An Introduction to Computational Chemistry

Jorge Alarcón Ochoa

July 14, 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

In order to do quantum mechanics there is a very basic principle underlying the mathematical description of the subject which relies on the decomposition of quantities into a given basis.

The main point from this section is this: vectors can be represented in different basis. Vectors are things, ideas and choosing a different basis for their representation is solely a matter of what language you want to speak.

1.1 Vectors

Grab your pencil a piece of paper, we will draw a vector! We could start by imagining a Cartesian plane. In this cartesian plane there will be an x axis that runs from side to side, and a y axis that runs from the top of the paper to the bottom of the paper. Now, a vector can be drawn by specifying component vectors, each lying along one of the two axis, as shown in figure(1.1).

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

$$\boldsymbol{a} = (\boldsymbol{a_1}, \boldsymbol{a_2}) = a_1\hat{x} + a_2\hat{y} = <a_1, a_2> \cdot <\hat{x}, \hat{y}>$$
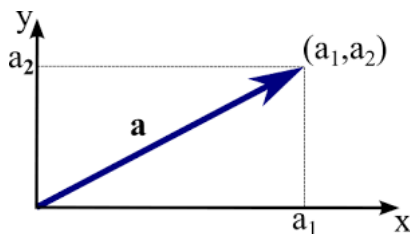
Figure 1: **2 dimensional Cartesian plane**

First, $\vec{a} = (\boldsymbol{a_1}, \boldsymbol{a_2})$ is what we see on the picture. In this case we say the vector $\boldsymbol{a}$ is the sum of the vector $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$.

$\vec{a} = a_1\hat{x} + a_2\hat{y}$ 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 (are normal vectors)so when we write $a_1\hat{x}$ we are saying that we want to move a_1 unit steps along the $\hat{x}$ direction. Notice that $\hat{x}$ and $\hat{y}$ form a right angle, this condition makes these two basis vectors orthonormal. 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 expression, $< a_1, a_2 > \cdot < \hat{x}, \hat{y} >$, is a somewhat more sophisticated way of again specifying the vector $\boldsymbol{a}$. Here $< a_1, a_2 >$ 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 $< a_1, a_2 > \cdot < \hat{x}, \hat{y} >$ is trying to say is that have a vector $< a_1, a_2 >$ and another vector $< \hat{x}, \hat{y} >$ and by taking the dot product between these two we are projecting the first onto a Cartesian plane.

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 visualize!

1.2 Vectors in Quantum Mechanics

So here is where vectors come in handy. Imagine you go to an ice cream shop. For simplicity, imagine you are going to a shop where you can only chose one ice cream flavour and the type of cone. This too could be depicted in a vector space by using 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 represented by

$$|ice\,cream\,cone\rangle = 3 \times fla\hat{v}our + 5 \times c\hat{o}ne = 3\hat{x} + 5\hat{y}.$$

Here $fla\hat{v}our$ and $c\hat{o}ne$ are orthonormal basis vectors and the numerical values correspond to a definite physical state. We used the so-called braket notation, where the $|ice\,cream\,cone\rangle$ denotes a vector.

2 Quantum Mechanics

2.1 Time-Dependent Schrödinger's Equation

In order to start with quantum mechanics we need to specify two thing: 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.

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

$$H \ket{\psi} = i\hbar \frac{\partial \ket{\psi}}{\partial t}. \tag{1}$$

So the way in which our quantum state, $\bra{\psi}$, changes in time is governed by the Hamiltonian operator. It just so happens that the Hamiltonian corresponds to the energy of our state.

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

$$H = -\frac{\hbar^2}{m} \nabla^2 + V \tag{2}$$

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 $\ket{\psi}$. For example the first term in equation(2) takes the second derivative of $\ket{\psi}$ 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),

$$H \ket{\Psi} = -\frac{\hbar^2}{m} \nabla^2 \ket{\Psi} + V \ket{\Psi} = i\hbar \frac{\partial \ket{\Psi}}{\partial t}. \tag{3}$$

The trick is to use separation of variable (any quantum mechanics textbook will have the details on this operation, I personally recommend Griffiths). Assume there is a solution $\ket{\Psi}$ to equation(3) that can be written as $\ket{\Psi(\vec{r}, t)} = \ket{\psi(\vec{r})}\ket{\phi(t)}$. Separation of variables 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.

