The Hartree–Fock Method

Jannik P. Roth

March 3, 2023

Abstract

This document is an overview of the theory behind the Hartree–Fock theory. After covering the quantum mechanical basics, the Hartree–Fock approximation is explained in detail. It is mainly based on common literature such as Ref. [1] and Ref. [2] but also contains elements of lectures I heard during my studies at the University of Bonn (Prof. Dr. Thomas Bredow) and the LMU Munich (Prof. Dr. Regina de Vivie-Riedle and Prof. Dr. Christian Ochsenfeld).

This document is part of a the `eduHF` repository which can be found at github.com/jannik-roth/eduHF. If you find any mistake, please contact me via mail jannikproth@gmail.com

Contents

1 Introduction

1.1 What is this document about and who should use it?

Well, it's about computational chemistry, especially Hartree–Fock[1]. Honestly, this document starts with a short revision of quantum mechanic basics and then deals in detail with the theory behind the Hartree–Fock theory. I will show you the derivation step by step and how it is applied in a quantum chemistry program. After that we will start writing our own Hartree–Fock program. Here you will definitely strengthen your understanding tremendously[2]. After that we will start having fun[3] with our Hartree–Fock program. We will append it with various different things going from our own integral code, to a convergence accelerator and do some interesting stuff like population analysis or analytical gradients.

Who should use this document? That's probably the most important question. Short answer: Anyone that is interested in understanding computational chemistry. It is probably best suited for chemistry students who have some entry level university knowledge in math and physics. An introductory course in quantum mechanics is advantageous but not a must and can easily be made up with a good textbook. I tried writing this document in a simple and easy-to-understand style[4]. In my opinion anyone that is willing to and has a knack for science can understand it. Although as already said a good basis in mathematics and physics is necessary. I think it is ideal for chemistry students in their second year. There you should have a good basis and should get through this document without a lot of trouble. However, don't let this hold you back in trying it out earlier. This whole document isn't about any mysterious and mythical derivations and you should not get the impression that any of computational chemistry is like that. Yeah, it is intimidating and might seem a very complicated at first but you have to start simple and then work yourself up. And Hartee–Fock is the basis of almost all computational chemistry. So go ahead and take a look at it!

1.2 Recommended Literature

There are a lot of different books on the basics on quantum mechanics and the Hartree–Fock theory. Here is a short list what I would recommend:

- MODERN QUANTUM CHEMISTRY by A. Szabo and N. S. Ostlund
  This is the standard textbook when it comes to Hartree–Fock and it does a brilliant job at explaining it. The introduction on the mathematical basics which are necessary for understanding is short but enough for what we are doing. Although I don't especially like its take on the electron correlation methods (CI, CC, PT) I think it is an excellent book to study Hartree–Fock. Another big advantage: it's quite cheap only around €15.

- MOLECULAR QUANTUM MECHANICS by P. W. Atkins
  A good book for beginners in the typical Atkins style (you probably know it from his Physical Chemistry book). Easy to understand, not heavy on the mathematics. In my opinion a bit too easy on that. But this is a personal preference. If you just want to understand the basic concepts this is the book to go for.

---

[1]I mean that's kind of the title, isn't it?
[2]I did atleast. It wasn't until I wrote my second Hartree–Fock program that I completely understood it.
[3]No, not that kind of fun you dirty mind!
[4]English isn't my first language so don't expect a new Shakespeare.

- INTRODUCTION TO COMPUTATIONAL CHEMISTRY by F. Jensen
  This one gives a good overview on the different aspects on computational chemistry. It usually gives a short and easy to understand overview. However, it doesn't go into too much detail. I still pick it up regularly if I want just a short and easy introduction on different topics. Overall it is a great at giving you an overview.

- QUANTUM CHEMISTRY by I. N. Levine
  I would say it is similar to Atkins' MOLECULAR QUANTUM MECHANICS in that it is good for beginners and takes you through the basics of quantum mechanics. I personally prefer Levine's style over Atkins'.

- MOLECULAR ELECTRONIC-STRUCTURE THEORY by T. Helgaker, J. Olsen and P. Jorgensen
  Oh boy, that's a book. If you really want to get into the detail and math behind computational chemistry here you go. Not for beginners but probably the most comprehensive book of them all. I only recommend this one if you already have a good understanding of computational chemistry and really want to dive into the specifics.

To summarize: for the basics in quantum chemistry use Atkins or Levine, for Hartree–Fock use Szabo/Ostlund, for an overview in computational chemistry Jensen and for the real math and details Helgaker/Olsen/Jorgensen.

Remember: Don't immediately buy any textbooks! They are very expensive (except for Szabo/Ostlund). Use your library and try and have a look at them. Do you like their style? Is the scope what you need or want? Only if you really want to have one of these books at home you should buy one. It always recommend getting a used one as these are often times considerably cheaper. You can look online but some university libraries sell older books too. If you are studying ask the student council. They sometimes offer a service of selling and buying textbooks. Only buy a new one if it is really necessary. Your bank account and the environment will thank you.

2 Quantum Mechanical Basics

Here we will however give a short repetition of the basics of quantum mechanics. These basics are essential for the rest of this document[5]. We will start with the absolute basics like the Schrödinger equation and the Hamilton Operator but will also talk about some more specific details like the Slater Determinants which will be important for later usage.

2.1 The Schrödinger Equation

Arguably one of the most famous equations in physics. A lot of people have heard about it, some know what it is used for, and even fewer people can pronounce the 'ö' in it[6]. So what is it exactly? In its most general form its called the time-dependent Schrödinger equation.

$$i\hbar \frac{\partial}{\partial t} \Psi\left(\boldsymbol{r}, t\right) = \hat{\mathcal{H}} \Psi\left(\boldsymbol{r}, t\right) \tag{1}$$

As we can see it is a differential equation (because of the differential operator, duh) and it describes the wave function Ψ for a quantum mechanical system. See what was done there? We just give a nice and clean looking equation and say that it describes almost all our problems. Before starting to solve it, let's make it a bit easier. We see that the wave function Ψ depends on the time t. Let's get rid of this. Why can we do this? When solving these kinds of problems a good approach is the product approach. We split the wave function which depends on time and space into the product of two functions, one depending only on the spatial component while the other takes care of the temporal component. Usually we are interested in the so called stationary states. These stationary states are, who would've thought, standing waves. The temporal part of the wave function is a simple standing wave (like a sin function, or more general a complex exponential). This means we can simplify the time-dependent Schrödinger equation (eq. (1)) and obtain the time-independent Schrödinger equation (eq. (2)).

$$\hat{\mathcal{H}} \Psi\left(\boldsymbol{r}, t\right) = E \Psi\left(\boldsymbol{r}, t\right) \tag{2}$$

Nice, know we even have E the energy of the system in our equation! That's what we wanted! But, how do we solve it? What is this wave function Ψ? And what is that fancy $\mathcal{H}$ with the cute hat on top?

2.2 Hamilton Operator

We will first try to answer what the fancy $\mathcal{H}$ in the time-dependent Schrödinger equation (eq. (1)) means. This is the so called Hamilton Operator or just Hamiltonian. One way to denote that it is a operator is by using a the hat symbol ˆ on top. What is an operator? Well, an operator 'does something' with a function and gives back a function or a number. You already know quite a lot of operators. A simple example is the differential operator, e.g. $\frac{\mathrm{d}}{\mathrm{d}x}$ or the integral operator $\int_0^t$. Don't worry if you don't grasp it immediately. It is a useful concept that will show up later a lot and can be understood best with examples.
Now, what is the Hamilton operator exactly? The Hamiltonian mechanics are another way of formulating classical physics. It is completely equivalent to Newton's mechanics but can easily be applied to quantum mechanical systems. The most important property is the

---

[5]Of course you can skip them if you are already quite familiar with these. However, I would advise to at least skim over this section just to be sure you got everything that is necessary for later

[6]Ok, you start with the 'ay' sound in 'day'. Now stand in front of the mirror and slowly round your lips to a circle. Congratulations, you look ridiculous but at least you can pronounce 'Schrödinger' now.

associated quantity which is for the Hamilton operator the energy of the system. This means, that by applying the Hamilton operator on our system we get the energy back. But how does the Hamilton Operator look like? We start simple by splitting the Hamiltonian in two parts: kinetic energy and potential energy.

$$\hat{\mathcal{H}} = \hat{T} + \hat{V} \tag{3}$$

See, we now have an operator $\hat{T}$ for the kinetic energy and an operator $\hat{V}$ for the potential energy. Let's get a bit more specific now. We want to describe molecules which contain positively charged nuclei and negatively charged electrons. What contributions to the energy do we have know? It is clear, that both nuclei and electrons have kinetic energy. This covers the kinetic part, now the potential energy. Electrons and nuclei feel an attractive force according to Coulomb's law. But electrons and electrons repel each other. The same goes for the nuclei who also repel each other. Attractive forces gain a negative sign and repulsive forces a positive. This means we can write down the Hamiltonian for our system as

$$\hat{\mathcal{H}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{1}{2m_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{r_{AB}} \tag{4}$$

Eq. (4) is the full Hamiltonian for a system consisting of N electrons and M nuclei. The electrons are labeled with lowercase letter i, j, ... and nuclei with uppercase letters A, B, ... m_A and Z_A is the mass and charge of nucleus A, r_{ij} is the distance between electron i and j (analogously for r_{iA} and r_{AB}). The kinetic part are the first two sums and the potential part are the last three (double) sums. This is the most general form of the Hamiltonian for our systems and it really doesn't get more complicated. But we can simplify it even more.

You may have also noticed that eq. (4) looks very easy and is missing some constants (like the $\frac{1}{4\pi\varepsilon_0}$ in front of the Coulomb terms). This is because we use atomic units. (Hartree) atomic units are just a different system of units that are used quite a lot in quantum chemistry. In this system, the reduced Planck constant $\hbar$, the elementary charge e, the Bohr radius a_0 and the mass of an electron m_e is set to unity ($\hbar = e = a_0 = m_e = 1$). This leads to a whole different set of units that can be derived from these four physical constants[7]. The most important one is probably the Hartree (E_h) which is the unit of energy in atomic units ($1\,E_h = 27.211\,\text{eV} = 2625.5\,\text{kJ}\,\text{mol}^{-1}$). Often times people are just too lazy to exactly write down which atomic unit it is exactly and just say a.u.. This is not good because it can lead to a lot of confusion. Imagine something like $5\,\text{a.u.}$, is it a length, an energy, a charge or a polarizability? That's why you should always write down what it is exactly and not just go with the lazy a.u.[8]

2.2.1 Born–Oppenheimer Approximation

You know, even though the full Hamiltonian in eq. (4) doesn't look to bad at first, it would lead to a lot of problems later which makes solving the Schrödinger equation considerably more difficult[9]. That's why we make it easier. When comparing the nuclei with electrons we

---

[7]Since the conversion between atomic units and SI units is not always trivial, just google it or go to wikpedia.

[8]Yes, I am angrily looking at you software developers who often write all of there output just in a.u. because you are either too lazy or don't think it is necessary to write down what you actually mean with your numbers.

[9]Trust me on this one, you don't really want to get into that

note that nuclei weigh considerably more. One proton is roughly 1836 times heavier than an electron. So we just say that they move so slow (seen from the perspective of an electron) that they are basically standing still. This is called the Born–Oppenheimer Approximation. One way to understand the consequences is to imagine the electrons as moving particles in a field of nuclei with fixed positions. In general, this approximation is quite good. However, it can break down once we have to energy surfaces which come quite close to each-other, but we won't deal with this. Now, what do we actually gain from this approximation? We can simplify our Hamiltonian because first, the kinetic energy of the nuclei can be neglected and second, the nuclei-nuclei repulsion is now a constant. This simplifies our Hamiltonian in eq. (4) to

$$\hat{\mathcal{H}}^{\mathrm{BO}} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + V_{\mathrm{NN}} \tag{5}$$

Now, this is the full Hamiltonian within the Born–Oppenheimer approximation. Since this approximation is so common, it is considerably rarer to find theories which don't use it. We will only work with this approximation and will drop the explicit notation $^{\mathrm{BO}}$. However, do not forget this! It's a good idea to always keep in mind what kind of approximations you do and when they break down.

