

An Introduction to Computational Chemistry

Jorge Alarcón Ochoa

July 13, 2016

”The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.”

—Dirac

1 Introduction

1.1 Vectors

Grab your pencil a piece of paper for we will draw a vector! We could start by imagining a Cartesian plane, with an x axis that runs from side to side, and a y axis that runs from the top of the piece of paper to the bottom of the piece of paper. Now, a vector can be drawn by specifying the components along the x and y axis as shown in figure(1.1).

A more ”mathematical” way of saying this is as follows,

$$\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2) = a_1\hat{x} + a_2\hat{y} = \langle a_1, a_2 \rangle \cdot \langle \hat{x}, \hat{y} \rangle$$

First, $\vec{a} = (\mathbf{a}_1, \mathbf{a}_2)$ is what we see on the picture. In this case we say the vector \mathbf{a} is the sum of the vector \mathbf{a}_1 and \mathbf{a}_2 .

$\vec{a} = a_1\hat{x} + a_2\hat{y}$ is a more explicit way of saying the same thing. In this case a_1 and a_2 are just numbers, whereas \hat{x} and \hat{y} specify the direction. \hat{x} and \hat{y} what we call basis vectors. Basis vectors have unit length so when we write $a_1\hat{x}$ we are saying that we want to move a length a_1 along the \hat{x} direction. Notice that \hat{x} and \hat{y} form a right angle, this condition makes these two basis vectors

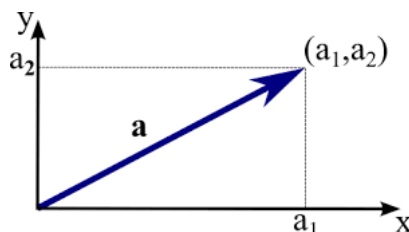


Figure 1: 2 dimensional Cartesian plane

orthonormal. When two vectors are orthogonal what we mean to say is that they form right angles. When we say that a vector is a normal vector, usually, one is saying that the length of the vector is unity, that is one (i.e. \hat{x} and \hat{y}). So orthonormal vectors are normal vectors that are orthogonal to each other. Notice how this is a more fundamental description of a vector, we explicitly write the length (a_1 and a_2) along with the direction (\hat{x} and \hat{y}).

The last step, $\langle a_1, a_2 \rangle \cdot \langle \hat{x}, \hat{y} \rangle$ is a somewhat more sophisticated way of again specifying the vector \mathbf{a} . Here $\langle a_1, a_2 \rangle$ is an object by itself, it needs not be specified in a Cartesian plane, it could equally live in a curvilinear plane or any other type of space. What $\langle a_1, a_2 \rangle \cdot \langle \hat{x}, \hat{y} \rangle$ is trying to say is that have a vector $\langle a_1, a_2 \rangle$ and another vector $\langle \hat{x}, \hat{y} \rangle$ and we take their dot product.

The actual definition of the dot product runs backwards, that is $a_1\hat{x} + a_2\hat{y} = \langle a_1, a_2 \rangle \cdot \langle \hat{x}, \hat{y} \rangle$. We multiply the first component of the first vector with the first component of the second vector and then we add the product of the second component of the first vector with the second component of the second vector and so on for however many components both vectors have.

This notion of vectors, normal vectors, basis vectors, and orthogonality can be expanded for any number of dimensions. Of course when working with more than three dimensions it is not as easy to draw!

1.2 Vectors in Quantum Mechanics

So here is where vectors come in handy. Imagine you go to an ice cream shop. For simplicity, we will only worry about the ice cream flavour and the type of cone. This too could be depicted in a vector space, we could then have two orthogonal directions: one for flavour and the other for the type of cone. We could then assign numerical values to the various options, for example in this made up model, the vector describing an vanilla on a waffle cone could be written as follows,

$$|\text{ice cream cone}\rangle = 3 \times \hat{\text{flavour}} + 5 \times \hat{\text{cone}} = 3\hat{x} + 5\hat{y}$$

Here $\hat{\text{flavour}}$ and $\hat{\text{cone}}$ are orthonormal basis vectors and the numerical values correspond to a definite physical state. We used the so-called bracket notation, where the $|\text{ice cream cone}\rangle$ denotes a vector.

This sounds like a very strange manner of talking about the physical world, nonetheless I promise you will see the value of this abstraction in what follows.

2 Quantum Mechanics

2.1 Time-Dependent Schrödinger's Equation

In order to start with quantum mechanics we need to specify two things: physical states and the equation governing the motion of said states.