This sort of problems are called eigenvalue problems. This is because when the operator H acts on the state $\langle\psi|$ it returns a number (E is the energy) times the same state $\langle\psi|$. This is not always the case in quantum mechanics.

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) is does not take into account relativistic effects nor spin-orbit effects, it only takes into account the electrostatic potential energy. If you are looking at this equation thinking it isn't right it is because we are using atomic units. For example, the kinetic energy becomes

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

while the electrostatic energy between two electrons ($Z = 1$), in atomic units, 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).

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 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 for fast enough length 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 the solve for $|\Psi_e(\vec{r})\rangle$.

First off we want to see how we can separate the molecular Hamiltonian into nuclear and electronic parts. The Hamiltonian (equation(4)) can also be written as

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

where 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,

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

so the equation we need to solve is

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

Nuclear dynamics can be solved for in a similar manner or one may even use classical mechanics in some cases.

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. Orbitals are either spatial orbitals (electrons with no spin) or spin orbitals (electrons with spin). We can then express the orbital $|\chi\rangle$ as,

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

$\langle s \rangle$ being the spin of the electron.

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

We now need to apply another condition to our wave function, this the condition of antisymmetry. Fermions need to have an antisymmetric wave function $|\Psi_e\rangle$,

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

each index in the wave function represents an electron with spatial coordinates and spin.

The way of obtaining antisymmetry is by writing the electronic wave function as a slater determinants of orbitals (an orbital per column and an electron per row):

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

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.

What we gain is that we can express our solution in terms of slater determinants,

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

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 our steps towards a solution of the electronic wave functions,

1. Invoke the Born-Oppenheimer approximation.

2. Solve the Electronic Schrödinger equation.

3. Solve the for Nuclear motion (via quantum or classical mechanics).

Although this begs the question, how do we actually solve the electronic Schrödinger equation. An answer to this is by using the Hartree-Fock method which is the subject of this post.

4.1 The HF Method

Actually there are a couple things that we need to add to our previous algorithm, these assumptions are the basis of the HF procedure,

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 the for Nuclear motion (via quantum or classical mechanics).

4.2 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, $|\tilde{x}\rangle$ (i.e. 6-31G*),

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

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 physicist use plane-waves (periodic systems) whereas chemist tend to use atom centred Gaussian functions (molecules).

The rest of the HF method relies on solving the HF equations in order to solve the electronic eigenvalue problem.

5 Computation

In this section I will describe how to carry out the Hartree-Fock-Roothan procedure using Python 3. The code can be found at:

`https://github.com/alejandrox1/blog/tree/master/HF/code`.

5.1 More Notation

Let's 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,

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

i and j denoting electrons.

and the energy becomes

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

We will use the 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

$$\langle i|\,h\,|i\rangle = \int d\vec{x_1}\, \chi_i^*(\vec{x_1})h(\vec{r_1})\chi_j(\vec{x_1}).$$

5.2   Hartree-Fock Equations

The next step is to 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

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

where the overlap integral is defined by

$$\langle i|j\rangle = \int d\vec{x}\, \chi_i^*(\vec{x_1})\chi_j(\vec{x_1})$$

This procedure will give us the Hartree-Fock equations (which are derived in the next section)

$$\begin{aligned}
h(\vec{x_1})\chi_i(\vec{x_1}) &+ \sum_{j\,\neq i}\left(\int d\vec{x_2}\, |\chi_j(\vec{x_2})\frac{1}{r_{12}}|^2\right)\chi_i(\vec{x_1}) \\
&- \sum_{j\,\neq i}\left(\int d\vec{x_2}\, \chi_j^*(\vec{x_2})\chi_i(\vec{x_2})\frac{1}{r_{12}}\right)\chi_j(\vec{x_1}) = \epsilon_{ij}\chi_j(\vec{x_1}).
\end{aligned} \tag{6}$$

Now define the Coulomb operator $\mathcal{J}_i$ and the exchange operator $\mathcal{K}_i$, so that equation(6) becomes

$$\left(h(\vec{x_1}) + \sum_{j\neq i}[\mathcal{J}_i(\vec{x_1}) - \mathcal{K}_i(\vec{x_1})]\right)\chi_j(\vec{x_1}) = \epsilon_{ij}\chi_j(\vec{x_1}). \tag{7}$$

$\mathcal{J}_i$ 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_i.

$$\mathcal{K}_i(\vec{x_1})\chi_j(\vec{x_1}) = \left(\int d\vec{x_2}\, \chi_j^*(\vec{x_2})\chi_i(\vec{x_2})\frac{1}{r_{12}}\right)\chi_j(\vec{x_1})$$