2.3 Wave Functions

We previously covered what the fancy $\hat{\mathcal{H}}$ in the Schrödinger equation was, let's now talk about the wave function Ψ. It is an abstract mathematical description of our quantum system and the whole concept is not easy to grasp. We know from quantum mechanics that we can not describe everything as a discrete particle since they can behave like waves on a very small scale. Wave functions are an ideal way to describe these systems but the biggest problem with wave functions is its interpretation: In a classical description we can say that an electron has an exact position and momentum. This is easy to imagine since we think of it as a shrunk down bowling ball floating around in a certain direction. With wave functions this is considerably more difficult because they do not have a physical meaning[10]. What do we do now? Well, Max Born comes to the rescue with the Born rule. The square of the wave function at a certain point is proportional to the probability density $p(\boldsymbol{r}, t)$.

$$p(\boldsymbol{r}, t) = |\Psi(\boldsymbol{r}, t)|^2 \tag{6}$$

This means that the larger the square of our wave function, the larger the probability of finding our particle there.
Now what does a wave function look like? First overall it is important to distinguish between two cases: 1) unbound (free) particles and 2) bound particles. Unbound particles feel no outer force (this means, that the potential is zero) and the solution of the Schrödinger equation just yield a plane wave which is boring. More interesting is the second case of bound particles since that's what we have in molecules. We have four requirements for the wave function of these bound particles:

1. The wave function Ψ must be quadratically normalizable

$$1 \overset{!}{=} \int_{-\infty}^{+\infty} \Psi^*(\boldsymbol{r}) \Psi(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \tag{7}$$

---

[10]Yep, you are maybe as shocked as I was when I first heard it. But don't fall into despair it will make sense soon.

Why? Well, we need to know where the particle actually is and it can't be everywhere. We chose a normalization of unity because it is most convenient. We simply take care of it here so we don't have to deal with it later in the process.

2. The wave function Ψ must be continuously differentiable. We already saw that in eq. (5) where we need the second derivative to work with the Hamiltonian. However, there is one small exception: If the potential has an infinity point, then the function doesn't have to continuously differentiable at this point.

3. The wave function Ψ must be continuous which is necessary for point 2.

4. The wave function must be a clear function. This is necessary for the wave function to even be called 'function'.

These are all the requirements for the wave functions of bound particle. As we can see they still leave us a lot of freedom and we can come up with quite different shapes which all fulfill above properties. But is that really all there is to it? Unfortunately not quite. Nature complicates things again and we will have to consider the Pauli exclusion principle when talking about wave functions.

2.3.1 The Paul exclusion principle

The Pauli exclusion principle, named after Wolfgang Pauli, is a fundamental principle which states that the total wave function has to asymmetric under exchange of two fermions. And now a bit slower so that we understand it: fermions are particles with odd integer spin $(\frac{1}{2}, \frac{3}{2}, \dots)$. Electrons are thereby fermions because their spin is $\pm\frac{1}{2}$. We can describe an electron with four coordinates, 3 spatial (x, y, z or just $\boldsymbol{r}$) and one spin coordinate (σ), all these four coordinates will be abbreviated with $\boldsymbol{\tau}$. Assume now we have a wave function Ψ for a system with more than one electron (we will label them with simple arabic number) $\Psi = \Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2})$. The Pauli exclusion principle states then

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}) = -\Psi(\boldsymbol{\tau_2}, \boldsymbol{\tau_1}) \tag{8}$$

You probably know another formulation of this principle. Two electrons cannot occupy the same quantum state (identical quantum numbers) simultaneously. This actually follows from the more general formulation above[11]. We can also extend this requirement into a more general form

$$\Psi\left(\boldsymbol{\tau_1}, \dots, \boldsymbol{\tau_i}, \dots, \boldsymbol{\tau_N}\right) = -\Psi\left(\boldsymbol{\tau_i}, \dots, \boldsymbol{\tau_1}, \dots, \boldsymbol{\tau_N}\right) \tag{9}$$

When we try to solve for our wave function, which we do by solving the Schrödinger equation, we will have to make sure that the Pauli exclusion principle is met. How can we do that? First we need to consider how to actually construct the wave function for a system.

2.3.2 Orbitals

Solving the time-independent Schrödinger equation (2) leads to the wave function Ψ for the whole system. In quantum chemistry, the wave function Ψ is usually constructed out of orbitals which are wave functions for a single electron. It is advantageous to differentiate between spatial orbitals $\phi(\boldsymbol{r})$ and spin orbitals $\varphi(\boldsymbol{\tau})$ where spin orbitals are the product of

---

[11]Just take a two electron system where $\boldsymbol{\tau_1} = \boldsymbol{\tau_2} = \boldsymbol{\tau}$. Inserting in eq. (8) leads to $\Psi(\boldsymbol{\tau}, \boldsymbol{\tau}) = -\Psi(\boldsymbol{\tau}, \boldsymbol{\tau})$ which can only be true for $\Psi = 0$.

a spatial orbital with a spin function (either α or β for spin up and down, respectively). We want to make things easier again so we will only consider the so called restricted case where the spatial orbital for the two electrons who differ only in their spin coordinate is identical. This means that we can construct two spin orbitals $\varphi(\boldsymbol{\tau})$ from just one spatial orbital $\phi(\boldsymbol{r})$ by multiplying it with the spin functions α and β.

$$\varphi_i\left(\boldsymbol{\tau}\right) = \left\{ \begin{array}{l} \phi_i\left(\boldsymbol{r}\right)\alpha\left(\sigma\right) \\ \phi_i\left(\boldsymbol{r}\right)\beta\left(\sigma\right) \end{array} \right. \tag{10}$$

Applying the Pauli exclusion principle we know that each spin orbital $\varphi(\boldsymbol{\tau})$ can occupy only one electron and each spatial orbital $\phi(\boldsymbol{r})$ can occupy two electrons. A better description would allow the spatial orbital to be different for the two electrons who only differ in spin. This is the so called unrestricted case and is actually not that more difficult. However, it is advantageous to first start and understand the restricted case and then later move on to te unrestricted case.

The set of spatial orbitals $\{\phi\}$ is chosen to fulfil orthonormality.

$$\int_{-\infty}^{+\infty} \phi_i(\boldsymbol{r})\phi_j(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = \delta_{ij} \tag{11}$$

Here, δ_{ij} is the Kronecker delta which is defined as follows.

$$\delta_{ij} = \left\{ \begin{array}{ll} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{array} \right. \tag{12}$$

We will later take a closer look on how to ensure that the set of spatial orbitals $\{\phi\}$ are otrhonormal. Just note, that because the spatial orbitals and the spin functions are both orthonormal sets, the product of the two (which are our spin orbitals for the restricted case) are also orthonormal.

Now we have functions for single electrons but how can we use them to construct the wave function Ψ for our many-electron system?

2.3.3 Hartree Product

The most simple approach is just to use a product approach. The many-electron wave function is the product of each occupied spin orbital.

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}) = \varphi_1(\boldsymbol{\tau_1})\varphi_2(\boldsymbol{\tau_2}) \tag{13}$$

This is known as the Hartree Product or in more general

$$\Psi(\boldsymbol{\tau_1}, \ldots, \boldsymbol{\tau_N}) = \prod_i^N \varphi_i(\boldsymbol{\tau_i}) \tag{14}$$

You already see the problem, don't you? If we would just use Hartree Products Wolfgang Pauli would be very sad and would maybe in a rush of furious anger smack us with his Nobel Prize in Physics. Just take a look at a simple two electron system by exchanging the coordinates

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}) = \varphi_1(\boldsymbol{\tau_1})\varphi_2(\boldsymbol{\tau_2}) \tag{15}$$
$$\Psi(\boldsymbol{\tau_2}, \boldsymbol{\tau_1}) = \varphi_1(\boldsymbol{\tau_2})\varphi_2(\boldsymbol{\tau_1}) \tag{16}$$

And now we have our problem

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}) \neq -\Psi(\boldsymbol{\tau_2}, \boldsymbol{\tau_1}) \tag{17}$$

The Hartree Product is not antisymmetric and thereby violates the Pauli exclusion principle. Ugh, we have to be creative. If you think about it you may come to the solution that looks something like this

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}) = \frac{1}{\sqrt{2}} \left(\varphi_1(\boldsymbol{\tau_1})\varphi_2(\boldsymbol{\tau_2}) - \varphi_1(\boldsymbol{\tau_2})\varphi_2(\boldsymbol{\tau_1}) \right) \tag{18}$$

Just try exchanging the coordinates and you will see that it is antisymmetric. You might be wondering why the factor $\frac{1}{\sqrt{2}}$ is standing there. This is just to make sure that

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Psi^*(\boldsymbol{\tau_1}, \boldsymbol{\tau_2})\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2})\mathrm{d}\boldsymbol{\tau_1}\mathrm{d}\boldsymbol{\tau_2} = 1 \tag{19}$$

which is convenient when calculating properties later on. Alright, seems like we solved our problem with the Pauli exclusion principle at least for two-electron systems. Is there a way to generalize this?

2.3.4 Slater Determinants

Yes, there is. If you squint your eyes and get some flashbacks to high school mathematics you might say that eq. (18) looks kind of like the determinant of a fancy matrix.

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}) = \frac{1}{\sqrt{2}} \begin{vmatrix} \varphi_1(\boldsymbol{\tau_1}) & \varphi_2(\boldsymbol{\tau_1}) \\ \varphi_1(\boldsymbol{\tau_2}) & \varphi_2(\boldsymbol{\tau_2}) \end{vmatrix} \tag{20}$$

Note that in the columns we have the spin orbitals φ and in the rows we have our different electrons. Can we generalize this? Of course!

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}, \ldots, \boldsymbol{\tau_N}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_i(\boldsymbol{\tau_1}) & \varphi_j(\boldsymbol{\tau_1}) & \cdots & \varphi_k(\boldsymbol{\tau_1}) \\ \varphi_i(\boldsymbol{\tau_2}) & \varphi_j(\boldsymbol{\tau_2}) & \cdots & \varphi_k(\boldsymbol{\tau_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_i(\boldsymbol{\tau_N}) & \varphi_j(\boldsymbol{\tau_N}) & \cdots & \varphi_k(\boldsymbol{\tau_N}) \end{vmatrix} \tag{21}$$

Again, the factor $\frac{1}{\sqrt{N!}}$ just takes care of our normalization. But why does this weird looking determinant now take care of the Pauli exclusion principle? Exchange the coordinates of two electrons is like exchanging two rows in the matrix. If we change two rows (or columns) in a matrix, the determinant changes in sign. Nice! Now what happens if we have two electrons in the same spin orbital? This is like having two identical columns. If a matrix has two identical columns (or rows) its determinant is zero. Perfect! This construction, which is called a Slater determinant, takes care of the Pauli exclusion principle. It is extremely convenient for us now since we do not have to check manually whether our solution for the Schrödinger equation fulfills the antisymmetric behaviour because we included this requirement in our construction. This way we do not have to think about it any further and can just continue with our adventure of getting a better understanding on quantum chemistry. The only annoying thing its kind of bulky representation. I mean, who wants to write down the huge matrix for a ten electron system? That's why we use a short-hand notation where we only list the diagonal elements

$$\Psi(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}, \ldots, \boldsymbol{\tau_N}) = |\varphi_i(\boldsymbol{\tau_1})\, \varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N})\rangle \tag{22}$$

2.4 Bra and Ket

People are lazy[12], quantum chemists especially[13]. That is the reason they use so many short-hand notations. You might have already noticed that we have been a bit sloppy when it comes to specifying integrals (and their limits). Let's take a look at the normalized many-electron wave function.

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\Psi^*(\boldsymbol{\tau_1},\boldsymbol{\tau_2})\Psi(\boldsymbol{\tau_1},\boldsymbol{\tau_2})\mathrm{d}\boldsymbol{\tau_1}\mathrm{d}\boldsymbol{\tau_2}=1 \tag{23}$$

We know that the little star * denotes the complex conjugate of the wave function and we also know that $\boldsymbol{\tau}$ contains the spatial ($\boldsymbol{r}$ or x,y,z) and the spin component (σ) which is either spin up or spin down. The integration is carried out from $-\infty$ to $+\infty$ but how does this work with $\boldsymbol{\tau}$? What we actually mean with this integral is to be over the complete space. Instead of writing one integration sign for every of the three spatial coordinates, we just write one

$$\int_{-\infty}^{+\infty}\Psi\mathrm{d}\boldsymbol{r}=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\Psi\mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{24}$$

And what about the spin coordinate? How can we integrate over the spin coordinate if it is either spin up ($+\frac{1}{2}$ or α) or spin down ($-\frac{1}{2}$ or β)? Then the integral is quite simple and just a sum over both possible spin states

$$\int_{-\infty}^{+\infty}\Psi\mathrm{d}\sigma=\sum_{\sigma}\Psi^{\sigma}=\Psi^{\alpha}+\Psi^{\beta} \tag{25}$$

In quantum chemistry it is quite common to have these integrals over the complete space and over both spin states. Because we don't want to write every integral sign we will be using the bra-ket notation (also called Dirac-notation). It uses 'bras' $\langle$, 'kets' $\rangle$ and the vertical line | and is defined as follows:

$$\langle\Psi|=\Psi^*(\boldsymbol{\tau_1},\boldsymbol{\tau_2},\ldots,\boldsymbol{\tau_N}) \tag{26}$$

$$|\Psi\rangle=\Psi(\boldsymbol{\tau_1},\boldsymbol{\tau_2},\ldots,\boldsymbol{\tau_N}) \tag{27}$$

$$\langle\Psi\,|\,\Psi\rangle=\int\cdots\int_{-\infty}^{+\infty}\Psi^*(\boldsymbol{\tau_1},\boldsymbol{\tau_2},\ldots,\boldsymbol{\tau_N})\Psi(\boldsymbol{\tau_1},\boldsymbol{\tau_2},\ldots,\boldsymbol{\tau_N})\mathrm{d}\boldsymbol{\tau_1}\mathrm{d}\boldsymbol{\tau_2}\ldots\mathrm{d}\boldsymbol{\tau_N} \tag{28}$$

As you can see this will shorten equations a lot and that's why we will use it a lot[14]. But there is a little bit more to it: Mathematically the ket $|\Psi\rangle$ is behaving like a column vector that describes a physical state. The according bra $\langle\Psi|$ can be understood as the transpose, so it is a row vector. The product $\langle\Psi\,|\,\Psi\rangle$ is then the vector product between these two vectors. We can eve further extend this notation by including operators like

$$\left\langle\Psi\,\middle|\,\hat{\mathcal{H}}\,\middle|\,\Psi\right\rangle=\int\cdots\int_{-\infty}^{+\infty}\Psi^*(\boldsymbol{\tau_1},\boldsymbol{\tau_2},\ldots,\boldsymbol{\tau_N})\hat{\mathcal{H}}\Psi(\boldsymbol{\tau_1},\boldsymbol{\tau_2},\ldots,\boldsymbol{\tau_N})\mathrm{d}\boldsymbol{\tau_1}\mathrm{d}\boldsymbol{\tau_2}\ldots\mathrm{d}\boldsymbol{\tau_N} \tag{29}$$

I can not stress enough how important this notation is. Really, everyone uses it. It is not difficult to understand and will become natural very fast to you.