We denote a quantum state by the symbol $|\psi\rangle$. The $|\rangle$ notation by itself refers to the fact that we are dealing with a vector, a quantity that has a direction. The Greek letter ψ as usual denotes the state. Going back to the ice cream example ψ could denote the combination of ice cream flavour and the type of cone.

Although, as of right now we have nothing particular in mind for our quantum state. ψ could refer to the position of a particle, whether the particle is standing up or sitting down, whatever we want to describe.

The other detail to discuss is the equation that tells us how quantum states change. This happens to be the famous Schrödinger's equation, which looks like this

$$H |\psi\rangle = i\hbar \frac{\partial |\psi\rangle}{\partial t} \quad (1)$$

One would read equation(1) as follows: the rate of change of the quantum state $|\psi\rangle$ with respect to time is proportional to the product of the Hamiltonian operator acting on the quantum state $|\psi\rangle$.

An operator is just a set of instructions. For example, the Hamiltonian of a single electron can be written as

$$H = -\frac{\hbar^2}{m} \nabla^2 + V \quad (2)$$

where the first term in equation(2) is the kinetic energy and V represents the potential. Both of the terms in equation(2) are also operators, they do something to the quantum state $|\psi\rangle$. For example the first term in equation(2) takes the second derivative of $|\psi\rangle$ and divides it by the mass of the electron.

2.1.1 The Hamiltonian

There is a nice correspondence with classical mechanics that I would like to point out. Let me begin by first stating that the momentum (in one dimension) in quantum mechanics (another operator) is written as

$$p = -i\hbar \frac{\partial}{\partial x}$$

Back to classical mechanics the Energy is given by,

$$E = \frac{1}{2}mv^2 + V = \frac{p^2}{2m} + V$$

If we now use the quantum mechanical momentum, we would have

$$E = -\frac{\hbar^2}{m} \frac{\partial^2}{\partial x^2} + V$$

Which as it turns out looks exactly like the one dimensional Hamiltonian.

2.2 Time-Independent Schrödinger's Equation

There is a trick to solving equation(1),

$$H |\Psi\rangle = -\frac{\hbar^2}{m} \nabla^2 |\Psi\rangle + V |\Psi\rangle = i\hbar \frac{\partial |\Psi\rangle}{\partial t} \quad (3)$$

is by using separation of variable (any quantum mechanics textbook will have the details on this operation, I personally recommend Griffiths.). Assume there is a solution $|\Psi\rangle$ to equation(3) that can be written as $|\Psi(\vec{r}, t)\rangle = |\psi(\vec{r})\rangle |\phi(t)\rangle$, which usually works, as long as the potential is not a function of time.

This procedure gives us

$$|\phi(t)\rangle = e^{-iEt/\hbar}$$

and the time-independent schrödinger equation,

$$H |\psi\rangle = E |\psi\rangle$$

Thus, in order to know how a system behaves we need to figure out how to solve the time-independent schrödinger equation for a given Hamiltonian which has become an eigenvalue problem.

2.2.1 Eigenvalues and Eigenvectors

We saw that the time-independent schrödinger equation is given by

$$H |\psi\rangle = E |\psi\rangle$$

Here H is an operator acting on an eigenvector (or eigenket) $|\psi\rangle$ that equals the energy (eigenvalue) E times the eigenvector $|\psi\rangle$. $|\psi\rangle$ are called eigenkets because when H acts on them the result is a number, eigenvalue, times the same eigenket, this is not always the case with operators.

3 Molecular Mechanics

We saw previously how to construct the Hamiltonian for a one-electron system. The Hamiltonian of a molecule in turn could look something like this,

$$H = -\sum_A \frac{1}{2M_A} \nabla_A^2 - \sum_A \frac{1}{2} \nabla_i^2 + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} - \sum_{Ai} \frac{Z_A}{r_{Ai}} + \sum_{i>j} \frac{1}{r_{ij}} \quad (4)$$

Here i, j refer to electrons while A, B refer to nuclei. $r_{ij} = |\vec{r}_i - \vec{r}_j|$ the vector from the source point j to the field point i .

Notice that the molecular Hamiltonian in equation(4) does not take into account relativistic effects nor spin-orbit effects, it only takes looks at kinetic energy and electrostatic potential energy. The reason equation(4) looks peculiar is because we are using atomic units, as such the kinetic energy becomes

$$\frac{\hbar^2}{2m_e} \nabla^2 \longrightarrow \frac{1}{2} \nabla^2$$

while the electrostatic energy between two electrons ($Z = 1$) looks like

$$\frac{(Ze)(Ze)}{4\pi\epsilon_0} \frac{1}{r_{ij}} \longrightarrow \frac{1}{r_{ij}}$$

3.1 Born-Oppenheimer Approximation

Now that we have a molecular Hamiltonian, let's try and simplify our problem a bit more. Recall that we went from the time-dependent schrödinger equation to the time-independent schrödinger equation by employing separation of variables (time and space).