We can further simply 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,

$$f(\vec{x_1})\chi_j(\vec{x_1}) = \epsilon_i\chi_j(\vec{x_1})$$

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, denoting the atomic orbital basis functions by $\tilde{\chi}$, so that

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

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 called 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,

$$\boldsymbol{FC = SC\epsilon} \tag{8}$$

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 to 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, so we will derive it now. Again, we use the method of Lagrange multipliers to minimize the electronic energy with respect to the orbitals while keeping these orthonormal,

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

We want to minimize the Lagrangian by finding a set of orbitals such that

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

We will linearly vary the orbitals χ_i, that is, $\chi_i = \chi_i \to \delta\chi_i$, 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 off, looking at the constraint $\epsilon_{ij} (\langle i|j\rangle - \delta_{ij})$ the only term that will be affected upon $i \to i + \delta i$ is $\langle i|j\rangle$. We will only keep terms linear on δi, so

$$\langle i|j\rangle \to \langle \delta i|j\rangle + \langle i|\delta j\rangle \tag{9}$$

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

$$\delta E = \sum_i \left(\langle \delta i|h|i\rangle + \langle i|h|\delta i\rangle \right) + \frac{1}{2} \sum_{ij} \left([\delta ii|jj] + [i\delta i|jj] + [ii|\delta jj] + [ii|j\delta j] \right)$$

$$- \frac{1}{2} \sum_{ij} \left([\delta ij|ji] + [i\delta j|ji] + [ij|\delta ji] + [ij|j\delta i] \right) \tag{10}$$

Look at equations (10) and (9), there is a symmetry between the terms in each. 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,

$$\sum_{ij} \epsilon_{ij} \left(\langle \delta i|j\rangle + \langle i|\delta j\rangle \right). \tag{11}$$

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,

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

Also, it turns out that ϵ is a hermitian matrix, so

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

10

Look at this, the second term in equation(11) is the complex conjugate of the first term.

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

$$[\delta ij|kl] = \int d\vec{x_1} d\vec{x_2} \, \delta i^*(\vec{x_1}) j(\vec{x_1}) \left(\frac{1}{r_{12}}\right) k^*(\vec{x_2}) l(\vec{x_2}),$$

then the complex conjugate of this quantity would be,

$$\int d\vec{x_1} d\vec{x_2} \, j^*(\vec{x_1}) \delta i(\vec{x_1}) \left(\frac{1}{r_{12}}\right) l^*(\vec{x_2}) k(\vec{x_2}) = [j\delta i|lk].$$

From equation(10), the two electron terms are

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

and

$$\frac{1}{2} \sum_{ij} ([\delta ij|ji] + [i\delta j|ji] + [ij|\delta ji] + [ij|j\delta i]) \tag{13}$$

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

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

Starting with the first term of equation(12),

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

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,

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

We started with the complex conjugate of the first term in equation(12) 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

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

We can now write equation(12) as

$$\frac{1}{2} \sum_{ij} ([\delta ii|jj] + [i\delta i|jj]) + c.c. \tag{14}$$

where *c.c.* stands for the complex conjugate terms. using $[ij|kl] = [ji|kl]$, we can further simplify equation(14),

$$\sum_{ij}[\delta ii|jj] + c.c. \tag{15}$$

Doing the same for equation(13), beginning with $[\delta ij|ji]$ and $[i\delta j|ji]$, we get

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

and

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

this time there is no need of renaming. We can now rewrite equation(13) as

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

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

$$\sum_{ij}[\delta ij|ji] + c.c. \tag{16}$$

Putting everything together, we have,

$$\delta \mathcal{L} = \sum_{i}\langle \delta i|h|i\rangle + \sum_{ij}\left([\delta ii|jj] - [\delta ij|ji]\right) - \sum_{ij}\epsilon_{ij}\langle \delta i|j\rangle + c.c. = 0 \tag{17}$$