---

[12]Trust me, they are.

[13]Trust me again, I've made my experience with them.

[14]I already told you, quantum chemists are lazy and we are on our way to become one.

2.4.1 Slater–Condon Rules

Next up we must learn how to use Slater Determinants in our calculation. We will take a look at how the Hamiltonian acts on a Slater determinant in general. Ok, let's take a look back at the Hamiltonian in eq. (5). First we will ignore the core-core potential V_{NN} since it is a constant. We will then split the Hamiltonian into two parts: One part ($\hat{\mathcal{O}}_1$) will have the one-electron operators and the other ($\hat{\mathcal{O}}_2$) the two-electron operators.

$$\hat{\mathcal{H}} = \hat{\mathcal{O}}_1 + \hat{\mathcal{O}}_2 \tag{30}$$

$$\hat{\mathcal{O}}_1 = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} = \sum_{i=1}^{N} \hat{h}_i \tag{31}$$

$$\hat{\mathcal{O}}_2 = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{32}$$

To do this evaluation we will introduce a different way of writing the Slater-Determinant. In a Slater-Determinant every electron is once in each orbital. We will know write this using a permutational operator $\hat{\mathcal{P}}$ which permutes every electron into every orbital

$$|\Psi\rangle = \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} (-1)^{P_n} \hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \tag{33}$$

Alright, the $\frac{1}{\sqrt{N!}}$ is a normalization factor, P_n is the number of permutation and $(-1)^{P_n}$ thereby takes care of the correct sign. In a Slater-Determinant this was done through the way of calculating the determinant. $\hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\}$ is the permutational operator acting on the set of our spin orbitals to ensure antisymmetry.

Let's star simple by considering the integral

$$\langle \Psi \,|\, \Psi \rangle = 1 \tag{34}$$

Substituting eq. (33) yields

$$\begin{aligned}
\langle \Psi \,|\, \Psi \rangle = &\frac{1}{\sqrt{N!}} \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} \sum_{n=1}^{M!} (-1)^{P_n}(-1)^{P_m} \\
&\times \int \cdots \int \mathrm{d}\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \right. \\
&\left. \times \hat{\mathcal{P}}_{\mathrm{m}} \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right]
\end{aligned} \tag{35}$$

Ok, let's deal with that step by step. We will just consider one integral in this double sum. Start with an easy one:

$$\langle \varphi_i(\boldsymbol{\tau_1})\varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_i(\boldsymbol{\tau_1})\varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \rangle \tag{36}$$

Here, the same electron is in the same orbital in the bra and in the ket. Since we carry out the integration over one $\boldsymbol{\tau}$ after the other we can write these easier.

$$\langle \varphi_i(\boldsymbol{\tau_1}) \,|\, \varphi_i(\boldsymbol{\tau_1}) \rangle \cdot \langle \varphi_j(\boldsymbol{\tau_2}) \,|\, \varphi_j(\boldsymbol{\tau_2}) \rangle \cdot \ldots \cdot \langle \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_k(\boldsymbol{\tau_N}) \rangle \tag{37}$$

Remember the orthonormality condition we imposed on the molecular orbitals? Here it greatly simplifies the evaluation of above expression because $\langle \varphi_i(\boldsymbol{\tau}) \,|\, \varphi_i(\boldsymbol{\tau}) \rangle = \delta_{ij}$. This means we only got ones and the product of ones is a simple one!

$$
\begin{aligned}
&\langle \varphi_i(\boldsymbol{\tau_1}) \varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_i(\boldsymbol{\tau_1}) \varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \rangle \\
&= \underbrace{\langle \varphi_i(\boldsymbol{\tau_1}) \,|\, \varphi_i(\boldsymbol{\tau_1}) \rangle}_{1} \cdot \underbrace{\langle \varphi_j(\boldsymbol{\tau_2}) \,|\, \varphi_j(\boldsymbol{\tau_2}) \rangle}_{1} \cdot \ldots \\
&= 1
\end{aligned}
\tag{38}
$$

Only if all electrons (designated by their coordinates $\boldsymbol{\tau}$) are in the same molecular orbitals, the product gives a one. What if they are in different orbitals? Just take a look at:

$$
\langle \varphi_i(\boldsymbol{\tau_1}) \varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_j(\boldsymbol{\tau_1}) \varphi_i(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \rangle
\tag{39}
$$

We can take a look at the integrals separately again

$$
\langle \varphi_i(\boldsymbol{\tau_1}) \,|\, \varphi_j(\boldsymbol{\tau_1}) \rangle \cdot \langle \varphi_j(\boldsymbol{\tau_2}) \,|\, \varphi_i(\boldsymbol{\tau_2}) \rangle \cdot \ldots \cdot \langle \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_k(\boldsymbol{\tau_N}) \rangle
\tag{40}
$$

and now the orthonormality strikes again. Because i and j are different, the integral is zero and thereby the whole product is zero:

$$
\begin{aligned}
&\langle \varphi_i(\boldsymbol{\tau_1}) \varphi_j(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_j(\boldsymbol{\tau_1}) \varphi_i(\boldsymbol{\tau_2}) \ldots \varphi_k(\boldsymbol{\tau_N}) \rangle \\
&= \underbrace{\langle \varphi_i(\boldsymbol{\tau_1}) \,|\, \varphi_j(\boldsymbol{\tau_1}) \rangle}_{0} \cdot \underbrace{\langle \varphi_j(\boldsymbol{\tau_2}) \,|\, \varphi_i(\boldsymbol{\tau_2}) \rangle}_{0} \cdot \ldots \cdot \underbrace{\langle \varphi_k(\boldsymbol{\tau_N}) \,|\, \varphi_k(\boldsymbol{\tau_N}) \rangle}_{1} \\
&= 0
\end{aligned}
\tag{41}
$$

Ok, now that we learnt that, we can go back to eq. (35) and try to evaluate it. We learnt that all electrons must be in the same orbitals in the bra and in the ket, otherwise the integrals will result in zero. What does that mean for our double sum in eq. (35)? If you think about it you will realize that the permutation on the left side and the permutation on the right side have to be identical because otherwise some electrons will be in different orbitals left and right and the whole integrals is zero. How can we translate this in mathematical terms? Actually quite easy, in our case it means that $m = n$.

$$
\begin{aligned}
\langle \Psi \,|\, \Psi \rangle = &\frac{1}{\sqrt{N!}} \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} \sum_{n=1}^{N!} (-1)^{P_n} (-1)^{P_n} \\
&\times \int \cdots \int \mathrm{d}\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \right. \\
&\times \left. \hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right]
\end{aligned}
\tag{42}
$$

We go through it step by step. First, we get rid of the double sum because the index is identical.

$$
\sum_{n=1}^{N!} \sum_{n=1}^{N!} = \sum_{n=1}^{N!}
\tag{43}
$$

Next:

$$
(-1)^{P_n} (-1)^{P_n} = (-1)^{2P_n} = (-1)^{2P_n} = 1^{P_n} = 1
\tag{44}
$$

This one was easy! Now the integrals: If the permutation is always identical the electrons will always be in the same orbitals and the whole integrals will be one:

$$
\int \cdots \int \mathrm{d}\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \hat{\mathcal{P}}_{\mathrm{n}} \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right] = 1
\tag{45}
$$

13

Now we put it together:

$$\langle \Psi \mid \Psi \rangle = \frac{1}{\sqrt{N!}} \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} 1 \cdot 1 = \frac{1}{N!} N! = 1 \tag{46}$$

Yes, we did it! But was this a surprise? No, actually not. We want the wave function Ψ to be normalized so we already knew that the result had to be one. But still good to check that our math works out. Now it's time to move on to the one- and two-electron part of the Hamiltonian.

The one-electron part $\hat{\mathcal{O}}_1$

We first start with the easier one-electron part where we try to evaluate

$$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi \right\rangle = \left\langle \Psi \,\middle|\, \sum_i^N \hat{h}_i \,\middle|\, \Psi \right\rangle \tag{47}$$

To make our lives easier we will use the fact that electrons are indistinguishable particles[15]. This means that all the elements of the sum are identical and we can write:

$$\left\langle \Psi \,\middle|\, \sum_i^N \hat{h}_i \,\middle|\, \Psi \right\rangle = N \left\langle \Psi \,\middle|\, \hat{h}_1 \,\middle|\, \Psi \right\rangle \tag{48}$$

We we substitute:

$$\begin{aligned}
N \left\langle \Psi \,\middle|\, \hat{h}_1 \,\middle|\, \Psi \right\rangle =& N \frac{1}{\sqrt{N!}} \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} \sum_{n=1}^{M!} (-1)^{P_n} (-1)^{P_m} \\
& \times \int \cdots \int \mathrm{d}\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_n \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \right. \\
& \left. \times \hat{h}_1 \, \hat{\mathcal{P}}_m \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right]
\end{aligned} \tag{49}$$

Now, which terms actually contribute to the sum? Because $\hat{h}_1$ only acts on electron 1 we can separate again

$$\left\langle \varphi_i(\boldsymbol{\tau_1}) \,\middle|\, \hat{h}_1 \,\middle|\, \varphi_i(\boldsymbol{\tau_1}) \right\rangle \underbrace{\langle \varphi_j(\boldsymbol{\tau_2}) \mid \varphi_j(\boldsymbol{\tau_2}) \rangle}_{1} \cdot \ldots \cdot \underbrace{\langle \varphi_k(\boldsymbol{\tau_N}) \mid \varphi_k(\boldsymbol{\tau_N}) \rangle}_{1} = \left\langle \varphi_i(\boldsymbol{\tau_1}) \,\middle|\, \hat{h}_1 \,\middle|\, \varphi_i(\boldsymbol{\tau_1}) \right\rangle \tag{50}$$

If all electrons are in the same orbital in bra and ket, we simply get the one-electron operator acting on one orbitals. But how many possibilities are there? Well if electron 1 is fixed, then there are $(N-1)!$ possible ways to permutate the other electrons around and still get the same result. Of course only if we permutate the other orbitals in the bra and the ket the same way. What do we get if we don't do that?

$$\left\langle \varphi_i(\boldsymbol{\tau_1}) \,\middle|\, \hat{h}_1 \,\middle|\, \varphi_j(\boldsymbol{\tau_1}) \right\rangle \underbrace{\langle \varphi_j(\boldsymbol{\tau_2}) \mid \varphi_i(\boldsymbol{\tau_2}) \rangle}_{0} \cdot \ldots \cdot \underbrace{\langle \varphi_k(\boldsymbol{\tau_N}) \mid \varphi_k(\boldsymbol{\tau_N}) \rangle}_{1} = 0 \tag{51}$$

---

[15]We can't just point our finger at an electron and say this is electron 1 and then at another and say this is electron 2. There are also practical difficulties when it comes to pointing your finger at electrons. They are rather shy and don't like it when you point your finger at them.

Familiar, right? Again we need the permutations to be identical ($m = n$) else we end up with a zero.

$$\left\langle \Psi \left| \hat{\mathcal{O}}_1 \right| \Psi \right\rangle = N \frac{1}{\sqrt{N!}} \frac{1}{\sqrt{N!}} (N-1)! \sum_i^k \left\langle \varphi_i \left| \hat{h} \right| \varphi_i \right\rangle$$
$$= \sum_i^k \left\langle i \left| \hat{h} \right| i \right\rangle \tag{52}$$

See how we became even more lazy with the notation? Not only are we dropping the electron coordinates (the τ), we don't even write the symbol φ for the orbitals anymore and just use the index. This is unfortunately not uncommon in the literature because it increases readability but only if you know what's behind it. I'm afraid that you will have to get used to it. Anyway, we got our first proper result which we will use extensively in the future[16].

