

From quantum mechanical simulations of atoms and molecules to molecular dynamics

by

Milad Hobbi Mobarhan

THESIS
for the degree of
MASTER OF SCIENCE



Faculty of Mathematics and Natural Sciences
University of Oslo

August 30, 2013

Contents

I	Theory	5
1	Hartree-Fock Molecular Orbital Theory	7
1.1	Key Features of Quantum Mechanics	8
1.1.1	Hamiltonian Operator	9
1.1.2	The Variational Principle	10
1.1.3	The Born-Oppenheimer Approximation	12
1.2	Construction of Trial Wave Functions	13
1.2.1	The LCAO Basis Set Approach	14
1.3	Many-electron Wave Functions	16
1.3.1	Hartree Wave Function	17
1.3.2	Electron Spin and Antisymmetry	19
1.3.3	Slater Determinant	20
1.4	The Hartree-Fock Theory	24
1.4.1	Derivation of the Hartree-Fock Equation	24
1.4.2	Hartree-Fock Equation - Physical Picture	30
1.5	Closed- and Open-shell Systems	32
1.5.1	Restricted Hartree-Fock	32
1.5.2	LCAO Approach: The Roothaan Equation	35
1.5.3	Orthogonalization of the Basis	38
1.5.4	Unrestricted Hartree-Fock	40
1.5.5	LCAO Approach: Pople-Nesbet Equations	42
	Appendices	
A	Mathematics	47
A.1	Atomic units	47
A.2	Antisymmetrizer	48

List of Figures

1.1	HF spectrum: schematic representation on how the levels are filled for (a): the ground state of an even number of electrons, (b): the ground state of an odd number of electrons and (c): an excited spectrum of an even number of electrons.	31
1.2	Restricted and unrestricted spectrum: schematic representation of (a): spin restricted configuration and (b): spin unrestricted configuration.	33

List of Tables

Part I

Theory

Chapter 1

Hartree-Fock Molecular Orbital Theory

In molecules, electrons are shared among individual atoms to form chemical bonds, which will determine the structure of the molecule. But electrons are quantum mechanical particles and must be treated by the laws of quantum mechanics. Therefore, in order to describe the structure of molecules, we will need a theory where the quantum mechanical properties of the electrons are incorporated. Molecular orbital (MO) theory is one such theory, where the wave like characteristics of electrons is included. In this theory the electrons are no longer deterministically given defined coordinates, but their position is described according to a probability distribution function defining all the possible positions of the electrons. Moreover, the electrons are not assigned to individual bonds between atoms, but instead treated as moving under influence of the nuclei in the whole molecule. That is, they are "smeared out" across the molecule.

The basis of MO theory is the Hartree-Fock (HF) method, which is the cornerstone of *ab initio* electronic structure theory. The importance of HF method cannot be overemphasized in quantum chemistry and investigating this method in detail will provide a good understanding of the many-electron problem. Not only provide the HF method a way to handle the many-electron problem, but it is also a central starting point for more sophisticated and accurate methods, such as the non-variational Møller-Plesset and coupled-cluster method. Although, the HF method by itself is not so useful for high accurate studies, it is still the only method that can be applied on large molecular systems, and therefore a natural method to use in *ab initio* molecular dynamics schemes.

In this chapter we will first give an overview of the fundamental features of quantum mechanics, and thereafter gradually present how to handle the

many-electron problem, using the Hartree method and later the Hartree-Fock method. The fundamental equations arising from the HF method, will be derived and discussed, in addition to the application on closed- and open shell systems, using the restricted and unrestricted HF method. Note that, the overview of quantum mechanics is superficial, and only the central points in later discussion is included. For more detailed description of the quantum theory, the reader is referred to standard quantum mechanics books such as [1] written by D. Griffiths and [2] written by R. Shankar.

The first part of this chapter is heavily influenced by chapter 4 in [3], written by Christopher J. Cramer. The derivation of the HF equation is done in line with the derivation given in chapter 4 in [4], written by J. Thijssen, while the discussion of restricted and unrestricted HF is inspired by chapter 3 in [5], written by A. Szabo and N. S. Ostlund.