If we look at equation(4) we could try again to separate $|\Psi\rangle$ into nuclear and electronic parts, $|\Psi(\vec{R}, \vec{r})\rangle = |\Psi_N(\vec{R})\rangle |\Psi_e(\vec{r})\rangle$, but

$$\sum_{Ai} \frac{Z_A}{r_{Ai}}$$

the coupling between electrons and nuclei prevents us from doing so.

Nevertheless we still pursue this separation of variables in what is termed the Born-Oppenheimer Approximation. In practice this is valid due to the fact that nuclei are so much more massive than electrons and as such the electrostatic interaction between nuclei and electrons (the termed that prevents a complete separation) is constant fast enough time scales. For the sake of visualization think of it this way: If you are jumping up and down seldom do you need to take into account the momentum change you are imparting upon the entire planet Earth. For the same reason the much more lighter (about 2,000 times) than the nuclei.

Now that we have separated the nuclear from the electronic degrees of freedom we might as well fix the nuclei, \vec{R} , at a given position and solve for $|\Psi_e(\vec{r})\rangle$.

To do this, notice that we can further simplify our problem by rewriting the molecular Hamiltonian as follows,

$$H = T_{NN}(\vec{R}) + T_{ee}(\vec{r}) + V_{NN}(\vec{R}) + V_{Ne}(\vec{R}; \vec{r}) + V_{ee}(\vec{r})$$

e refers to electrons and N to nuclei.

First of all, the nuclear kinetic energy, T_{NN} , is a lot smaller than the electronic kinetic energy, T_{ee} , by the ratio of the electron's mass to the nuclear mass m_e/M_A . Also, V_{NN} is just a constant value, it will only shift the eigenvalues, energies by a constant amount. So for a fixed nuclear configuration, we obtain the electronic Hamiltonian,

$$H_e = T_{ee}(\vec{r}) + V_{Ne}(\vec{R}; \vec{r}) + V_{ee}(\vec{r})$$

so that our equation to solve becomes

$$H_e |\psi_e\rangle = E_e |\psi_e\rangle$$

3.2 Orbitals

In order to simplify our eigenvalue problem we will expand $|\psi_e\rangle$, our electronic wavefunction, in terms of one-electron functions, the so called orbitals. A useful kind of orbital is a spin orbitals, which take into account the spin of a given electron. We can then express the orbital $|\chi\rangle$ as,

$$|\chi\rangle = |\psi(\vec{r})\rangle |s\rangle$$

Notice that spatial orbitals are doubly occupied, a given spatial orbital $|\psi(\vec{r})\rangle$ can either have a spin up or a spin down.

Later on we will talk about Unrestricted Hartree-Fock methods which utilize spatial orbitals that are singly occupied, only one spatial orbital for a given spin.

3.3 Electronic Wavefunction

Because we are solely interested in chemistry our focus will be mostly upon fermions for which the total wavefunction $|\psi_e\rangle$ must be antisymmetric.

$$|\Psi_e(1, \dots, i, j, \dots, n)\rangle = -|\Psi_e(1, \dots, j, i, \dots, n)\rangle$$

Each index in the wavefunction represents an electron with spatial coordinates and spin.

The way of obtaining antisymmetry is by writing the electronic wave function as a Slater determinant of orbitals:

$$|\psi_i\rangle = |\chi_1\chi_2\chi_3\dots\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\chi_1(1)\rangle & |\chi_2(1)\rangle & \cdots & |\chi_N(1)\rangle \\ |\chi_1(2)\rangle & |\chi_2(2)\rangle & \cdots & |\chi_N(2)\rangle \\ \vdots & \vdots & \ddots & \vdots \\ |\chi_1(N)\rangle & |\chi_2(N)\rangle & \cdots & |\chi_N(N)\rangle \end{vmatrix}$$

The number of orbitals in a Slater determinant equals the number of electrons. Furthermore, any N-electron wave function can be expressed exactly as a linear combination of all possible N-electron Slater determinants formed from a complete set of spin orbitals.

We can in turn express quantities in terms of Slater determinants,

$$|\Psi\rangle = \sum_i c_i |\psi_i\rangle$$

where the sum runs over all the possible N-electron Slater determinants, which will be infinite if we have a complete set of spin orbitals $|\chi\rangle$.

