle
$$

(6.10)

For example, for the shaded region

$$
-\sum_{i=1,\text{shaded}}^{N} \left\langle \vec{r}_i \cdot \vec{F}_{i,\text{ext}} \right\rangle = L_x(PL_yL_z)
$$

(6.11)

Accounting for all sides

$$
-\sum_{i=1}^{N} \left\langle \vec{r}_i \cdot \vec{F}_{i,\text{ext}} \right\rangle = L_x(PL_yL_z) + L_y(PL_xL_z) + L_z(PL_xL_y) + 0 + 0 + 0
$$

$$
= DPV
$$

$$(6.12)$$

Thus

$$2\left(ND\frac{1}{2}k_BT\right) = -\sum_{i=1}^{N}\left\langle \vec{r}_i \cdot \vec{F}_{i,\text{int}}\right\rangle + DPV$$

$$\implies PV = Nk_BT + \frac{1}{D}\sum_{i=1}^{N}\left\langle \vec{r}_i \cdot \vec{F}_{i,\text{int}}\right\rangle$$

$$(6.13)$$

6.3   Thermostats and Barostats

One simple thermostat is <u>velocity scaling</u>.

$$v_i^{\text{new}}(t) = \sqrt{\frac{T_{\text{MD}}}{T(t)}}v_i^{\text{old}}(t) \tag{6.14}$$

Another example is the <u>Berendesen thermostat</u>.

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} = \frac{F_i}{m_i} + \frac{1}{2\tau}\left(\frac{T_{\text{MD}}}{T(t)} - 1\right)v_i \tag{6.15}$$

There is also the <u>Andersen thermostat</u>. Knowing that particles should follow a Maxwell-Boltzmann distribution, velocities of a prescribed number of particles are drawn from a Maxwell-Boltzmann distribution corresponding to the desired temperature and the resulting particles are allowed to interact with the system.

Consider <u>Langevin dynamics</u>.

$$m_i\frac{\mathrm{d}v_i}{\mathrm{d}t} = F_i(t) - m_i\Gamma v_i(t) + A_i(t) \tag{6.16}$$

where Γ is the friction coefficient and $A_i(t)$ is random noise that depends on temperature.

One famous example is the <u>Nosé-Hoover thermostat</u>. This uses the velocity Verlet algorithm.

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)\Delta t^2$$

$$v_i(t + \Delta t) = v_i(t) + \frac{a_i(t) + a_i(t + \Delta t)}{2}\Delta t$$

$$= v_i\left(t + \frac{\Delta t}{2}\right) + \frac{a_i(t + \Delta t)}{2}\Delta t$$

$$(6.17)$$

The thermostat follows

$$m_i a_i(t) = F_i(t) - \mu m_i v_i(t) \tag{6.18}$$

where μ evolves following

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \frac{1}{Q}\left[\sum_{i=1}^{N}\frac{1}{2}m_i v_i^2(t) - \frac{3N+1}{2}k_B T_{\mathrm{MD}}\right] \tag{6.19}$$

extra degree of freedom from μ

weight

deviation between actual and expected kinetic energies

Thus

$$\mu\left(t + \frac{\Delta t}{2}\right) = \mu(t) + \frac{1}{Q}\left[\sum_{i=1}^{N}\frac{1}{2}m_i v_i^2(t) - \frac{3N+1}{2}k_B T_{\mathrm{MD}}\right]\frac{\Delta t}{2}$$

$$\mu(t + \Delta t) = \mu(t + \frac{\Delta t}{2}) + \frac{1}{Q}\left[\sum_{i=1}^{N}\frac{1}{2}m_i v_i^2(t) - \frac{3N+1}{2}k_B T_{\mathrm{MD}}\right]\frac{\Delta t}{2} \tag{6.20}$$

Following the velocity Verlet equation and the equation describing the thermostat

$$v_i(t + \Delta t) = v_i\left(t + \frac{\Delta t}{2}\right) + \frac{1}{2m_i}[F_i(t + \Delta t) - \mu(t + \Delta t)m_i v_i(t + \Delta t)]\Delta t$$

$$= \frac{v_i\left(t + \frac{\Delta t}{2}\right) + \left(\frac{\Delta t}{2m_i}\right)F_i(t + \Delta t)}{1 + \frac{\Delta t}{2}\mu(t + \Delta t)} \tag{6.21}$$

Moving to barostats, one simple barostat is volume scaling. Recall the virial theorem relating pressure and volume. If the pressure differs from the desired value as the system evolves, the volume can be adjusted such that the pressure obtains the value required.

Another barostat is the Berendsen barostat. The coordinates are scaled by $\eta^{\frac{1}{3}}$, such that the volume is scaled by η, where

$$\eta(t) = 1 - \frac{\Delta t}{\tau_P}\gamma(P_{\mathrm{MD}} - P(t)) \tag{6.22}$$

with τ_P gives the relaxation time and γ gives the compressibility of the system.

There is also the <u>Andersen barostat</u>, which follows

$$\frac{\mathrm{d}^2[}{\mathrm{d}V^2}t = \frac{P(t) - P_{\mathrm{MD}}}{Q} \quad , \quad \begin{cases} r_i = V^{\frac{1}{3}} s_i \\ v_i = V^{\frac{1}{3}} \frac{\mathrm{d}s_i}{\mathrm{d}t} \end{cases}$$

$$\implies \frac{\mathrm{d}^2 s_i}{\mathrm{d}t^2} = \frac{F_i}{m_i V^{\frac{1}{3}}} - \frac{2}{3} \frac{\frac{\mathrm{d}V}{\mathrm{d}t}}{V} \left(\frac{\mathrm{d}s_i}{\mathrm{d}t} \right) \tag{6.23}$$

where s_i is an auxiliary variable.

Consider finally the <u>Parinello-Rahman barostat</u>. It is an extension of the Andersen barostat, but allows the shape of the simulation box to change.