1.1 Key Features of Quantum Mechanics

The foundation of quantum mechanics (QM) are the postulates; theoretical principles based on experimental observations which the applications of QM are built on. The fundamental idea in quantum theory is that any (chemical) system can be described by a wave function, Ψ , and any measurable property of the system can be obtained by letting appropriate operator act on the wave function. Mathematically this can be written as

$$\mathcal{O}\Psi = o\Psi \tag{1.1}$$

where \mathcal{O} is an operator and o is a scalar value for some property of the system. This equation is nothing else than an eigenvalue problem, if think of Ψ as an N -element column vector and \mathcal{O} as an $N \times N$ square matrix. This analogy hints that linear algebra may have central role in QM and this is indeed the case. Linear algebra is absolutely necessarily in order to express QM in its general form. It is the mathematical language of quantum mechanics.

The physical interpretation of the wave function is that the product of it with its complex conjugate (i.e. $|\Psi^*\Psi|$) represents the probability density for a system to be found within some region of a multi-dimensional space. This probability interpretation of the wave function, certainly place a few requirements on the wave function. One of the requirements is that the wave function has to be normalizable, i.e. the integral of $|\Psi^*\Psi|$ over all space must be 1. This has to be the case because the system has to be found somewhere. Without this, the probability interpretation of the wave function would be nonsense. In addition to normalizable, Ψ has to be single valued and continuous.

As written in the book of Christopher J. Cramer ([3]), the best description of Ψ at this point is that it is an oracle; when queried with questions by an operator, it returns answers. The big question now is how the wave function looks like, which indeed will depend on the system we are studying. Unfortunately, the form of the wave function is often not known and without the wave function we are not able to measure any property of the system. However, there exist several theories on how to approximate the wave function of a many-body system. We will return to this later in this chapter.

1.1.1 Hamiltonian Operator

Associated with each measurable parameter in a system, is a quantum mechanical operator. One very central operator in QM is the Hamiltonian operator, \mathcal{H} , because it returns the energy, E , of the system as an eigenvalue when Eq. (1.1) holds;

$$\mathcal{H}\Psi = E\Psi. \quad (1.2)$$

This equation is the time-independent or stationary Schrödinger equation and a lot of QM is about solving this equation for systems of interest.

We will mainly be concerned with molecular systems and a typical Hamiltonian for a system consisting of N_e electrons and N_n nuclei with charges Z_n reads

$$\begin{aligned} \mathcal{H} = & - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{n=1}^{N_n} \frac{\hbar^2}{2M_n} \nabla_n^2 + \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \\ & - \frac{1}{4\pi\epsilon_0} \sum_{n=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|} - \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{\substack{n,m=1 \\ n \neq m}}^{N_n} \frac{Z_n Z_m e^2}{|\mathbf{R}_n - \mathbf{R}_m|} \end{aligned} \quad (1.3)$$

The indices i and j refer to the electrons while n and m refer to the nuclei, m_e is the electron mass and M_n is the mass of nucleus n . The first two terms represent the kinetic energy of the electrons and nuclei, while the third and fourth term represent the Coulomb repulsion between the electrons and the Coulomb attraction between the electrons and nuclei, respectively. Finally, the last term contains the Coulomb repulsion between the nuclei. Note that this Hamiltonian depends on the set of positions $\{\mathbf{r}_i\}$ and $\{\mathbf{R}_n\}$ of the electrons and nuclei, respectively. Thus, we expect the wave function of the system also to depend on the positions of electrons and nuclei.

For a molecular system, the eigenvalue problem in Eq. (1.2) can in general have many solutions Ψ_i , each with associated eigenvalue E_i . These solutions will in fact form a complete basis, which we can, without loss of generality, assume to be orthonormal. Mathematically this means that

$$\int \Psi_i^* \Psi_j d\mathbf{r} = \delta_{i,j}, \quad (1.4)$$

where the integral is taken over a generalized $3(N_e + N_n)$ -dimensional volume element $d\mathbf{r}$. Note that $\mathbf{r} = \mathbf{r}(\{\mathbf{r}_i\}, \{\mathbf{R}_n\})$ and must not be confused with the electronic coordinate \mathbf{r}_i .