4 The Hartree-Fock Method

Summarizing the procedure for doing quantum mechanical calculations is as follows:

1. Invoke the Born-Oppenheimer approximation.
2. Solve the Electronic Schrödinger equation.
3. Solve for Nuclear motion (via quantum or classical mechanics).

This just begs the question, how do we actually solve the electronic Schrödinger equation. The answer to this is the base of the Hartree-Fock method which is by assuming the electronic wave function can be approximated by a product of orbitals. This takes us back to the use of Slater determinants as a basis.

4.1 The Variational Method

So we build an initial electronic wave function by a product of orbitals. We will vary the orbitals until the electronic energy reaches a minimum.

4.2 The HF Method

Adding the two last steps we obtain the Hartree-Fock method

1. Invoke the Born-Oppenheimer approximation.
2. Solve the Electronic Schrödinger equation.
 - (a) Express the electronic wave function as a single Slater determinant.
 - (b) Use the variational method to find the orbitals that minimize the electronic energy.
3. Solve for Nuclear motion (via quantum or classical mechanics).

4.3 Orbitals in HF

In the same manner that one can write the exact wave function as a linear combination of orbitals to simplify calculations, we will use orbitals that are linear combinations of basis functions, $|\xi\rangle$,

$$|\chi_i\rangle = \sum c_j |\tilde{x}\rangle$$

This is commonly referred to as the Linear Combination of Atomic Orbitals Molecular Orbital (LCAO-MO) theory

There are two types of basis functions. Normally physicists use plane-waves (periodic systems) whereas chemists tend to use atom centred Gaussian functions (molecules).

4.4 Problems with HF

The HF method is what is called a mean-field approximation, that is, each electron interacts only with the average electron cloud of the other electrons. This is due to the fact that we use a single Slater determinant for our wave function. To solve this one needs to refer to density functional theory and other methods that take into correlation between electrons.

Other problems that come into play are due to the finite size of the basis sets we utilized as orbitals.

In mind, HF is usually a poor choice for excited states, bond breaking, transition metals with unfilled d and f shells, and diradicals.

4.4.1 Some Terminology

Restricted HF Method: The system of study has a closed-shell with all orbitals doubly occupied.

Restricted Open-Shell HF Method: The system of study has an open-shell with all orbitals doubly occupied.

Unrestricted HF Method: As mentioned previously, for this method all orbitals are singly occupied.

5 Computation

5.1 More Notation

Like always we hope to simplify our problems as much as possible by introducing notation. We start by defining the one electron operator

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{R_{iA}}$$

and the two electron operator

$$v(ij) = \frac{1}{r_{ij}}$$

Using this new notation we can rewrite our electronic Hamiltonian as,

$$H_e = \sum_i h(i) + \sum_{i<j} v(i, j)$$

and the energy becomes

$$\begin{aligned} E &= \sum_i \langle i | h | i \rangle + \frac{1}{2} \sum_{ij} \langle ij | ij \rangle \\ &= \sum_i \langle i | h | i \rangle + \frac{1}{2} \sum_{ij} (\langle ij | ij \rangle - \langle ij | ji \rangle) \\ &= \sum_i \langle i | h | i \rangle + \frac{1}{2} \sum_{ij} ([ii | jj] - [ij | ji]) \end{aligned} \tag{5}$$

using chemist notation, where

$$\langle ik | jl \rangle = [ij | kl] = \int d\vec{x}_1 d\vec{x}_2 \chi_i^*(\vec{x}_1) \chi_j(\vec{x}_1) \left(\frac{1}{r_{12}} \right) \chi_k^*(\vec{x}_2) \chi_l(\vec{x}_2)$$

and

$$\langle i | h | i \rangle = \int d\vec{x}_1 \chi_i^*(\vec{x}_1) h(\vec{x}_1) \chi_i(\vec{x}_1)$$

5.2 Hartree-Fock Equations

All we have to do now is use the variational principle. We look for the orbitals that will minimize the energy while keeping the orbitals orthonormal. For this purpose we utilize the method of Lagrange multipliers by defining the quantity

$$\mathcal{L}[\{i\}] = E_{hf}[\{i\}] - \sum_{ij} \epsilon_{ij} (\langle i | j \rangle - \delta_{ij})$$

where the overlap integral is defined by

$$\langle i | j \rangle = \int d\vec{x} \chi_i^*(\vec{x}) \chi_j(\vec{x})$$

$$h(\vec{x}_1)\chi_i(\vec{x}_1) + \sum_{j \neq i} \left(\int d\vec{x}_2 |\chi_j(\vec{x}_2) \frac{1}{r_{12}}|^2 \right) \chi_i(\vec{x}_1) - \sum_{j \neq i} \left(\int d\vec{x}_2 \chi_j^*(\vec{x}_2) \chi_i(\vec{x}_2) \frac{1}{r_{12}} \right) \chi_j(\vec{x}_1) = \epsilon_{ij} \chi_j(\vec{x}_1) \quad (6)$$