We will make a short jump here to complete the Slater–Condon rules. We have to talk about excitations. Now a normal molecule at low temperatures is mainly in its ground state. This means, that the electrons occupy the energetically lowest lying orbitals. What happens if we excite such a molecule with heat, light or through something else? We can imagine one electron jumping from a low lying orbital to a higher one. How do we write that down? It's actually quite easy. A ground-state Slater-Determinant looks like

$$|\Psi\rangle = |\varphi_i \varphi_j \ldots \varphi_k\rangle \tag{53}$$

Exciting one electron from the ith orbital to a previously unoccupied ath orbital looks like this:

$$|\Psi_i^a\rangle = |\varphi_a \varphi_j \ldots \varphi_k\rangle \tag{54}$$

The notation is simple and easy to read: subscript is the orbital that we excited from and superscript is the orbital we excited to. If we excite only a single electron (like $|\Psi_i^a\rangle$) we speak of single excitations or just singles, if we excite two electrons (like $|\Psi_{ij}^{ab}\rangle$) we speak of double excitations or doubles and so on. We will always denote occupied orbitals with indices $i, j, k, \cdots$ and unoccupied/virtual orbitals with $a, b, c, \cdots$.

Now we will evaluate integrals of the type

$$\left\langle \Psi \left| \hat{\mathcal{O}}_1 \right| \Psi_i^a \right\rangle \tag{55}$$

Inserting the sum and using the indistinguishable character of electrons

$$\left\langle \Psi \left| \hat{\mathcal{O}}_1 \right| \Psi_i^a \right\rangle = \left\langle \Psi \left| \sum_i^N \hat{h}_i \right| \Psi_i^a \right\rangle = N \left\langle \Psi \left| \hat{h}_1 \right| \Psi_i^a \right\rangle \tag{56}$$

and now we simply need to repeat what we did above. Inserting eq. (33):

$$N \left\langle \Psi \left| \hat{h}_1 \right| \Psi_i^a \right\rangle = N \frac{1}{\sqrt{N!}} \frac{1}{\sqrt{N!}} \sum_{n=1}^{N!} \sum_{n=1}^{M!} (-1)^{P_n} (-1)^{P_m}$$
$$\times \int \cdots \int \mathrm{d}\tau_1 \ldots \tau_N \left[\hat{\mathcal{P}}_n \left\{ \varphi_i^*(\tau_1), \varphi_j^*(\tau_2), \ldots, \varphi_k^*(\tau_N) \right\} \right. \tag{57}$$
$$\left. \times \hat{h}_1 \, \hat{\mathcal{P}}_m \left\{ \varphi_a(\tau_1), \varphi_j(\tau_2), \ldots, \varphi_k(\tau_N) \right\} \right]$$

---

[16]Don't worry we will later summarize all of these in a nice table for you to print out.

We already learned that due to the orthonormality of the molecular orbitals

$$\langle i \,|\, j \rangle = \delta_{ij} \tag{58}$$

and because the excited determinant has an orbital a which is not in the ground-state determinant, this orbital has to be part of the operator i.e. it has to stand in

$$\left\langle i \,\middle|\, \hat{h}_1 \,\middle|\, a \right\rangle \tag{59}$$

because if it is in the simple overlap term, the result will be zero. Then we have $(N-1)!$ possible permutations for the other electrons in the orbitals which are present in both determinants. This cancels nicely with the prefactor and we end up with only one term.

$$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi_i^a \right\rangle = \left\langle i \,\middle|\, \hat{h} \,\middle|\, a \right\rangle \tag{60}$$

We went through this a little bit faster because it is exactly the same reasoning as above. We always have to make sure that we have only identical permutations and we end up with overlap integrals that always yield 1. If the orbitals are different we have to put them into the term with the operator. Since we only have place for one orbital in the operator we have to put our 'excited' electron in there to not get a zero.
Let's move on to double excitations:

$$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi_{ij}^{ab} \right\rangle \tag{61}$$

What do we expect? Well, first we notice that we only have place for one orbital in the part with the operator. Let's put the orbital a in there. But then we still got b left in the overlap part and we can't pair it up with another b in the ground-state determinant. Because b is orthogonal to every other orbital we will always get a zero. The reasoning is of course identical if we put b in the operator term. This means that the one-electron part is zero if we use a double-excited determinant.

$$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi_{ij}^{ab} \right\rangle = 0 \tag{62}$$

The two-electron part $\hat{\mathcal{O}}_2$
Let's move on to the more complicated two-electron part. What is it actually? It's just the Coulomb interaction!

$$\hat{\mathcal{O}}_2 = \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} = \frac{1}{2} \sum_i^N \sum_{j \neq i}^N \frac{1}{r_{ij}} \tag{63}$$

Try to get used to the summation and how it works. First overall it is important that in both summations $j \neq i$. This makes sure that we don't have self interaction. An electron does not feel repulsion from itself[17]. Then we need to take care of not double counting. This is done in the first summation by making sure that j is always larger than i. In the other case we do double count but divide by 2 to cancel it. Especially in the second summation it is easy to see that the first sum runs over all N electrons while the second one over $N-1$ electrons. Using the indistinguishable character of electrons[18] we can simplify wrtiting the two-electron part.

$$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_2 \,\middle|\, \Psi \right\rangle = \left\langle \Psi \,\middle|\, \sum_{j>i} \frac{1}{r_{ij}} \,\middle|\, \Psi \right\rangle = \frac{N(N-1)}{2} \left\langle \Psi \,\middle|\, \frac{1}{r_{12}} \,\middle|\, \Psi \right\rangle \tag{64}$$

---

[17] We would also get into trouble dividing by 0.

[18] Yep, I know it gets a bit boring to always write this but we are just trying to abuse the concepts as much as possible.

16

The prefactor $\frac{N(N-1)}{2}$ is just the number of possible pairing we can have. Inserting eq. (33) again:

$$
\begin{aligned}
\frac{N(N-1)}{2} \left\langle \Psi \left| \frac{1}{r_{12}} \right| \Psi \right\rangle =& \frac{N(N-1)}{2} \frac{1}{N!} \sum_{n=1}^{N!} \sum_{n=1}^{M!} (-1)^{P_n} (-1)^{P_m} \\
& \times \int \cdots \int d\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_n \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \right. \\
& \times \left. \frac{1}{r_{12}} \hat{\mathcal{P}}_m \left\{ \varphi_i(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right]
\end{aligned}
\tag{65}
$$

Now the biggest difference to before is that the operator we are using now can take two orbitals. Let's just use the simple example of the orbitals i and j. Which possible combinations can we produce? After a bit of thinking you will come up with these four integrals:

$$
\begin{aligned}
& \left\langle \varphi_i(1)\varphi_j(2) \left| \frac{1}{r_{12}} \right| \varphi_i(1)\varphi_j(2) \right\rangle - \left\langle \varphi_i(1)\varphi_j(2) \left| \frac{1}{r_{12}} \right| \varphi_j(1)\varphi_i(2) \right\rangle \\
& + \left\langle \varphi_j(1)\varphi_i(2) \left| \frac{1}{r_{12}} \right| \varphi_j(1)\varphi_i(2) \right\rangle - \left\langle \varphi_j(1)\varphi_i(2) \left| \frac{1}{r_{12}} \right| \varphi_i(1)\varphi_j(2) \right\rangle \\
& = \langle ij \,|\, ij \rangle - \langle ij \,|\, ji \rangle + \langle ji \,|\, ji \rangle - \langle ji \,|\, ij \rangle
\end{aligned}
\tag{66}
$$

Note that we used a short-hand notation again? Dropping the $\frac{1}{r_{12}}$, the electron designation and the φ symbols. This is the Dirac notation (sometimes called physicists notation) of the two electron integrals. The vertical line is the $\frac{1}{r_{12}}$ operator and to the right and the left of it are always the two orbitals in which the two electrons are. The left orbital is always electron 1 and the right electron 2. Now, what about the other $N-2$ electrons? Of course there are $(N-2)!$ possible permutations for these. To put everything together:

$$
\begin{aligned}
\frac{N(N-1)}{2} \left\langle \Psi \left| \frac{1}{r_{12}} \right| \Psi \right\rangle =& \frac{N(N-1)}{2} \frac{1}{N!} (N-2)! \sum_i^N \sum_j^N \langle ij \,|\, ij \rangle - \langle ij \,|\, ji \rangle \\
=& \frac{1}{2} \sum_i^N \sum_j^N \langle ij \,|\, ij \rangle - \langle ij \,|\, ji \rangle
\end{aligned}
\tag{67}
$$

Wait a moment, didn't we say that we need to make sure that $i \neq j$ in the summation? Don't we have self-interaction? No actually not. Take a closer look. If $i = j$ then the term becomes

$$
\langle ii \,|\, ii \rangle - \langle ii \,|\, ii \rangle = 0
\tag{68}
$$

and thus we have no self-interaction[19]. The factor $\frac{1}{2}$ takes care of double counting again. To sum it up we got

$$
\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi \right\rangle = \frac{1}{2} \sum_i^N \sum_j^N \langle ij \,|\, ij \rangle - \langle ij \,|\, ji \rangle
\tag{69}
$$

for the two-electron part. Now, let's move on to single excitations!

$$
\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi_i^a \right\rangle = ?
\tag{70}
$$

---

[19]This is actually an amazing property because not all methods have it. In density functional theory self-interaction can pose quite a big problem.

First using the indistinguishable property as above and inserting leads to

$$\frac{N(N-1)}{2} \left\langle \Psi \left| \frac{1}{r_{12}} \right| \Psi_i^a \right\rangle = \frac{N(N-1)}{2} \frac{1}{N!} \sum_{n=1}^{N!} \sum_{n=1}^{M!} (-1)^{P_n} (-1)^{P_m}$$
$$\times \int \cdots \int d\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_n \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \quad (71)$$
$$\times \frac{1}{r_{12}} \hat{\mathcal{P}}_m \left\{ \varphi_a(\boldsymbol{\tau_1}), \varphi_j(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right]$$

Here we need to make sure that the orbitals i and a are in the operator because otherwise we will get zero. Using the orbitals i, j and a as an example, we only need to compute

$$\frac{1}{2} \sum_j^N \langle ij \,|\, aj \rangle - \langle ij \,|\, ja \rangle + \langle ji \,|\, ja \rangle - \langle ji \,|\, aj \rangle \quad (72)$$

Compare this with what we had for the one-electron part above. We need to drop the summation over i because it has to stay inside the operator all the time. Using the symmetry in the Dirac notation we can simplify this to

$$\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi_i^a \right\rangle = \sum_j^N \langle ij \,|\, aj \rangle - \langle ij \,|\, ja \rangle \quad (73)$$

We don't need to take care when it comes to doble counting because we don't actually have two sums.

Next are double excitations!

$$\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi_{ij}^{ab} \right\rangle = ? \quad (74)$$

Same method as before: indistinguishable electrons and permutational operator:

$$\frac{N(N-1)}{2} \left\langle \Psi \left| \frac{1}{r_{12}} \right| \Psi_{ij}^{ab} \right\rangle = \frac{N(N-1)}{2} \frac{1}{N!} \sum_{n=1}^{N!} \sum_{n=1}^{M!} (-1)^{P_n} (-1)^{P_m}$$
$$\times \int \cdots \int d\boldsymbol{\tau_1} \ldots \boldsymbol{\tau_N} \left[\hat{\mathcal{P}}_n \left\{ \varphi_i^*(\boldsymbol{\tau_1}), \varphi_j^*(\boldsymbol{\tau_2}), \ldots, \varphi_k^*(\boldsymbol{\tau_N}) \right\} \quad (75)$$
$$\times \frac{1}{r_{12}} \hat{\mathcal{P}}_m \left\{ \varphi_a(\boldsymbol{\tau_1}), \varphi_b(\boldsymbol{\tau_2}), \ldots, \varphi_k(\boldsymbol{\tau_N}) \right\} \right]$$

It is quite clear here that we need to put i, j and a, b in the operator because we will get zero otherwise. Thus we only need to consider

$$\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi_{ij}^{ab} \right\rangle = \frac{1}{2} \left[\langle ij \,|\, ab \rangle - \langle ij \,|\, ba \rangle + \langle ji \,|\, ba \rangle - \langle ji \,|\, ab \rangle \right] \quad (76)$$

Comparing to the integral for single excitations we lost another summation. Using the symmetry of the integrals we simplify to

$$\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi_{ij}^{ab} \right\rangle = \langle ij \,|\, ab \rangle - \langle ij \,|\, ba \rangle \quad (77)$$

Last but not least, triple excitations!

$$\left\langle \Psi \left| \hat{\mathcal{O}}_2 \right| \Psi_{ijk}^{abc} \right\rangle = ? \quad (78)$$

You can probably already guess this will be a short one. We have to fit three orbitals inside the operator to not get zero ($\langle i\,|\,j \rangle = \delta_{ij}$). Our operator can only hold two orbitals. This means we will always get a zero:

$$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_2 \,\middle|\, \Psi_{ijk}^{abc} \right\rangle = 0 \tag{79}$$

Puhh, that was rather difficult. But we did a lot of work which will be important later on. You just need to understand this once and from now on we will only use the results. Nevertheless it is good to see how to arrive at these Slater–Condon rules.