The question now is how to find the energy E_j associated with Ψ_j . We consider Eq. (1.2) for a specific Ψ_j , and multiply with a general Ψ_i^* from left and integrate over the generalized space, which gives us

$$\int \Psi_i^* \mathcal{H} \Psi_j d\mathbf{r} = E_j \delta_{i,j}. \quad (1.5)$$

When the wave function Ψ_j is known, we can find the associated eigenvalue E_j just by solving the integral on the left hand side with $i = j$. This recipe is quite straightforward, but, as we will see, solving this integral can be anything but straightforward!

1.1.2 The Variational Principle

A very important element in several many-body theories is the variational principle, which we will use frequently in the rest of this thesis. We will therefore give a detailed description of it and discuss the consequences, in this section.

We consider Ψ to be some appropriate wave function for a system and define an arbitrary but complete set of orthonormal wave functions Ψ_i which satisfy Eq. (1.2). Using this complete set, we can express Ψ as a linear combination, defined as

$$\Psi = \sum_i c_i \Psi_i \quad (1.6)$$

where c_i are unknown coefficients that define how the orthonormal set of Ψ_i combine to form Ψ . Although we don't yet know anything about the coefficients, we do know that the normality of Ψ imposes a constraint on the coefficients, which easily can be derived from

$$\begin{aligned}
1 &= \int |\Psi^* \Psi| d\mathbf{r} \\
&= \int \sum_i c_i^* \Psi_i^* \sum_j c_j \Psi_j d\mathbf{r} \\
&= \sum_{i,j} c_i^* c_j \int \Psi_i^* \Psi_j d\mathbf{r} \\
&= \sum_{i,j} c_i^* c_j \delta_{i,j} \\
&= \sum_i |c_i^* c_i|
\end{aligned} \tag{1.7}$$

What $|c_i^* c_i|$ tells us is the probability that a measurement of the energy would yield the value E_i . It is therefore clear that the sum of these probabilities should be 1 as shown in Eq. (1.7). The interpretation of coefficients becomes more clear if we insert Eq. (1.6) into Eq. (1.5), which gives us

$$\begin{aligned}
\int \Psi^* \mathcal{H} \Psi d\mathbf{r} &= \int \left(\sum_i c_i^* \Psi_i^* \right) \mathcal{H} \left(\sum_j c_j \Psi_j \right) d\mathbf{r} \\
&= \sum_{i,j} c_i^* c_j \int \Psi_i^* \mathcal{H} \Psi_j d\mathbf{r} \\
&= \sum_{i,j} c_i^* c_j E_j \delta_{i,j} \\
&= \sum_i |c_i^* c_i| E_i
\end{aligned} \tag{1.8}$$

Thus, the energy associated with the wave function Ψ is determinable from all of the coefficients c_i and eigenvalues E_i . However, note that, what we have found is not the result of an individual energy measurement, but a probabilistic expected value of the energy, also called the expectation value of \mathcal{H} , which is the average of repeated energy measurements on an ensemble of identically prepared systems.

Now we let E_0 be the lowest value in the set of energies and combine the results from Eqs. (1.7) and (1.8) to write

$$\int \Psi^* \mathcal{H} \Psi d\mathbf{r} - E_0 \int |\Psi^* \Psi| d\mathbf{r} = \sum_i |c_i^* c_i| (E_i - E_0). \tag{1.9}$$

Since $|c_i^* c_i|$ is always real and positive and $(E_i - E_0) \geq 0$ by definition, the right hand side will always be greater or equal to zero. This means that we have

$$\frac{\int \Psi^* \mathcal{H} \Psi d\mathbf{r}}{\int |\Psi^* \Psi| d\mathbf{r}} \geq E_0 \quad (1.10)$$

The last equation is the variational principle in mathematical notation and is much used in problems where we are interested in the ground state energy of a system. The variational principle gives us a way to judge the quality of an approximated wave function, which not necessarily needs to be represented by a linear combination. In fact, we can construct our guess wave function in any manner we wish, within the requirements the wave function has to satisfy. Any variations in the guess wave function which lower its energy are necessarily making the approximate energy closer to the exact answer and the guess wave function closer to the true ground state wave function.