Defining the Coulomb operator \mathcal{J}_i and the exchange operator \mathcal{K}_i

$$\left(h(\vec{x}_1) + \sum_{j \neq i} [\mathcal{J}_i(\vec{x}_1) - \mathcal{K}_i(\vec{x}_1)] \right) \chi_j(\vec{x}_1) = \epsilon_{ij} \chi_j(\vec{x}_1) \quad (7)$$

These operators are called as such because of their functional forms. \mathcal{J}_i corresponds to the average Coulomb interaction between the orbital χ_i and all other electrons

$$\mathcal{J}_i(\vec{x}_i) = \int d\vec{x}_2 |\chi_j(\vec{x}_2) \frac{1}{r_{12}}|^2$$

On the other hand, the exchange operator comes from the antisymmetry requirement of the total wave function. The name is due to the fact that it looks like the Coulomb operator if you exchange the orbital χ_i for χ_j .

$$\mathcal{K}_i(\vec{x}_1) \chi_j(\vec{x}_1) = \left(\int d\vec{x}_2 \chi_j^*(\vec{x}_2) \chi_i(\vec{x}_2) \frac{1}{r_{12}} \right) \chi_j(\vec{x}_1)$$

We can further simplify equation(7) by defining the Fock operator

$$f(\vec{x}_1) = h(\vec{x}_1) + \sum_{j \neq i} [\mathcal{J}_i(\vec{x}_1) - \mathcal{K}_i(\vec{x}_1)]$$

so equation(7) now can be written as follows,

$$f(\vec{x}_1) \chi_j(\vec{x}_1) = \epsilon_{ij} \chi_j(\vec{x}_1)$$

Which is an eigenvalue problem!

The HF equation can either be solved exactly (exact HF), or by introducing a set of basis functions (Hartree-Fock-Roothan equations). In the case of HFR equations, the solutions depend on the orbitals used. This is the reason that the HFR method is called a self-consistent field (scf) approach.

We will then proceed with the HFR method. First of all we denote the atomic orbital basis functions by $\tilde{\chi}$, so that

$$\chi_j = \sum_{\nu}^K C_{\nu j} \tilde{\chi}_{\nu}$$

Hence, for each orbital i we have

$$f(\vec{x}_1) \sum_{\nu} C_{\nu i} \tilde{\chi}_{\nu}(\vec{x}_1) = \epsilon_i \sum_{\nu} C_{\nu i} \tilde{\chi}_{\nu}(\vec{x}_1)$$

Taking this previous expression, multiplying by $\tilde{\chi}_{\mu}^*(\vec{x}_1)$ and integrating with respect to \vec{x}_1 , we have

$$\sum_{\nu} C_{\nu i} \int d\vec{x}_1 \tilde{\chi}_{\mu}^*(\vec{x}_1) f(\vec{x}_1) \tilde{\chi}_{\nu}(\vec{x}_1) = \epsilon_i \sum_{\nu} C_{\nu i} \int d\vec{x}_1 \tilde{\chi}_{\mu}^*(\vec{x}_1) \tilde{\chi}_{\nu}(\vec{x}_1)$$

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i}$$

where, F is called the Fock matrix

$$F_{\mu\nu} = \int d\vec{x}_1 \tilde{\chi}_{\mu}^*(\vec{x}_1) f(\vec{x}_1) \tilde{\chi}_{\nu}(\vec{x}_1)$$

and S is the overlap matrix

$$S_{\mu\nu} = \int d\vec{x}_1 \tilde{\chi}_{\mu}^*(\vec{x}_1) \tilde{\chi}_{\nu}(\vec{x}_1)$$

In vector notation we have,

$$\mathbf{F}\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\epsilon} \quad (8)$$

Here $\boldsymbol{\epsilon}$ is a diagonal matrix containing the orbital energies ϵ_i . Notice that equation(8) is almost an eigenvalue problem except for the overlap matrix \mathbf{S} . In order to deal with this problem we apply a basis transformation to go from an orthogonal basis where we can get rid of \mathbf{S} .