Writing equation(17) out

$$\begin{aligned}
\delta \mathcal{L} = &\sum_{i}\int d\vec{x_1}\,\delta\chi_i^*(\vec{x_1})h\chi_i(\vec{x_1})\\
&+ \sum_{ij}\left(\int d\vec{x_1}d\vec{x_2}\,\delta\chi_i^*(\vec{x_1})\chi_i(\vec{x_1})\left(\frac{1}{r_{12}}\right)\chi_j^*(\vec{x_2})\chi_j(\vec{x_2})\right.\\
&\left.- \int d\vec{x_1}d\vec{x_2}\,\delta\chi_i^*(\vec{x_1})\chi_j(\vec{x_1})\left(\frac{1}{r_{12}}\right)\chi_j^*(\vec{x_2})\chi_i(\vec{x_2})\right)\\
&- \sum_{ij}\epsilon_{ij}\int d\vec{x_1}\delta\chi_i^*(\vec{x_1})\chi_j(\vec{x_1}) + c.c.
\end{aligned} \tag{18}$$

or equivalently,

$$\begin{aligned}
\delta \mathcal{L} = &0\\
= &\sum_{i}\int d\vec{x_1}\,\delta\chi_i^*(\vec{x_1})\left[h\chi_i(\vec{x_1}) + \sum_{j}\left(\chi_i(\vec{x_1})\int d\vec{x_2}\left(\frac{1}{r_{12}}\right)\chi_j^*(\vec{x_2})\chi_j(\vec{x_2})\right.\right.\\
&\left.\left.- \chi_j(\vec{x_1})\int d\vec{x_2}\left(\frac{1}{r_{12}}\right)\chi_j^*(\vec{x_2})\chi_i(\vec{x_2})\right) - \sum_{j}\epsilon_{ij}\chi_j(\vec{x_1})\right] + c.c.
\end{aligned} \tag{19}$$

$\delta\chi_i^*$ is arbitrary, so it is the quantity inside the brackets that must be zero for any orbital i. The HF equations are obtained by rearranging the term in brackets,

$$\left[h\chi_i(\vec{x_1}) + \sum_{j\neq i} \chi_i(\vec{x_1}) \left(\int d\vec{x_2} \left(\frac{1}{r_{12}}\right) \chi_j^*(\vec{x_2})\chi_j(\vec{x_2}) \right) \right.$$
$$\left. - \sum_{j\neq i} \chi_j(\vec{x_1}) \left(\int d\vec{x_2} \left(\frac{1}{r_{12}}\right) \chi_j^*(\vec{x_2})\chi_i(\vec{x_2}) \right) \right] = \sum_j \epsilon_{ij}\chi_j(\vec{x_1}) \tag{20}$$

Which are the HF equation we derived in the previous section.

5.4 Solving the Hartree-Fock-Roothan Equations

5.4.1 Input

Recall that the problem we wanted to solve was

$$\boldsymbol{FC} = \boldsymbol{SC\epsilon} \tag{21}$$

which can be written as

$$\sum_\nu F_{\mu\nu}C_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu}C_{\nu i} \tag{22}$$

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 the necessary information 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(21) into an eigenvalue problem, so we need to find a transformation that will do just that.

We need to diagonalize the overlap matrix **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

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

To see how this works, from equation(21) assume that there exists a new matrix **C**, such that

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

and

$$C = XC'$$

so that equation(21) becomes

$$FXC' = SXC'\epsilon.$$

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

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

then we can write equation(21) as

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

As it turns out $X \equiv S^{-1/2}$.

5.4.3 Self-Consistent Field Iteration

For the actual implementation of the HF procedure see `hf.py`.

- Begin by calculating the Fock matrix:

  `F = func.makefock(H.Hcore, D, H.twoe)`

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

- 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) = {}\n".format(EN + H.Enuc))
  ```

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

- Test convergence of the density matrix

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

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

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

  i denotes the iteration.

6   Some Other Details

I want to mention some terminology that quantum chemists use: in a restricted HF method the system of study has a closed-shell with all orbitals doubly occupied. For a restricted open-shell HF method the system of study has an open-shell with all orbitals doubly occupied. Lastly, for an unrestricted HF method, as mentioned previously, all orbitals are singly occupied.

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. The HF method can also be viewed as the lowest order perturbation expansion with respect to interactions between 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.

It is worth mentioning that HF usually struggles with transition metals, diradicals and breaking of bonds. Otherwise, HF provides relatively good geometries and energies but is very poor when trying to obtain vibrational frequencies.

7   References

- For more background in quantum mechanics see David J. Griffiths "Introduction to Quantum Mechanics, 2nd ed.".

- These notes are based on : notes compiled by the Sherill group (`http://vergil.chemistry.gatech.edu/notes/`) and on Attila Szabo, Neil S. Ostlund " Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory (Dover Books on Chemistry)".