1.1.3 The Born-Oppenheimer Approximation

The molecular Hamiltonian, given in equation Eq. (1.3), contains of pairwise attraction and repulsion terms, implying that no particle is moving independently of all of the others. Because of this interdependency, the stationary Schrödinger equation may be quite unpleasant to solve. However, this problem can be alleviated by the Born-Oppenheimer (BO) approximation, which luckily is entirely justified in most cases.

The BO approximation consist of separating the degrees of freedom connected with the motion of the nuclei from those of electrons. This approximation turns out to be a really good assumption because of the high ratio between nuclear and electronic masses. It is therefore intuitively clear that the nuclei move much more slowly than the electrons, and can be considered as fixed. This leads to a Hamiltonian for the electrons in the field generated by a static configuration of nuclei, and a separate Schrödinger equation for nuclei in which the electronic energy enters as a potential. We are mainly interested in the Hamiltonian for the electrons, which reads

$$\mathcal{H}_{el} = - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{4\pi\epsilon_0} \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \frac{1}{4\pi\epsilon_0} \sum_{n=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|} \quad (1.11)$$

The motion of the nuclei is neglected and only the electrostatic energy of the nuclei should be added to the energy of the electrons to arrive at the total energy. Thus, the (stationary) electronic Schrödinger equation reads

$$(\mathcal{H}_{el} + V_N) \Psi_{el}(\{\mathbf{r}_i\}; \{\mathbf{R}_n\}) = E_{el} \Psi_{el}(\{\mathbf{r}_i\}; \{\mathbf{R}_n\}) \quad (1.12)$$

where V_N is the nuclear–nuclear potential energy term in Eq. (1.3) and the subscript "el" emphasizes the use of BO approximation. Note that, the electronic wave function is a function of electronic degrees of freedom, but depends parametrically on the nuclear coordinates, which is the reason for using semicolon instead of comma in the variable list for Ψ_{el} . Moreover, V_N is constant for a given set of fixed nuclear coordinates, which means that we can solve Eq. (1.12) without V_N , since the wave function is invariant to the appearance of constant terms in the Hamiltonian.

The BO approximation is a very important result and the foundation of important concepts in quantum chemistry, such as the concept of potential energy surface (PES); a plot of the energy of a collection of nuclei and electrons against the geometric coordinates of the nuclei. A PES reflects the relationship between the energy of a molecule (or a collection of molecules) and its geometry, which is a very important relationship in many cases. However, without the BO approximation, this concept wouldn't be meaningful at all.

1.2 Construction of Trial Wave Functions

We have so far shown that by applying the Born-Oppenheimer approximation, we can concentrate on solving the (stationary) electronic Schrödinger equation given in Eq. (1.12). Solving this equation, however, is still a very comprehensive tasks and often not doable at all, at least without further approximations. The difficulty lies in the pairwise repulsion term in the Hamiltonian, which implies that no electron moves independently of all other electrons; the motion of electrons are correlated¹. Moreover, we also have a serious dimensionality problem. Solving Schrödinger equation for a system consisting of N_e electrons, involves solving a partial differential equation in $3N_e$ dimensions, which becomes quickly not feasible, as the number particles increases, using any of the standard methods for solving partial differential equations.

From the discussion above it seems quite hopeless to try to solve the electronic Schrödinger equation exactly. Instead of trying to solve this equation directly, let us try to approach the problem in a different way. What we

¹The term correlation is central in quantum chemistry, and is used to describe the interdependency between the electrons, usually beyond what is described in the Hartree-Fock method.

can do is to make a guess on the wave function of the system of interest and use the variational method to judge the quality of our guess wave function, also called the trial wave function. The one with lowest energy eigenvalue, is believed to be closest to the ground state energy and the best one to use for computing other properties by the application of other operators. The question now is how we go about constructing the trial wave function. This is what we are going to cover in the following sections. We start simple with a molecular system consisting of just one electron and move thereafter to many-electron systems.