5.3 Derivation of the Hartree-Fock Equations

In the previous section we presented equation(6) without derivation, here we will derive them. Again, we use the method of Lagrange multipliers to minimize the electronic energy with respect to the orbitals while keeping these orthonormal,

$$\mathcal{L}[\{i\}] = E_{hf}[\{i\}] - \sum_{ij} \epsilon_{ij} (\langle i|j\rangle - \delta_{ij})$$

We are interested then in calculating

$$\delta\mathcal{L}_i[\{i\}] = 0$$

We want to find a minimum for our Lagrangian. We will do this by linearly varying the orbitals χ_i , that is $\chi_i \rightarrow \chi_i + \delta\chi_i$, while keeping only terms of first order in $\delta\chi_i$.

From here on we will change notation, the orbital χ_i will solely be denoted by i . So let's go term by term.

First of, looking at the constraint $\epsilon_{ij} (\langle i|j\rangle - \delta_{ij})$ the only term that will be affected upon $i \rightarrow i + \delta i$ is $\langle i|j\rangle$. We will only keep terms linear on δi , so

$$\langle i|j\rangle \rightarrow \langle \delta i|j\rangle + \langle i|\delta j\rangle \quad (9)$$

We ignored the terms that look like $\langle \delta i|\delta j\rangle$ because this is not linear on δi and $\langle i|j\rangle$ because we are only interested in the change.

This term was simple, now let's look at $E_{hf}[i]$. from equation(5), one can clearly see that

$$\begin{aligned} \delta E = \sum_i (\langle \delta i|h|i\rangle + \langle i|h|\delta i\rangle) + \frac{1}{2} \sum_{ij} ([\delta i i|j j] + [i \delta i|j j] + [i i|\delta j j] + [i i|j \delta j]) \\ - \frac{1}{2} \sum_{ij} ([\delta i j|j i] + [i \delta j|j i] + [i j|\delta j i] + [i j|j \delta i]) \end{aligned} \quad (10)$$

At this point is where the ricks come in. look at equations (10) and (9), there is a symmetry between the terms. Let's make it clearer by playing around with equation(9), this time we will include the summation and the Lagrange multipliers, so we have,

$$\sum_{ij} \epsilon_{ij} (\langle \delta i | j \rangle + \langle i | \delta j \rangle)$$

Let's begin with the second term,

$$\sum_{ij} \epsilon_{ij} \langle i | \delta j \rangle = \left[\sum_{ij} \epsilon_{ij}^* \langle i | \delta j \rangle^* \right]^*$$

surely taking a double complex conjugate doesn't hurt, it gives us what we started with. Now, because we are summing over i and j it is our right to change the names of the dummy variables if we so desire. So we will change i to j and j to i ,

$$\left[\sum_{ij} \epsilon_{ij}^* \langle i | \delta j \rangle^* \right]^* = \left[\sum_{ij} \epsilon_{ji}^* \langle j | \delta i \rangle^* \right]^*$$

There is another thing that comes in handy, turns out that ϵ is a hermitian matrix, and so

$$\left[\sum_{ij} \epsilon_{ij}^* \langle i | \delta j \rangle^* \right]^* = \left[\sum_{ij} \epsilon_{ij} \langle j | \delta i \rangle^* \right]^* = \left[\sum_{ij} \epsilon_{ij} \langle \delta i | j \rangle \right]^*$$

So after all that work we ended up with a term that is the complex conjugate of the first term in $\sum_{ij} \epsilon_{ij} (\langle \delta i | j \rangle + \langle i | \delta j \rangle)$

From equation(10), we can see that we have a similar set up in $\sum_i (\langle \delta i | h | i \rangle + \langle i | h | \delta i \rangle)$, for h is a hermitian matrix and so $\langle i | h | \delta i \rangle$ is the complex conjugate of $\langle \delta i | h | i \rangle$.

As it turns out, the same thing happens for the rest of the terms in equation(10). You can check this yourself by noting that because

$$[\delta i j | k l] = \int d\vec{x}_1 d\vec{x}_2 \delta i^*(\vec{x}_1) j(\vec{x}_1) \left(\frac{1}{r_{12}} \right) k^*(\vec{x}_2) l(\vec{x}_2)$$

the complex conjugate would be,

$$\int d\vec{x}_1 d\vec{x}_2 j^*(\vec{x}_1) \delta i(\vec{x}_1) \left(\frac{1}{r_{12}} \right) l^*(\vec{x}_2) k(\vec{x}_2) = [j \delta i | l k]$$

and so from equation(10), the two electron terms are

$$\frac{1}{2} \sum_{ij} ([\delta i i | j j] + [i \delta i | j j] + [i i | \delta j j] + [i i | j \delta j]) \quad (11)$$

and

$$\frac{1}{2} \sum_{ij} ([\delta i j | j i] + [i \delta j | j i] + [i j | \delta j i] + [i j | j \delta i]) \quad (12)$$