Here they are summarized again:

$\hat{\mathcal{O}}_1 = \sum_{i=1}^{N} \hat{h}_i$	
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi \right\rangle$	$\sum_i^N \left\langle i \,\middle|\, \hat{h} \,\middle|\, i \right\rangle$
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi_i^a \right\rangle$	$\left\langle i \,\middle|\, \hat{h} \,\middle|\, a \right\rangle$
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_1 \,\middle|\, \Psi_{ij}^{ab} \right\rangle$	0

Table 1: Slater–Condon rules for the one-electron part.

$\hat{\mathcal{O}}_2 = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}}$	
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_2 \,\middle|\, \Psi \right\rangle$	$\frac{1}{2} \sum_i^N \sum_j^N \langle ij\,|\,ij \rangle - \langle ij\,|\,ji \rangle$
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_2 \,\middle|\, \Psi_i^a \right\rangle$	$\sum_j^N \langle ij\,|\,aj \rangle - \langle ij\,|\,ja \rangle$
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_2 \,\middle|\, \Psi_{ij}^{ab} \right\rangle$	$\langle ij\,|\,ab \rangle - \langle ij\,|\,ba \rangle$
$\left\langle \Psi \,\middle|\, \hat{\mathcal{O}}_2 \,\middle|\, \Psi_{ijk}^{abc} \right\rangle$	0

Table 2: Slater–Condon rules for the two-electron part.

2.5 Now we are ready

Puh, a lot of work but we made good progress. We will use the results of this section extensively in the coming sections. Maybe you have already seen some of it but it is nevertheless a good exercise to repeat these every once in a while. Now we are ready to get started with the theory behind Hartree–Fock!

3 The Hartree–Fock Approximation

In the previous section, we refreshed our knowledge on all the quantum mechanical basics that we need to better understand the Hartree–Fock approximation. Now it is our time to get started!

I would first like to mention some general remarks. The Hartree–Fock approximation is the most general and basic approximation within quantum chemistry. It is not perfect, and by today's standards quite lackluster. If you'd try to publish something with just a Hartree–Fock calculation, the reviewers would probably just laugh at you. Then, why is it so important? It is the basis for (almost) all other more sophisticated quantum chemical techniques such as configuration interaction (CI), coupled cluster (CC) and even density-functional theory (DFT). To understand these 'modern' methods you first need to get behind the Hartree–Fock approximation. You can trust me, once you understood Hartree–Fock, the rest will be a lot easier.

3.1 Optimizing a Trial Function

We first need to talk about how to optimize the wave function of our quantum system. We let $\tilde{\Phi}$ be a normalized trial function for a many-particle system. We are mainly interested in the energy E of the system and express it as a functional of the trial function $\tilde{\Phi}$. A functional is similar to a function: while a function takes a number and gives back a number, a functional takes a function and gives a number. We can use the Hamilton operator to write:

$$E\left[\tilde{\Phi}\right] = \left\langle \tilde{\Phi} \,\middle|\, \hat{\mathcal{H}} \,\middle|\, \tilde{\Phi} \right\rangle \tag{80}$$

A common strategy when it comes to optimization of functions is to introduce an infinitesimal change $\delta\tilde{\Phi}$ in it and then later setting this change to zero. We do this with our trial function ($\tilde{\Phi} \to \tilde{\Phi} + \delta\tilde{\Phi}$).

$$
\begin{aligned}
E\left[\tilde{\Phi} + \delta\tilde{\Phi}\right] &= \left\langle \tilde{\Phi} + \delta\tilde{\Phi} \,\middle|\, \hat{\mathcal{H}} \,\middle|\, \tilde{\Phi} + \delta\tilde{\Phi} \right\rangle \\
&= E\left[\tilde{\Phi}\right] + \left\{ \left\langle \tilde{\Phi} + \delta\tilde{\Phi} \,\middle|\, \hat{\mathcal{H}} \,\middle|\, \tilde{\Phi} \right\rangle + \left\langle \tilde{\Phi} \,\middle|\, \hat{\mathcal{H}} \,\middle|\, \tilde{\Phi} + \delta\tilde{\Phi} \right\rangle \right\} + \ldots \\
&= E\left[\tilde{\Phi}\right] + \delta E + \ldots
\end{aligned}
\tag{81}
$$

As already stated above, we aim to optimize the wave function to minimize the energy. This is equivalent to the first derivative being zero (we want to be in a minimum where every change leads to an increase in the energy). This is identical to collecting all terms which are first-order in the variation of $\delta\tilde{\Phi}$ and set it to zero.

$$\delta E = 0 \tag{82}$$

This ensures that the energy is stationary with respect to any change in the trial function.

The trial function can be constructed as a linear combination of fixed set basis functions $\{|\Psi_i\rangle\}$. Using these kinds of linear combinations is very common in quantum chemistry and will be used a lot.

$$\left|\tilde{\Phi}\right\rangle = \sum_{i=1}^{N} c_i \left|\Psi_i\right\rangle \tag{83}$$

The energy, which to minimize, can be calculated quite simply.

$$E = \left\langle \tilde{\Phi} \, \middle| \, \hat{\mathcal{H}} \, \middle| \, \tilde{\Phi} \right\rangle = \sum_{ij} c_i^* c_j \left\langle \Psi_i \, \middle| \, \hat{\mathcal{H}} \, \middle| \, \Psi_j \right\rangle \tag{84}$$

During the minimization, the trial function $\tilde{\Phi}$ is kept normalized.

$$\left\langle \tilde{\Phi} \, \middle| \, \tilde{\Phi} \right\rangle - 1 = 0 = \sum_{ij} c_i^* c_j \left\langle \Psi_i \, \middle| \, \Psi_j \right\rangle - 1 \tag{85}$$

This is just extremely convenient and allows us to deal less with annoying normalization problems. This constraint is introduced via a Lagrange multiplier. You do not know what a Lagrange multiplier is? It is a very convenient mathematical construct that allows us to introduce constraints (or side conditions) when solving our problem. They are heavily used in physics and we will use them here and later a couple of times. Unfortunately, I won't go into too much detail behind it. If you are not familiar with them I'd advise you to learn a bit more about them by yourself.

Minimize the energy with respect to the coefficients c_i leads to the Lagrangian given by eq (86) where we used the energy E as the Lagrange multiplier. This is (for a normalized function) identical to adding a zero.

$$\mathcal{L} = \left\langle \tilde{\Phi} \, \middle| \, \hat{\mathcal{H}} \, \middle| \, \tilde{\Phi} \right\rangle - E \left(\sum_{ij} c_i^* c_j \left\langle \Psi_i \, \middle| \, \Psi_j \right\rangle - 1 \right) \tag{86}$$

Now we set the first variation in $\mathcal{L}$ to zero:

$$\begin{aligned}
\delta\mathcal{L} = &\sum_{ij} \delta c_i^* c_j \left\langle \Psi_i \, \middle| \, \hat{\mathcal{H}} \, \middle| \, \Psi_j \right\rangle - E \sum_{ij} \delta c_i^* c_j \left\langle \Psi_i \, \middle| \, \Psi_j \right\rangle \\
&+ \sum_{ij} c_i^* \delta c_j \left\langle \Psi_i \, \middle| \, \hat{\mathcal{H}} \, \middle| \, \Psi_j \right\rangle - E \sum_{ij} c_i^* \delta c_j \left\langle \Psi_i \, \middle| \, \Psi_j \right\rangle .
\end{aligned} \tag{87}$$

Note that the energy E is real[20]. Collecting terms and rearranging leads to eq. (88).

$$\sum_i \delta c_i^* \left(\sum_j H_{ij} c_j - E S_{ij} c_j \right) + \text{c.c.} = 0 \tag{88}$$

with

$$H_{ij} = \left\langle \Psi_i \, \middle| \, \hat{\mathcal{H}} \, \middle| \, \Psi_j \right\rangle \tag{89}$$

$$S_{ij} = \left\langle \Psi_i \, \middle| \, \Psi_j \right\rangle \tag{90}$$

Here we have a definition for the elements of the Hamiltonian matrix $\boldsymbol{H}$ as well as a definition for an overlap matrix $\boldsymbol{S}$. The c.c. just stands for the complex conjugate of all the previous terms. This shortens the equations and is done quite often. You usually just deal with one half of the equation then because the same usually goes for the second part. For eq. (88) to be fulfilled, the term in the brackets has to be zero because the $\delta c_i^\star$ are arbitrary. This leads to the simple result:

$$\boldsymbol{Hc} = E\boldsymbol{Sc} \tag{91}$$

where we used a matrix notation. We will use the matrix notation quite often in the future. Not only do we save time by not having to write down all these indices but they are also easy to implement in code later on. Nevertheless, you should get familiar with both notations, as sums and as matrix multiplication.

---

[20]I mean, what are imaginary energies?

3.2 Energy Minimization for a Single Determinant

You might have noticed that in the previous section we build our trial function $\tilde{\Phi}$ as a linear combination of multiple Slater-Determinants Ψ. Now we go back a bit and just use a single Slater-Determinant. This simplifies things tremendously for us and still gives us an adequate description. You will see later, that a different more sophisticated quantum chemical methods reintroduce more than one Slater-Determinant to get a better description. However, in Hartree–Fock we are satisfied with just a single Slater-Determinant. This is the reason why Hartree–Fock is often described as a single determinant picture or method. We will later also dive a bit into the shortcomings of this.

When using a single Slater-Determinant of the form $|\Psi_0\rangle = |\varphi_1\varphi_2\ldots\varphi_N\rangle$ the energy can be expressed as functional of the spin orbitals $\{\varphi_i\}$. Analogously, the minimization of the energy with respect to the spin orbitals can be done while keeping the spin orbitals orthonormal[21]. This constraint can be formulated as

$$\langle \varphi_i \,|\, \varphi_j \rangle - \delta_{ij} = 0 \tag{92}$$

which leads to the Lagrangian[22]

$$\mathcal{L}\left[\{\varphi_i\}\right] = E_0\left[\{\varphi_i\}\right] - \sum_{i=1}^{N}\sum_{j=1}^{N}\varepsilon_{ij}\left(\langle \varphi_i \,|\, \varphi_j \rangle - \delta_{ij}\right) \tag{93}$$

with the Lagrange multipliers ε_{ij}. Note that $\varepsilon_{ij} = \varepsilon_{ji}^{\star}$ since $\langle \varphi_i \,|\, \varphi_j \rangle = \langle \varphi_j \,|\, \varphi_i \rangle^{*}$. E_0 is the expectation value of the determinant.

$$E_0\left[\{\varphi_i\}\right] = \sum_{i=1}^{N}\left\langle i \,\middle|\, \hat{h} \,\middle|\, j \right\rangle + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(ii|jj) - (ij|ji) \tag{94}$$

where the Mulliken (chemists) notation is being used. Remember this notation since it is quite common and very useful.

$$(ij|kl) = \iint \varphi_i^*(\boldsymbol{\tau_1})\varphi_j(\boldsymbol{\tau_1})\frac{1}{r_{12}}\varphi_k^*(\boldsymbol{\tau_2})\varphi_l(\boldsymbol{\tau_2})\mathrm{d}\boldsymbol{\tau_1}\mathrm{d}\boldsymbol{\tau_2} \tag{95}$$

To minimize the energy, the spin orbitals are varied by an infinitesimal amount

$$\varphi_i \to \varphi_i + \delta\varphi_i. \tag{96}$$

Setting the first-order variation of the Lagrangian to zero yields eq. (97).

$$\delta\mathcal{L} = \delta E_0 - \sum_{i=1}^{N}\sum_{j=1}^{N}\varepsilon_{ij}\delta\langle \varphi_i \,|\, \varphi_j \rangle \tag{97}$$

with

$$\delta\langle \varphi_i \,|\, \varphi_j \rangle = \langle \delta\varphi_i \,|\, \varphi_j \rangle + \langle \varphi_i \,|\, \delta\varphi_j \rangle. \tag{98}$$

---

[21]You see the pattern here?