We will hereafter use lowercase Greek letters for one-electron wave functions and uppercase Greek letters for many-electron wave functions. We also drop the subscript "el" in Eq. (1.12), and unless otherwise specified, all wave functions are electronic wave functions with parametrically dependency on nuclear coordinates. Further, we will refer to the eigenfunctions in Eq. (1.2) as *molecular orbitals* (MO) or *atomic orbitals* (AO) if we are dealing with atoms.

1.2.1 The LCAO Basis Set Approach

Consider a molecular system consisting of just one electron. The electronic wave function for such system will depend on the fixed nuclear coordinates and the three Cartesian coordinates of the single electron. If our system were to have only one nucleus as well, we have the Hydrogen atom, where Eq. (1.2) can be solved *exactly*. The resulting eigenfunctions in this case are the well-known hydrogenic atomic orbitals; $1s, 2s, 2p, 3s, 3p, 3d$, etc.. The properties and derivation of these eigenfunctions is discussed in detail in [1]. Now, our suggestion is that these orbitals may be useful, as functions, for constructing molecular orbitals, in more complex systems. In fact, we may represent a MO ϕ , as a linear combination of AOs;

$$\phi = \sum_{i=1}^M c_i \varphi_i \quad (1.13)$$

where each AO, φ_i , is multiplied by a corresponding coefficient c_i , reflecting the contribution to the MO. This representation, is known as the linear combination of atomic orbitals approach (LCAO).

It is very important to consider the basis set with AOs just as *functions* used to construct MOs. We anticipate as chemists, that AO are efficient functions for representing MOs, but this should not restrict our mathematical flexibility. For example, imagine a molecular system consisting of several hydrogen atoms. From our knowledge about the ground state of an undis-

turbed Hydrogen atom, we could naively limit ourselves to just use $1s$ function to represent MOs. But this chemical idea definitely restrict us, from a mathematical standpoint. Ultimately, we have a mathematical problem on our hand, we try to represent an arbitrary function by a combination of more convenient functions, and more AOs will bring the basis closer to spanning the true MO space. It should also be emphasized that sometimes it may even be more efficient to use other types of functions to represent MOs than the hydrogenic atomic orbitals².

The distinction between orbitals and functions is critical in computational MO theory and one should avoid conceptually mix them. This is discussed in more detail in [3].

The secular equation

We have seen that MOs can be constructed as linear combination of AOs, as shown in Eq. (1.13). But the coefficients, reflecting the contribution of each AO to the MO, are not yet known, so we need a recipe for finding them. Assuming real functions, we have from Eqs. (1.10) and (1.13);

$$\begin{aligned}
 E &= \frac{\int \left(\sum_i c_i \varphi_i \right) \mathcal{H} \left(\sum_j c_j \varphi_j \right) d\mathbf{r}}{\int \left(\sum_i c_i \varphi_i \right) \left(\sum_j c_j \varphi_j \right) d\mathbf{r}} \\
 &= \frac{\sum_{i,j} c_i c_j \int \varphi_i \mathcal{H} \varphi_j d\mathbf{r}}{\sum_{i,j} c_i c_j \int \varphi_i \varphi_j d\mathbf{r}} \\
 &= \frac{\sum_{i,j} c_i c_j H_{ij}}{\sum_{i,j} c_i c_j S_{ij}} \tag{1.14}
 \end{aligned}$$

where we have introduced the shorthand notation H_{ij} and S_{ij} for the integrals, which often are referred to as matrix elements, but also known as "resonance integral" and "overlap integral", respectively.

²As we will see in the next chapter, Gaussian functions turns out to be a good candidate for representing MOs.