Before we begin simplifying these two equations recall that chemist notation has an eight-fold symmetry. For example, if we have $[ij|kl]$, this can be written as

$$[ij|kl] = [kl|ij] = [kl|ji] = [ji|kl] = [ij|lk] = [ji|lk] = [lk|ji] = [lk|ij]$$

Starting with the first term of equation(11),

$$[\delta ii|jj]^* \rightarrow [i\delta i|jj]$$

using $[ij|kl] = [kl|ij]$,

$$[\delta ii|jj]^* \rightarrow [i\delta i|jj] \rightarrow [jj|i\delta i]$$

and because we are free to rename dummy indices due to the fact that we are summing with respect to i and j ,

$$[\delta ii|jj]^* \rightarrow [i\delta i|jj] \rightarrow [jj|i\delta i] \rightarrow [ii|j\delta j]$$

So we started with the complex conjugate of the first term in equation(11) and ended with the fourth term of the same equation. You can do the same for the second term, it will turn out to be the complex conjugate of the third term

$$[i\delta i|jj]^* \rightarrow [\delta ii|jj] \rightarrow [jj|\delta ii] \rightarrow [ii|\delta jj]$$

Rewriting equation(11),

$$\frac{1}{2} \sum_{ij} ([\delta ii|jj] + [i\delta i|jj]) + c.c. \quad (13)$$

using $[ij|kl] = [ji|kl]$, we get

$$\sum_{ij} [\delta ii|jj] + c.c. \quad (14)$$

Doing the same for equation(12), beginning with $[\delta ij|ji]$,

$$[\delta ij|ji] \rightarrow [j\delta i|ij] \rightarrow [ij|j\delta i]$$

and

$$[i\delta j|ji] \rightarrow [\delta ji|ij] \rightarrow [ij|\delta ji]$$

this time no remaining. We can now rewrite equation(12)

$$\frac{1}{2} \sum_{ij} ([\delta ij|ji] + [i\delta j|ji]) + c.c.$$

Ladies and gentleman, let me proceed with yet another trick. We will now proceed to rename indices and use the symmetry relation $[ik|kl] = [ji|lk]$ on the second term of the previous equation,

$$[i\delta j|ji] \rightarrow [j\delta i|ij] \rightarrow [\delta ij|ji]$$

And thus we obtain

$$\sum_{ij} [\delta ij|ji] + c.c. \quad (15)$$

Putting everything together,

$$\delta\mathcal{L} = \sum_i \langle \delta i | h | i \rangle + \sum_{ij} ([\delta ii | jj] - [\delta ij | ji]) - \sum_{ij} \epsilon_{ij} \langle \delta i | j \rangle + c.c. = 0 \quad (16)$$

Writing equation(16) out

$$\begin{aligned} \delta\mathcal{L} = & \sum_i \int d\vec{x}_1 \delta\chi_i^*(\vec{x}_1) h\chi_i(\vec{x}_1) \\ & + \sum_{ij} \left(\int d\vec{x}_1 d\vec{x}_2 \delta\chi_i^*(\vec{x}_1) \chi_i(\vec{x}_1) \left(\frac{1}{r_{12}} \right) \chi_j^*(\vec{x}_2) \chi_j(\vec{x}_2) \right. \\ & \quad \left. - \int d\vec{x}_1 d\vec{x}_2 \delta\chi_i^*(\vec{x}_1) \chi_j(\vec{x}_1) \left(\frac{1}{r_{12}} \right) \chi_j^*(\vec{x}_2) \chi_i(\vec{x}_2) \right) \\ & - \sum_{ij} \epsilon_{ij} \int d\vec{x}_1 \delta\chi_i^*(\vec{x}_1) \chi_j(\vec{x}_1) + c.c. \end{aligned} \quad (17)$$

or

$$\begin{aligned} \delta\mathcal{L} = & 0 \\ = & \sum_i \int d\vec{x}_1 \delta\chi_i^*(\vec{x}_1) \left[h\chi_i(\vec{x}_1) + \sum_j \left(\chi_i(\vec{x}_1) \int d\vec{x}_2 \left(\frac{1}{r_{12}} \right) \chi_j^*(\vec{x}_2) \chi_j(\vec{x}_2) \right. \right. \\ & \quad \left. \left. - \chi_j(\vec{x}_1) \int d\vec{x}_2 \left(\frac{1}{r_{12}} \right) \chi_j^*(\vec{x}_2) \chi_i(\vec{x}_2) \right) - \sum_j \epsilon_{ij} \chi_j(\vec{x}_1) \right] + c.c. \end{aligned} \quad (18)$$