[22]As I told you before, Langrangian multipliers are extremely convenient to introduce these kinds of constraints

This leads to:

$$\delta E_0 = \sum_{i=1}^{N} \left\langle \delta\varphi_i \left| \hat{h} \right| \varphi_i \right\rangle + \left\langle \varphi_i \left| \hat{h} \right| \delta\varphi_i \right\rangle$$

$$+ \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\delta\varphi_i\varphi_i|\varphi_j\varphi_j) + (\varphi_i\delta\varphi_i|\varphi_j\varphi_j) + (\varphi_i\varphi_i|\delta\varphi_j\varphi_j) + (\varphi_i\varphi_i|\varphi_j\delta\varphi_j)$$

$$- \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\delta\varphi_i\varphi_j|\varphi_j\varphi_i) + (\varphi_i\delta\varphi_j|\varphi_j\varphi_i) + (\varphi_i\varphi_j|\delta\varphi_j\varphi_i) + (\varphi_i\varphi_j|\varphi_j\delta\varphi_i)$$

$$= \sum_{i=1}^{N} \left\langle \delta\varphi_i \left| \hat{h} \right| \varphi_i \right\rangle + \sum_{i=1}^{N} \sum_{j=1}^{N} \left[(\delta\varphi_i\varphi_i|\varphi_j\varphi_j) - (\delta\varphi_i\varphi_j|\varphi_j\varphi_i) \right] + \text{c.c.} \tag{99}$$

and

$$\sum_{ij} \varepsilon_{ji} \left[(\delta\varphi_i|\varphi_j) + (\varphi_i|\delta\varphi_j) \right] = \sum_{ij} \varepsilon_{ji} (\delta\varphi_i|\varphi_j) + \sum_{ij} \varepsilon_{ij} (\varphi_j|\delta\varphi_i)$$

$$= \sum_{ij} \varepsilon_{ji} (\delta\varphi_i|\varphi_j) + \sum_{ij} \varepsilon_{ji}^* (\delta\varphi_i|\varphi_j)^*$$

$$= \sum_{ij} \varepsilon_{ji} (\delta\varphi_i|\varphi_j) + \text{c.c.} \tag{100}$$

Puh, quite a lot. Take your time and try to retrace how we came here. We set the variation in the energy to zero and just have to add a small δ in front of everything. Then we used the braket notation and shuffled some indices around.

Eq. (97) can be further simplified.

$$\delta\mathcal{L} = \sum_{i=1}^{N} \left\langle \delta\varphi_i \left| \hat{h} \right| \varphi_i \right\rangle + \sum_{i=1}^{N} \sum_{j=1}^{N} \left[(\delta\varphi_i\varphi_i|\varphi_j\varphi_j) - (\delta\varphi_i\varphi_j|\varphi_j\varphi_i) \right] - \sum_{ij} \varepsilon_{ji} (\delta\varphi_i|\varphi_j) + \text{c.c.}$$

$$= 0 \tag{101}$$

This is much shorter but still kind of lengthy. We try to simplify as much as possible, especially with the two electron integrals. Defining a Coulomb and exchange operator of the form

$$\hat{\mathcal{J}}_j (1) \varphi_i (1) = \left[\int \varphi_j^* (2) \frac{1}{r_{12}} \varphi_j (2) \, \mathrm{d}\boldsymbol{\tau_2} \right] \varphi_i (1) \tag{102}$$

$$\hat{\mathcal{K}}_j (1) \varphi_i (1) = \left[\int \varphi_j^* (2) \frac{1}{r_{12}} \varphi_i (2) \, \mathrm{d}\boldsymbol{\tau_2} \right] \varphi_j (1) \tag{103}$$

allows rewriting to obtain

$$\delta\mathcal{L} = \sum_{i=1}^{N} \int \delta\varphi_i^* (1) \left[\hat{h} (1) \varphi_i (1) + \sum_{j=1}^{N} \left(\hat{\mathcal{J}}_j (1) - \hat{\mathcal{K}}_j (1) \right) \varphi_i (1) - \sum_{j=1}^{N} \varepsilon_{ji}\varphi_j (1) \right] \mathrm{d}\boldsymbol{\tau_1} + \text{c.c.}$$

$$= 0. \tag{104}$$

Here, the dependency on the coordinates has been abbreviated as $\varphi(\boldsymbol{\tau_1}) \equiv \varphi(1)$.[23]

Since $\delta\varphi_i^*$ is arbitrary, the expression in the brackets has to be zero. This leads to the Fock operator $\hat{f}$.

$$\left[\hat{h}_j(1) + \sum_{j=1}^{N}\left(\hat{\mathcal{J}}_j(1) - \hat{\mathcal{K}}_j(1)\right)\right]\varphi_i(1) = \sum_{j=1}^{N}\varepsilon_{ji}\varphi_j(1)$$

$$\hat{f}\ket{\varphi_i} = \sum_{j=1}^{N}\varepsilon_{ji}\ket{\varphi_j} \tag{105}$$

And there we have our Fock operator! This is a a big part of the mathematics behind the Hartree–Fock method. Now we mainly have to deal with how to actually solve for the fock operator to obtain our orbitals and energies. I know, this section was very heavy on mathematics. You might want to go read this a couple of times to fully grasp it. It took me a long time to understand all the steps behind it, sometimes it is also useful to try and do these mathematical transformations by hand. However, I hope I could explain the main steps here.

3.3   Canonical Hartree–Fock Equations

So, now it is time to further simplify eq. (105). Notice that we have not really talked about the spin orbitals that we used. We just assumed that we somehow know them and used their properties extensively. Well, kind of a reason for that is that these spin orbitals are not unique. Yeah, you heard (or read) right. These orbitals are not unique and we can choose different set of orbitals if we want to. Why is that the case? These orbitals are like the base states. A good analogy for base states are the base vectors that we can use in three dimensional space. We want these vectors to fulfill some properties like being orthogonal to each other and having a length of unity (notice the similarities to the orthonormality?) but we really don't need that. We can rotate around our base vectors, elongate or shorten them and even change the angle between them (as long as they don't become colinear), these base vectors are still a viable choice and we can reach every single point in space with a linear combination of them. The same goes for our spin orbitals. But what kind of spin orbitals do we use then? We just use the ones that are most convennient for our mathematics. This is what we will deal with in the next few sections. After that we will deal with how to actually construct these spin orbitals out of a known set of basis functions.

Let's get started! We will use a unitary transformation of the set of spin orbitals to obtain a new set of spin orbitals. This unitary transformation $\boldsymbol{U}$ preserves the orthonormality of the set.[24] Suppose the wave function is expressed as a Slater-Determinant

$$\ket{\Psi_0} = \frac{1}{\sqrt{N!}}\det(\boldsymbol{M}) \tag{106}$$

where the matrix $\boldsymbol{M}$ contains all spin orbitals in the original set of spin orbitals. Transforming yields $\boldsymbol{M'}$.

$$\boldsymbol{M'} = \boldsymbol{MU} \tag{107}$$

---

[23]Mainly because we are just lazy and don't want to write as much.
[24]This is kind of the definition of a unitary transformation.

Since the determinant of a matrix product equals the product of the determinants of the matrices (multiplicative map), the transformation just gives rise to an additional factor. Using the unitary property of $\boldsymbol{U}$ it can be shown:

$$\det\left(\boldsymbol{M}'\right) = \det\left(\boldsymbol{MU}\right) = \det\left(\boldsymbol{M}\right)\det\left(\boldsymbol{U}\right) \tag{108}$$

Using

$$\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{1} \tag{109}$$

leads to:

$$\det\left(\boldsymbol{U}^\dagger\boldsymbol{U}\right) = \det\left(\boldsymbol{U}^\dagger\right)\det\left(\boldsymbol{U}\right) = \det\left(\boldsymbol{U}\right)^*\det\left(\boldsymbol{U}\right) = \left|\det\left(\boldsymbol{U}\right)\right|^2 = \det\left(\boldsymbol{1}\right) = 1. \tag{110}$$

It follows that

$$\det\left(\boldsymbol{U}\right) = \exp\left(\mathrm{i}\varphi\right) \tag{111}$$

The transformed determinant differs from the original determinant only by a phase factor. If $\boldsymbol{U}$ is real, the factor simplifies to ± 1. This shows that the spin orbitals of a single determinant are not unique and have no particular physical significance. Let that sink in. You often see some plots of these orbitals and how they are spread over the molecule. These orbitals have no physical significance since we could just use another set of orbitals that look different and still describe the exact same quantum mechanical system.

Next, this invariance is used to simplify eq. (105). Only the Coulomb and exchange operator depend on the spin orbitals.

$$\sum_i \hat{\mathcal{J}}_i'(1) = \sum_i \int \varphi_i'^*(2)\frac{1}{r_{12}}\varphi_i'(2)\mathrm{d}\boldsymbol{\tau_2}$$

$$= \sum_{jk}\left[\sum_i U_{ji}^* U_{ki}\right]\int \varphi_j^*(2)\frac{1}{r_{12}}\varphi_k(2)\mathrm{d}\boldsymbol{\tau_2}$$

with

$$\sum_i U_{ji}^* U_{ki} = \left(\boldsymbol{UU}^\dagger\right)_{ki} = \delta_{ki} \tag{112}$$

and therefore

$$\sum_i \hat{\mathcal{J}}_i'(1) = \int \varphi_j^*(2)\frac{1}{r_{12}}\varphi_j(2)\mathrm{d}\boldsymbol{\tau_2} = \sum_j \hat{\mathcal{J}}_j(1). \tag{113}$$

The sum of the Coulomb operator is therefore invariant under unitary transformation. The reasoning for the exchange operator is analogous. This shouldn't really be a surprise here but it is nevertheless good to see that the math works as we want it to.

Next, consider the Lagrange multiplier ε_{ij}.

$$\left\langle \varphi_k \left| \hat{f} \right| \varphi_i \right\rangle = \sum_j \varepsilon_{ji}\left\langle \varphi_k \,|\, \varphi_i \right\rangle = \varepsilon_{ki} \tag{114}$$

Now applying this to the transformed spin orbitals:

$$\varepsilon_{ij}' = \int \varphi_i'^*(1)\hat{f}(1)\varphi_j'(1)\mathrm{d}\boldsymbol{\tau_1}$$

$$= \sum_{kl} U_{ki}^* U_{lj}\int \varphi_k^*(1)\hat{f}(1)\varphi_l(1)\mathrm{d}\boldsymbol{\tau_1} \tag{115}$$

$$= \sum_{kl} U_{ki}^* \varepsilon_{kl} U_{lj}$$

This can also be written in the matrix form.

$$\boldsymbol{\varepsilon}' = \boldsymbol{U}^\dagger \boldsymbol{\varepsilon} \boldsymbol{U} \tag{116}$$

$\boldsymbol{\varepsilon}$ is hermitian and it is therefore always possible to find a unitary transformation which diagonalizes $\boldsymbol{\varepsilon}$.[25] Note that up to this point the matrix $\boldsymbol{U}$ has not been defined.[26]

To summarize, the spin orbitals can be transformed using a unitary transformation in a way which results in a diagonal matrix of Lagrange multipliers. This simplifies eq. (105) to

$$\hat{f} \left|\varphi_i'\right\rangle = \varepsilon_i' \left|\varphi_i'\right\rangle. \tag{117}$$

The set of spin orbitals is called canonical spin orbitals. They are delocalized and show the symmetry of the molecule (the irreducible representations of the point group).[27] In the following, the prime symbol will be dropped.

$$\hat{f} \left|\varphi_i\right\rangle = \varepsilon_i \left|\varphi_i\right\rangle \tag{118}$$

3.4   Restricted Closed-Shell Hartree–Fock

A restricted set of spin orbitals can be constructed as a simple product between a spatial orbital and a spin function (see eq. (**??**)). For the closed-shell case, every spatial orbital is either doubly occupied or not occupied at all. This allows dropping the spin index and only working with the spatial orbitals.[28] In order to do that, the spin functions have to be integrated out.
Starting with the main result from the previous section:

$$\hat{f}\left(\boldsymbol{\tau_1}\right)\varphi_i\left(\boldsymbol{\tau_1}\right) = \varepsilon_i\varphi_i\left(\boldsymbol{\tau_1}\right). \tag{119}$$

Assuming the spin orbital $\varphi_i\left(\boldsymbol{\tau_i}\right)$ has a spin function α (the same reasoning goes for β), multiplying by α^* from the left and integrating leads to:

$$\left[\int \alpha^*(\sigma_1)\hat{f}\left(\boldsymbol{\tau_1}\right)\alpha(\sigma_1)\mathrm{d}\sigma_1\right]\phi_j\left(\boldsymbol{r_1}\right) = \varepsilon_j\phi_j\left(\boldsymbol{r_1}\right). \tag{120}$$

The Fock operator can be written as:

$$\hat{f}\left(\boldsymbol{\tau_1}\right) = \hat{h}\left(\boldsymbol{r_1}\right) + \sum_c^N \hat{\mathcal{J}}_c\left(\boldsymbol{\tau_1}\right) - \hat{\mathcal{K}}_c\left(\boldsymbol{\tau_1}\right). \tag{121}$$

Defining a closed-shell Fock operator leads to:

$$\hat{f}\left(\boldsymbol{r_1}\right) = \int \alpha^*\left(\sigma_1\right)\hat{f}\left(\boldsymbol{\tau_1}\right)\alpha\left(\sigma_1\right)\mathrm{d}\sigma_1. \tag{122}$$

---

[25]This is just a property of hermitian matrices. Accept it or look it up and accept it then.