We are interested in the coefficients that minimize the energy, so we require the derivative of the energy with respect to each of the coefficients to be zero;

$$\frac{\partial E}{\partial c_k} = 0 \quad \forall k. \quad (1.15)$$

This gives rise to M equations that can be solved to find the M coefficients $\{c_i\}$. After performing the differentiation, these equations read

$$\sum_i^M c_i (H_{ki} - ES_{ki}) = 0 \quad \forall k. \quad (1.16)$$

This is a linear algebra problem, which has a non-trivial solution if and only if $\det(H - ES) = 0$, i.e.

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} & \dots & H_{1M} - ES_{1M} \\ H_{21} - ES_{21} & H_{22} - ES_{22} & \dots & H_{2M} - ES_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ H_{M1} - ES_{M1} & H_{M2} - ES_{M2} & \dots & H_{MM} - ES_{MM} \end{vmatrix} = 0 \quad (1.17)$$

This equation is called the secular equation and has in general M roots, E_j , each giving rise to a different set of coefficients $\{c_{ij}\}$ (by solving Eq. (1.16)), which will represent the MO ϕ_j as

$$\phi_j = \sum_{i=1}^M c_{ij} \varphi_i \quad (1.18)$$

Here, the index j is used for the MO and index i for AOs used to construct the MO. The set of coefficients which gives the lowest energy eigenvalue, thus define the ground state MO of a one-electron system and higher energies define excited states.

1.3 Many-electron Wave Functions

In the previous section, we discussed how the wave function for a one-electron molecular system can be constructed. We will now, take the next step and show how the wave function for a many-electron system can be approximated. We will do this in several stages and increase the quality of the trial wave function gradually. As we will see soon, the wave function will be based on the one-electron MOs, so that we can benefit from the experience we have gained by studying the one-electron system.

1.3.1 Hartree Wave Function

The motion of one electron, in a many-electron system, depends on the motion of all the other electrons, which is due to the repulsion term in the Hamiltonian given in Eq. (1.11). We say that the electrons are correlated. Let us, however, for the moment totally ignore the interaction between electrons and simply drop the repulsion term. The Hamiltonian in this case is separable and can be expressed as

$$\mathcal{H}_0 = \sum_{i=1}^{N_e} \mathcal{h}_i \quad (1.19)$$

where N_e is the total number of electrons and \mathcal{h}_i is the one-electron Hamiltonian defined as (in atomic units, see appendix A.1)

$$\mathcal{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{n=1}^{N_n} \frac{Z_n}{|\mathbf{r}_i - \mathbf{R}_n|} \quad (1.20)$$

The nice thing about this Hamiltonian is that its eigenfunctions can be constructed as a product of eigenfunctions of the one-electron Hamiltonian. That is

$$\Phi_H = \phi_1 \phi_2 \dots \phi_N \quad (1.21)$$

where Φ_H is an eigenfunction of the simplified \mathcal{H}_0 given in Eq. (1.19), while $\{\phi_i\}$ are eigenfunctions of their respective operator $\{\mathcal{h}_i\}$. This can easily be shown by

$$\begin{aligned} \mathcal{H}_0 \Phi_H &= \mathcal{H}_0 \phi_1 \phi_2 \dots \phi_N \\ &= \sum_i^{N_e} \mathcal{h}_i \phi_1 \phi_2 \dots \phi_N \\ &= (\mathcal{h}_1 \phi_1) \phi_2 \dots \phi_N + \phi_1 (\mathcal{h}_2 \phi_2) \dots \phi_N + \dots + \phi_1 \phi_2 \dots (\mathcal{h}_N \phi_N) \\ &= (\epsilon_1 \phi_1) \phi_2 \dots \phi_N + \phi_1 (\epsilon_2 \phi_2) \dots \phi_N + \dots + \phi_1 \phi_2 \dots (\epsilon_N \phi_N) \\ &= \sum_i^{N_e} \epsilon_i \phi_1 \phi_2 \dots \phi_N \\ &= \left(\sum_i^{N_e} \epsilon_i \right) \Phi_{HP}, \end{aligned} \quad (1.22)$$

where we have used that

$$\hat{h}_i \phi_i = \epsilon_i \phi_i, \quad (1.23)$$

with ϵ_i as the energy eigenvalue of ϕ_i . Thus, the energy eigenvalue of Φ_H is simply given as the sum of the one-electron energy eigenvalues. The approximated wave function in Eq. (1.21) is known as the Hartree product, thereby the subscript "H", and is one of the simplest trial wave functions for a many-body system.