5.4 Solving the Hartree-Fock-Roothan Equations

5.4.1 Input

Recall that the problem we wanted to solve was

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (19)$$

which can be written as

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i} \quad (20)$$

Here S is the overlap matrix

$$S_{\mu\nu} = \int d\vec{x}_1 \chi_{\mu}^*(\vec{x}_1) \chi_{\nu}(\vec{x}_1)$$

F is the Fock matrix

$$F_{\mu\nu} = \int d\vec{x}_1 \chi_{\mu}^*(\vec{x}_1) f(\vec{x}_1) \chi_{\nu}(\vec{x}_1)$$

and the fock operator is,

$$f(\vec{x}_1) \chi_j(\vec{x}_1) = \epsilon_i \chi_j(\vec{x}_1)$$

$$= h(\vec{x}_1)\chi_i(\vec{x}_1) + \sum \left(\chi_i(\vec{x}_1) \int d\vec{x}_2 \left| \chi_j(\vec{x}_2) \frac{1}{r_{12}} \right| - \chi_j(\vec{x}_1) \int d\vec{x}_2 \chi_j^*(\vec{x}_2) \chi_i(\vec{x}_2) \frac{1}{r_{12}} \right)$$

In order to obtain these we will read them from text files. The overlap matrix can be found in 's.dat', the nuclear repulsion integrals in 'enuc.dat'. The rest of the information needed to construct the matrix \mathbf{F} can be obtained from the kinetic energy integrals 't.dat', the potential energy integrals 'v.dat', and the two electron integrals 'eri.dat'.

Calculating these quantities will be the subject of a later post.

5.4.2 Orthogonalization

Our first objective is to transform equation(19) into an eigenvalue problem, so we need to find a transformation that will do just that.

We need to diagonalize the overlap matrix \mathbf{S} . We do this by solving

$$SL_S = \Lambda_S L_S$$

Λ_S is a diagonal matrix of eigenvalues and L_S is the corresponding matrix of eigenvectors. The eigenvector matrix should obey the condition

$$L_S L_S^T = \mathbf{1}$$

Once this is done, we can build the $S^{-1/2}$ matrix, defined as

$$S^{-1/2} = L_S \Lambda^{-1/2} L_S^T$$

From equation(19) assume that there exists a new matrix \mathbf{C} , such that

$$C' = X^{-1}C$$

and

$$C = XC'$$

so that equation(19) becomes

$$FXC' = SXC' \epsilon$$

If we multiply both sides by X^{-1} and define

$$F' \equiv X^{-1}FX$$

then we can write equation(19) as

$$F'C' = C' \epsilon$$

5.4.3 Self-Consistent Field Iteration

- Begin by calculating the Fock matrix:

$$\mathbf{F} = \text{func.makefock}(\mathbf{H.Hcore}, \mathbf{D}, \mathbf{H.twoe})$$

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda\sigma} D_{\lambda\sigma} (2(\mu\nu|\lambda\sigma) - (\mu\lambda|\nu\sigma))$$

- Transform the Fock matrix into the orthonormal AO basis:

```
Fprime = func.fprime(S_half, F)
```

$$F' = \tilde{S}^{-1/2} F S^{-1/2}$$

$\tilde{S}^{-1/2}$ is the transpose of matrix $S^{-1/2}$.

- Diagonalize the transformed Fock matrix:

```
E, Cprime = np.linalg.eigh(Fprime)
```

$$F' C' = C' \epsilon$$

- Transform the resulting eigenvectors into the original (non-orthogonal) basis:

```
C = np.dot(S_half, Cprime)
```

$$C = S^{-1/2} C'$$

- Construct the new density matrix from C :

```
D, OLDD = func.makedensity(C, D, H.Nelec)
```

$$D_{\mu\nu} = \sum_i^{occ} C_{\mu}^i C_{\nu}^i$$

where i indexes the columns of C and the summation includes the occupied spatial molecular orbitals.

- Compute the new Electronic and total energies:

```
EN = func.currentenergy(D, H.Hcore, F)
OUTPUT.write("\nTOTAL E(SCF) = {}".format(EN + H.Enuc))
```

$$E_e = \sum_{\mu\nu}^{AO} D_{\mu\nu} (H_{\mu\nu} + F_{\mu\nu})$$

- Test convergence of the density matrix

```
delta = func.deltap(D, OLDD)
deltaE = (EN - oldEN)
```

$$rms_D = \left[\sum_{\mu\nu}^{AO} (D_{\mu\nu}^i - D_{\mu\nu}^{i-1}) \right]^{-1/2} < \delta_1$$

$$\Delta E = E_e^i - E_e^{i-1} < \delta_2$$

i denotes the iteration.