[26]Yep, we like to keep things as open as possible and only really figure them out once we need it.

[27]I want to repeat: These orbitals have in theory no physical significance. Or in other words: they are as signifcant as any other set of spin orbitals that we can get by any other orthonormal transformation from the original set of spin orbitals.

[28]We just want our lives to be easier. Although there is of course an unrestricted case that is actually not that much more difficult. I think it is just easier to start with the restricted one and maybe later move to the unrestricted case

Carrying out the integration over the spin coordinate σ_1 in the single electron operator $\hat{h}$ and writing out the Coulomb and exchange operator eq. (123) is obtained.

$$
\begin{aligned}
\hat{f}\left(\boldsymbol{r_1}\right)\phi_j\left(\boldsymbol{r_1}\right) =&\hat{h}\left(\boldsymbol{r_1}\right)\phi_j\left(\boldsymbol{r_1}\right) \\
&+\sum_i^N \int \alpha^*\left(\sigma_1\right)\varphi_c^*\left(\boldsymbol{\tau_2}\right)\frac{1}{r_{12}}\varphi_i\left(\boldsymbol{\tau_2}\right)\alpha\left(\sigma_1\right)\phi_j\left(\boldsymbol{r_1}\right)\mathrm{d}\sigma_1\mathrm{d}\boldsymbol{\tau_2} \\
&-\sum_i^N \int \alpha^*\left(\sigma_1\right)\varphi_c^*\left(\boldsymbol{\tau_2}\right)\frac{1}{r_{12}}\varphi_i\left(\boldsymbol{\tau_1}\right)\alpha\left(\sigma_2\right)\phi_j\left(\boldsymbol{r_2}\right)\mathrm{d}\sigma_1\mathrm{d}\boldsymbol{\tau_2} \\
=&\varepsilon_j\phi_j\left(\boldsymbol{r_1}\right)
\end{aligned}
\tag{123}
$$

For the closed-shell case, the sum can be split in two parts: One sums over all electrons in an orbital with an α spin function and the other sums over the electrons with a β spin function. This allows for the splitting of the integration over $\mathrm{d}\boldsymbol{\tau}$ in two parts $\mathrm{d}\boldsymbol{r}\mathrm{d}\sigma$.

$$
\begin{aligned}
\hat{f}\left(\boldsymbol{r_1}\right)\phi_j\left(\boldsymbol{r_1}\right) =&\hat{h}\left(\boldsymbol{r_1}\right)\phi_j\left(\boldsymbol{r_1}\right) \\
&+\sum_i^{N/2} \int \alpha^*\left(\sigma_1\right)\phi_i^*\left(\boldsymbol{r_2}\right)\alpha^*\left(\sigma_2\right)\frac{1}{r_{12}}\phi_i\left(\boldsymbol{r_2}\right)\alpha\left(\sigma_2\right)\alpha\left(\sigma_1\right)\phi_j\left(\boldsymbol{r_1}\right)\mathrm{d}\sigma_1\mathrm{d}\sigma_2\mathrm{d}\boldsymbol{\tau_2} \\
&+\sum_i^{N/2} \int \alpha^*\left(\sigma_1\right)\phi_i^*\left(\boldsymbol{r_2}\right)\beta^*\left(\sigma_2\right)\frac{1}{r_{12}}\phi_i\left(\boldsymbol{r_2}\right)\beta\left(\sigma_2\right)\alpha\left(\sigma_1\right)\phi_j\left(\boldsymbol{r_1}\right)\mathrm{d}\sigma_1\mathrm{d}\sigma_2\mathrm{d}\boldsymbol{\tau_2} \\
&-\sum_i^{N/2} \int \alpha^*\left(\sigma_1\right)\phi_i^*\left(\boldsymbol{r_2}\right)\alpha^*\left(\sigma_2\right)\frac{1}{r_{12}}\phi_i\left(\boldsymbol{r_1}\right)\alpha\left(\sigma_1\right)\alpha\left(\sigma_2\right)\phi_j\left(\boldsymbol{r_2}\right)\mathrm{d}\sigma_1\mathrm{d}\sigma_2\mathrm{d}\boldsymbol{\tau_2} \\
&-\sum_i^{N/2} \int \alpha^*\left(\sigma_1\right)\phi_i^*\left(\boldsymbol{r_2}\right)\beta^*\left(\sigma_2\right)\frac{1}{r_{12}}\phi_i\left(\boldsymbol{r_1}\right)\beta\left(\sigma_1\right)\alpha\left(\sigma_2\right)\phi_j\left(\boldsymbol{r_2}\right)\mathrm{d}\sigma_1\mathrm{d}\sigma_2\mathrm{d}\boldsymbol{\tau_2} \\
=&\varepsilon_j\phi_j\left(\boldsymbol{r_1}\right)
\end{aligned}
\tag{124}
$$

Now the integration over both spin coordinates can be carried out. First, note that the last sum equals zero due to the spin functions being orthogonal.[29] The first two Coulomb sums are identical and can be summarized:

$$
\begin{aligned}
\hat{f}\left(\boldsymbol{r_1}\right)\phi_j\left(\boldsymbol{r_1}\right) =&\hat{h}\left(\boldsymbol{r_1}\right)\phi_j\left(\boldsymbol{r_1}\right) \\
&+\left[2\sum_i^{N/2} \int \phi_i^*\left(\boldsymbol{r_2}\right)\frac{1}{r_{12}}\phi_i\left(\boldsymbol{r_2}\right)\mathrm{d}\boldsymbol{r_2}\right]\phi_j\left(\boldsymbol{r_1}\right) \\
&-\left[\sum_i^{N/2} \int \phi_i^*\left(\boldsymbol{r_2}\right)\frac{1}{r_{12}}\phi_j\left(\boldsymbol{r_2}\right)\mathrm{d}\boldsymbol{r_2}\right]\phi_i\left(\boldsymbol{r_1}\right).
\end{aligned}
\tag{125}
$$

We define a Coulomb and exchange operator for the closed-shell case as:

$$
\hat{\mathcal{J}}_i(1)\phi_j\left(\boldsymbol{r_1}\right) = \left[\int \phi_i^*\left(\boldsymbol{r_2}\right)\frac{1}{r_{12}}\phi_i\left(\boldsymbol{r_2}\right)\mathrm{d}\boldsymbol{r_2}\right]\phi_j\left(\boldsymbol{r_1}\right)
\tag{126}
$$

$$
\hat{\mathcal{K}}_i(1)\phi_j\left(\boldsymbol{r_1}\right) = \left[\int \phi_i^*\left(\boldsymbol{r_2}\right)\frac{1}{r_{12}}\phi_j\left(\boldsymbol{r_2}\right)\mathrm{d}\boldsymbol{r_2}\right]\phi_i\left(\boldsymbol{r_1}\right)
\tag{127}
$$

---

[29]This shows that Hartree–Fock theory only takes exchange between electrons with equal spin into account.

This simplifies the above expression:

$$\hat{f}(1) = \hat{h}(1) + \sum_{i}^{N/2} 2\hat{\mathcal{J}}_i(1) - \hat{\mathcal{K}}_i(1). \tag{128}$$

The closed-shell ground state energy therefore becomes:

$$E_0 = 2\sum_{i}^{N/2} h_i + \sum_{i}^{N/2}\sum_{j}^{N/2} 2J_{ij} - K_{ij}. \tag{129}$$

This is a nice result! We got rid of the spin and already have a nice epxression for our energy. Now we need to take care of the matrix elements but first we need to talk about the molecular orbitals and how to construct them.

3.5 The Roothaan–Hall Equations

After eliminating the spin for the closed-shell case, the construction of spatial orbitals is of interest. Roothaan [3] and Hall [4] algebraized this problem by introducing a set of known basis functions $\{\chi_\mu\}$. The unknown molecular orbitals can be expressed as a linear expansion.

$$\phi_i = \sum_{\mu=1}^{K} C_{\mu i}\chi_\mu \quad \text{with} \quad i = 1, 2, \ldots, K \tag{130}$$

Again we construct something unknown out of a set of known orbitals with a linear combination. This approach allows for an easy computation and implementation of the obtained equations. Above expansion is used to solve

$$\hat{f}(\boldsymbol{r_1})\phi_i(\boldsymbol{r_1}) = \varepsilon_i\phi_i(\boldsymbol{r_1}). \tag{131}$$

Substitution yields

$$\hat{f}(\boldsymbol{r_1})\sum_{\nu=1}^{K} C_{\nu i}\chi_\nu(\boldsymbol{r_1}) = \varepsilon_i\sum_{\nu=1}^{K} C_{\nu i}\chi_\nu(\boldsymbol{r_1}). \tag{132}$$

Multiplying with $\chi_\mu^*(\boldsymbol{r_1})$ from the left and integrating over $\boldsymbol{r_1}$ eq. (133) is obtained.

$$\sum_{\nu} C_{\nu i}\int \chi_\mu^*(\boldsymbol{r_1})\hat{f}(\boldsymbol{r_1})\chi_\nu(\boldsymbol{r_1})\,\mathrm{d}\boldsymbol{r_1} = \varepsilon_i\sum_{\nu} C_{\nu i}\int \chi_\mu^*(\boldsymbol{r_1})\chi_\nu(\boldsymbol{r_1})\,\mathrm{d}\boldsymbol{r_1} \tag{133}$$

Eq. (133) can be rewritten as a matrix equation. This can be done by introducing the overlap matrix $\boldsymbol{S}$ and the Fock matrix $\boldsymbol{F}$ with the elements:

$$S_{\mu\nu} = \int \chi_\mu^*(\boldsymbol{r_1})\chi_\nu(\boldsymbol{r_1})\,\mathrm{d}\boldsymbol{r_1} = \langle \chi_\mu \,|\, \chi_\nu \rangle, \tag{134}$$

$$F_{\mu\nu} = \int \chi_\mu^*(\boldsymbol{r_1})\hat{f}(\boldsymbol{r_1})\chi_\nu(\boldsymbol{r_1})\,\mathrm{d}\boldsymbol{r_1} = \left\langle \chi_\mu \,\middle|\, \hat{f} \,\middle|\, \chi_\nu \right\rangle. \tag{135}$$

From the definition it is evident that both matrices are hermitian. This ensures that the matrices can be diagonalized which will be necessary in the Hartree–Fock SCF scheme. If the integrals are real, the matrices are symmetric.
The Roothaan–Hall equations are obtained by rewriting in the matrix form.

$$\boldsymbol{FC} = \boldsymbol{SC\varepsilon} \tag{136}$$

The matrix C contains all the coefficients for the linear combination of the basis functions. The matrix ε is a diagonal matrix which contains the orbital energies. This result is quite fundamental and you should definitely memorize it as it is the core of the program we are going to write. Additionally, the equation will be used in other more sophisticated methods.

So what are we missing? We need an explicit expression for the Fock matrix so we can solve the Roothaan–Hall equations which will give use the coefficients C (and thereby the molecular orbitals) as well as the orbital energies ε and of course the total Hartree–Fock energy.

3.6 The Fock Matrix

In order to find an expression for the Fock matrix, it is convenient to first define a density matrix P.[30]
We aim to calculate the electron density at any point in space. The density can simply be written in terms of the molecular orbitals as:

$$\rho\left(\boldsymbol{r}\right) = 2 \sum_{i}^{occ} \phi_i^*\left(\boldsymbol{r}\right) \phi_i\left(\boldsymbol{r}\right). \tag{137}$$

where the summation is carried out over all occupied orbitals. The next goal is to express the density in terms of the basis functions. Substituting eq. (130) in above expression yields:

$$
\begin{aligned}
\rho\left(\boldsymbol{r}\right) &= 2 \sum_{i}^{occ} \sum_{\nu} C_{\nu i}^* \phi_\nu^*\left(\boldsymbol{r}\right) \sum_{\mu} C_{\mu i} \phi_\mu\left(\boldsymbol{r}\right) \\
&= \sum_{\mu\nu} \left[2 \sum_{i}^{occ} C_{\nu i}^* C_{\mu i} \right] \phi_\nu^*\left(\boldsymbol{r}\right) \phi_\mu\left(\boldsymbol{r}\right) \\
&= \sum_{\mu\nu} P_{\mu\nu} \phi_\nu^*\left(\boldsymbol{r}\right) \phi_\mu\left(\boldsymbol{r}\right)
\end{aligned}
\tag{138}
$$

where the elements of the density matrix P are defined as:

$$P_{\mu\nu} = 2 \sum_{i}^{occ} C_{\nu i}^* C_{\mu i}. \tag{139}$$

Since the coefficients of the expansion are usually real, the complex conjugation of the first coefficient can be omitted.
Important note: The density matrix is sometimes called D and is often defined without the factor 2 which comes from the unrestricted case that we are dealing with here. Whenever you read a paper or a tutorial make sure to get the definition of the density matrix or else your results will be different/wrong. Most common is the 'unrestricted' defition of the density matrix P (which is simply our defintion divided by two). However, one just has to be consistent throughout to obtain the correct results.