What we have done so far is simple and straightforward, but it is also a crime, at least from a chemical standpoint; we have totally ignored the interactions between the electrons! Let us now include the electron repulsion term in the molecular Hamiltonian and see if we can improve the quality of our Hartree product wave function. One alternative is to use the one-electron functions $\{\phi_i\}$, that minimize the expectation value of the Hamiltonian. This can in practice be done by applying variational calculus, which is described in details in section (1.4.1). It turns out that each of the optimal functions are eigenfunctions of their own operator \hat{h}_i , given as

$$\hat{h}_i = -\frac{1}{2}\nabla^2 - \sum_{n=1}^{N_n} \frac{Z_n}{|\mathbf{r} - \mathbf{R}_n|} + V_i^H\{j\} \quad (1.24)$$

where the last term represent the interaction with all other electrons, occupying orbitals $\{j\}$ and is given by

$$V_i^H\{j\} = \sum_{i \neq j} \int \frac{|\phi_j|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (1.25)$$

which is known as the Hartree potential. This term is very much like the second term in \hat{h}_i , but it involves an integration. This because electrons are treated as wave functions and their charge will therefore be spread out, so integration over all space is needed. The nuclei on the other hand are treated as point particles and their charge is therefore not spread out in space, ergo no integration needed.

Now, we want is to solve Eq. (1.23) with the improved Hamiltonians $\{\hat{h}_i\}$, in order to find optimal one-electron functions, which we later are going to use in the Hartree product. But, due to the interaction term, the one-electron Hamiltonians depend on the eigenfunctions themselves. How can we solve an equation when we need to know the solution of that equation, to actually formulate the equation? The solution is to apply the "self-consistent field" (SCF) procedure. What we do is basically to make a first guess on all of the one-electron wave functions $\{\phi_i\}$, which are used to construct the one-electron operators. We thereafter solve Eq. (1.23) to obtain a new set of

one-electron functions $\{\phi_i\}$. Intuitively, we expect this new set to be closer to the true optimal set, and hence it makes sense to repeat the calculations with the new set. We can keep doing this procedure, until the change in the calculated energy eigenvalues are less than some chosen threshold criterion.

Instead of totally ignore the repulsion term in the electronic Hamiltonian, we have added an effective interaction potential, which we hope describes some of the correlations between the electrons. This is indeed a better approximation, compared to just drop the repulsion term, but the Hartree-potential is still defective and corresponds to a "non-interacting" system. This doesn't mean that the electrons do not see each other—they indeed do—but the interaction is just included in an average way and hence their interaction is not accounted for instantaneously.

In addition to our simplified Hamiltonian, the wave function we have been using is also very simplified and usually referred to as the uncorrelated wave function or independent-particle wave function. From a statistical point of view, this approximation of the wave function is analogous to saying that the probability $P(A, B)$ for the event A and B is equal to probability of event A multiplied with the probability for event B . This would, of course, be true if events A and B were independent, but not otherwise.

1.3.2 Electron Spin and Antisymmetry

One very important thing we have totally ignored so far in this chapter, is the fact that electrons are identical fermions, which certainly has impact on the form of the wave function. This is because the wave function, has to obey the Pauli's antisymmetry principle, which states that the total wave function, including spin, for identical fermions must be antisymmetric with respect to exchange of the particles. Mathematically, this can be written as

$$\begin{aligned}
 \mathcal{P}_{ij}\Psi(\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N) \\
 &= \Psi(\mathbf{q}_1, \dots, \mathbf{q}_j, \dots, \mathbf{q}_i, \dots, \mathbf{q}_N) \\
 &= -\Psi(\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N)
 \end{aligned}
 \tag{1.26}$$

where \mathcal{P}_{ij} is the permutation operator, which interchanges the coordinates of particle i and j . Note that \mathbf{q} includes not only the three Cartesian coordinates, but also the spin, i.e. $\mathbf{q}_i = (\mathbf{r}_i, s_i)$, where s_i is the spin coordinate of particle i .