After defining the density matrix P, the next goal is to find an expression for the Fock matrix. In the last section the following expression for the Fock operator was obtained:

$$\hat{f}(1) = \hat{h}(1) + \sum_{i}^{occ} 2\hat{\mathcal{J}}_i(1) - \hat{\mathcal{K}}_i(1) \tag{140}$$

---

[30]We will later see why it is convenient here.

Applying this to the set of basis functions $\{\chi_\mu\}$ the matrix elements become:

$$
\begin{aligned}
F_{\mu\nu} &= \int \chi_\mu^* \left(\boldsymbol{r_1}\right) \hat{f} \left(\boldsymbol{r_1}\right) \chi_\nu \left(\boldsymbol{r_1}\right) \mathrm{d}\boldsymbol{r_1} \\
&= \int \chi_\mu^* \left(\boldsymbol{r_1}\right) \hat{h} \left(\boldsymbol{r_1}\right) \chi_\nu \left(\boldsymbol{r_1}\right) \mathrm{d}\boldsymbol{r_1} + \sum_i^{\text{occ}} \int \chi_\mu^* \left(\boldsymbol{r_1}\right) \left[2\hat{\mathcal{J}}_i \left(\boldsymbol{r_1}\right) - \hat{\mathcal{K}}_i \left(\boldsymbol{r_1}\right)\right] \chi_\nu \left(\boldsymbol{r_1}\right) \mathrm{d}\boldsymbol{r_1} \quad (141) \\
&= H_{\mu\nu}^{\text{core}} + \sum_i^{\text{occ}} 2 \left(\mu\nu|ii\right) - \left(\mu i|i\nu\right)
\end{aligned}
$$

where the core Hamiltonian $\boldsymbol{H}^{\textbf{core}}$ describes the one electron integrals.

$$
\begin{aligned}
H_{\mu\nu}^{\text{core}} &= T_{\mu\nu} + V_{\mu\nu} \\
&= \int \chi_\mu^* \left(\boldsymbol{r_1}\right) \left[-\frac{1}{2}\nabla^2\right] \chi_\nu \left(\boldsymbol{r_1}\right) \mathrm{d}\boldsymbol{r_1} + \int \chi_\mu^* \left(\boldsymbol{r_1}\right) \left[-\sum_A \frac{Z_a}{r_{iA}}\right] \chi_\nu \left(\boldsymbol{r_1}\right) \mathrm{d}\boldsymbol{r_1}
\end{aligned}
\quad (142)
$$

To complete the expression simple insertion of the linear expansion from eq. (130) into eq. (141) is necessary.

$$
\begin{aligned}
F_{\mu\nu} &= H_{\mu\nu}^{\text{core}} + \sum_i^{\text{occ}} \sum_{\lambda\sigma} C_{\lambda i} C_{\sigma i}^* \left[2 \left(\mu\nu|\sigma\lambda\right) - \left(\mu\lambda|\sigma\nu\right)\right] \\
&= H_{\mu\nu}^{\text{core}} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[\left(\mu\nu|\sigma\lambda\right) - \frac{1}{2}\left(\mu\lambda|\sigma\nu\right)\right] \\
&= H_{\mu\nu}^{\text{core}} + J_{\mu\nu} + K_{\mu\nu}
\end{aligned}
\quad (143)
$$

Here, the Coulomb matrix $\boldsymbol{J}$ and exchange matrix $\boldsymbol{K}$ are defined as follows:

$$
J_{\mu\nu} = \sum_{\lambda\sigma} P_{\lambda\sigma} \left(\mu\nu|\sigma\lambda\right) \tag{144}
$$

$$
K_{\mu\nu} = -\frac{1}{2}\sum_{\lambda\sigma} P_{\lambda\sigma} \left(\mu\lambda|\sigma\nu\right) \tag{145}
$$

Sometimes the sum of the Coulomb and exchange matrix is called $\boldsymbol{G}$. This matrix contains the two-electron part of the Fock matrix.

We have obtained an explicit expression for the Fock matrix and and now it is evident why the Roothaan–Hall equations have to be solved iteratively. The Fock matrix $\boldsymbol{F}$ depends on the density matrix $\boldsymbol{P}$ (see eq. (143)) which can be constructed from the coefficient matrix $\boldsymbol{C}$ (eq. (139)).

$$
\boldsymbol{F} = \boldsymbol{F}\left(\boldsymbol{C}\right) \tag{146}
$$

The Roothaan–Hall equations are therefore nonlinear and have to be solved iteratively:

$$
\boldsymbol{F}\left(\boldsymbol{C}\right)\boldsymbol{C} = \boldsymbol{S}\boldsymbol{C}\boldsymbol{\varepsilon} \tag{147}
$$

Now we are just missing one thing: the unitary transformation.

3.7 Orthogonalization of Basis Sets

First, consider the overlap Matrix $\boldsymbol{S}$: in an orthogonal basis it would simply be the unit matrix. This would simplify the Roothaan–Hall equations to

$$\boldsymbol{FC} = \boldsymbol{C\varepsilon} \tag{148}$$

which can be solved easily (usual matrix eigenvalue problem for which a lot of efficient algorithms exist). Transformation of the normalized basis set $\{\chi_\mu\}$ to an orthonormal set $\{\chi'_\mu\}$ with a unitary transformation matrix $\boldsymbol{X}$ can be achieved as seen in eq. (149).

$$\chi'_\mu = \sum_\nu X_{\nu\mu}\chi_\nu \tag{149}$$

In this orthonormal set the orthonormality is conserved.

$$\int \chi'^*_\mu(\boldsymbol{r})\,\chi'_\nu(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} = \delta_{\mu\nu} \tag{150}$$

Substituting the transformation in eq. (149) into the above expression yields:

$$\begin{aligned}
\int \chi'^*_\mu(\boldsymbol{r})\,\chi'_\nu(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r} &= \int \left[\sum_\lambda X^*_{\lambda\mu}\chi^*_\lambda(\boldsymbol{r})\right]\left[\sum_\sigma X_{\sigma\nu}\chi_\sigma(\boldsymbol{r})\right]\mathrm{d}\boldsymbol{r}\\
&= \sum_\lambda\sum_\sigma X^*_{\lambda\mu}\int \chi^*_\lambda(\boldsymbol{r})\,\chi_\sigma(\boldsymbol{r})\,X_{\sigma\nu}\mathrm{d}\boldsymbol{r}\\
&= \sum_\lambda\sum_\sigma X^*_{\lambda\mu}S_{\lambda\sigma}X_{\sigma\nu}\\
&= \delta_{\mu\nu}.
\end{aligned} \tag{151}$$

In matrix form this expression can simply be written as:

$$\boldsymbol{X}^\dagger \boldsymbol{S}\boldsymbol{X} = \boldsymbol{1} \tag{152}$$

There are different ways to obtain the transformation matrix $\boldsymbol{X}$. The most common way is called symmetric orthogonalization (or Löwdin orhtogonalization) where the inverse square root of $\boldsymbol{S}$ is used for $\boldsymbol{X}$.

$$\boldsymbol{X} \equiv \boldsymbol{S}^{-1/2} \tag{153}$$

From eq. (152) it is evident that this definition fulfills the properties of the transformation. To obtain $\boldsymbol{S}^{-1/2}$, an initial diagonalization of $\boldsymbol{S}$ using a unitary transformation $\boldsymbol{U}$ is done.

$$\boldsymbol{U}^\dagger \boldsymbol{S}\boldsymbol{U} = \boldsymbol{s} \tag{154}$$

The matrix $\boldsymbol{s}$ contains the eigenvalues on its diagonal. The matrix $\boldsymbol{s}^{-1/2}$ is formed by taking the inverse square root of the eigenvalues. Then a simple transformation is necessary to obtain $\boldsymbol{S}^{-1/2}$.

$$\boldsymbol{S}^{-1/2} = \boldsymbol{U}\boldsymbol{s}^{-1/2}\boldsymbol{U}^\dagger \tag{155}$$

Using the transformation simplifies the Roothaan–Hall equations. Introduce a new coefficient matrix $\boldsymbol{C}'$:

$$\boldsymbol{C}' = \boldsymbol{X}^{-1}\boldsymbol{C} \quad \text{and} \quad \boldsymbol{C} = \boldsymbol{X}\boldsymbol{C}' \tag{156}$$

Substituting this into the Roothaan–Hall equations yields:

$$\boldsymbol{FXC'} = \boldsymbol{SXC'\varepsilon}. \tag{157}$$

Multiplying with $\boldsymbol{X^\dagger}$ from the left and introducing the transformed Fock matrix $\boldsymbol{F'} = \boldsymbol{X^\dagger F X}$:

$$\boldsymbol{X^\dagger FXC'} = \boldsymbol{X^\dagger SXC'\varepsilon} \tag{158}$$

$$\boldsymbol{F'C'} = \boldsymbol{C'\varepsilon} \tag{159}$$

The transformed Roothaan–Hall equations are obtained which can be solved by diagonalizing $\boldsymbol{F'}$. When solved for $\boldsymbol{C'}$ transforming back to $\boldsymbol{C}$ (eq. (156)) is simple.

3.8   Recap of the Hartee–Fock scheme

Yes, that was it. We are done, we've got everthing we need to know to solve the Roothaan–Hall equations. You may be a bit confused because there was a quite a bit of maths and it isn't easy to see it, but here is a quick outline of the Hartree–Fock-SCF procedure.

1. Choose a molecule, its structure $\{R_i\}$ and an appropriate basis set $\{\chi_\mu\}$

2. Calculate the required integrals and save them:

$$S_{\mu\nu} = \langle \mu \,|\, \nu \rangle \tag{160}$$

$$T_{\mu\nu} = \left\langle \mu \,\left|\, -\frac{1}{2}\nabla^2 \,\right|\, \nu \right\rangle \tag{161}$$

$$V_{\mu\nu} = \left\langle \mu \,\left|\, \sum_A \frac{1}{|\boldsymbol{r} - \boldsymbol{R_A}|} \,\right|\, \nu \right\rangle \tag{162}$$

$$(\mu\nu|\lambda\sigma) = \iint \chi_\mu^*(\boldsymbol{r_1})\chi_\nu(\boldsymbol{r_1})\frac{1}{r_{12}}\chi_\lambda^*(\boldsymbol{r_2})\chi_\sigma(\boldsymbol{r_2})\mathrm{d}\boldsymbol{r_1}\mathrm{d}\boldsymbol{r_1} \tag{163}$$

3. Obtain the transformation matrix $\boldsymbol{X} = \boldsymbol{S}^{-1/2}$

4. Guess an initial density matrix $\boldsymbol{P}$. Here a zero-density guess $\boldsymbol{P} = \boldsymbol{0}$ can be used

5. Build a Fock Matrix $\boldsymbol{F}$

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + J_{\mu\nu} + K_{\mu\nu} \tag{164}$$

   with

$$H_{\mu\nu}^{\text{core}} = T_{\mu\nu} + V_{\mu\nu} \tag{165}$$

$$J_{\mu\nu} = \sum_{\lambda\sigma} P_{\lambda\sigma}\,(\mu\nu|\sigma\lambda) \tag{166}$$

$$K_{\mu\nu} = -\frac{1}{2}\sum_{\lambda\sigma} P_{\lambda\sigma}\,(\mu\lambda|\sigma\nu) \tag{167}$$

6. Transform the Fock matrix

$$\boldsymbol{F'} = \boldsymbol{X^\dagger F X} \tag{168}$$

7. Diagonalize $\boldsymbol{F'}$ to obtain a coefficient matrix $\boldsymbol{C'}$ and energy matrix $\boldsymbol{\varepsilon}$

$$\boldsymbol{F'C'} = \boldsymbol{C'\varepsilon} \tag{169}$$

8. Transform the coefficients

$$\boldsymbol{C} = \boldsymbol{X}\boldsymbol{C}' \tag{170}$$

9. Construct a new density matrix $\boldsymbol{P}$ from the coefficient matrix $\boldsymbol{C}$

$$P_{\mu\nu} = 2 \sum_i^{\text{occ}} C_{\nu i}^* C_{\mu i} \tag{171}$$

10. Check convergence. The easiest way is to determine the energy difference to the previous iteration

$$E_{\text{HF}} = \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \left(H_{\mu\nu}^{\text{core}} + F_{\mu\nu} \right) \tag{172}$$

- if converged $\rightarrow$ HF-SCF scheme finished
- if not converged $\rightarrow$ Go to 5.

References

(1) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry*, 1996.

(2) Helgaker, T.; Jorgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory*, 2013.

(3) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

(4) Hall, G. G.; Lennard-Jones, J. E. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1951**, *205*, 541–552.