The spin coordinate is an additional degrees of freedom, but it differs from the spatial degrees of freedom, in the sense that it is not continuous. In

fact, electrons' spin coordinate only takes two values; $\frac{1}{2}$ or $-\frac{1}{2}$. So $s = \pm\frac{1}{2}$ are the only possible arguments a spin function can have. In order to span the spin space we only need two functions $\alpha(s)$ and $\beta(s)$, with functional form;

$$\begin{aligned}\alpha\left(\frac{1}{2}\right) &= 1 & \alpha\left(-\frac{1}{2}\right) &= 0 \\ \beta\left(\frac{1}{2}\right) &= 0 & \beta\left(-\frac{1}{2}\right) &= 1\end{aligned}\tag{1.27}$$

where $\alpha(s)$ denotes spin up and $\beta(s)$ denotes spin down. These functions are by definition orthonormal;

$$\begin{aligned}\int \alpha^*(s) \alpha(s) ds &= \int \beta^*(s) \beta(s) ds = 1 \\ \int \alpha^*(s) \beta(s) ds &= \int \beta^*(s) \alpha(s) ds = 0\end{aligned}\tag{1.28}$$

Note that, integration in spin space is in reality a summation over two discrete values of s , i.e.

$$\int \xi(s) ds \Rightarrow \sum_s \xi(s) = \xi\left(\frac{1}{2}\right) + \xi\left(-\frac{1}{2}\right)\tag{1.29}$$

Therefore, integration with respect to \mathbf{q} , denotes a summation over s and an integral over the spatial degrees of freedom;

$$\int d\mathbf{q} \Rightarrow \sum_s \int d\mathbf{r}\tag{1.30}$$

1.3.3 Slater Determinant

The quality of the trial wave function will certainly increase, if it is antisymmetric, because it then will give a more complete description of the system. But the question is how we can use what we already have to construct *an antisymmetric* trial wave function, i.e. an antisymmetric trial wave function in terms of one-electron MOs. This can be achieved by representing the wave function as a Slater determinant (SD);

$$\Psi_{SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{q}_1) & \psi_2(\mathbf{q}_1) & \cdots & \psi_N(\mathbf{q}_1) \\ \psi_1(\mathbf{q}_2) & \psi_2(\mathbf{q}_2) & \cdots & \psi_N(\mathbf{q}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{q}_N) & \psi_2(\mathbf{q}_N) & \cdots & \psi_N(\mathbf{q}_N) \end{vmatrix} \quad (1.31)$$

where N is the total number of electrons and $\psi_i(\mathbf{q}_j)$ is molecular spin-orbital i occupied by electron j , which is the product of a spatial orbital (MO) and an electron spin eigenfunction:

$$\psi_i(\mathbf{q}_j) = \phi_i(\mathbf{r}_j)\xi_i(s_j) \quad (1.32)$$

where ϕ_i is a pure spatial function, while ξ_i is a spin function (either spin up (α) or spin down (β)). In some cases it is more convenient to write the SD in terms of the antisymmetrizer operator (see appendix A.2), defined as

$$\mathcal{A} \equiv \frac{1}{N!} \sum_P (-1)^{P_n} \mathcal{P}, \quad (1.33)$$

where we sum over all permutations \mathcal{P} , which permutes the coordinates of the spin-orbitals only, and not their label. The multiplied sign, $(-1)^{P_n}$, depends on whether the permutation can be written as product of an even or odd number of pair interchanges. By using the antisymmetrizer we can express the SD simply as

$$\Psi_{SD} = \sqrt{N!} \mathcal{A} \Psi_H \quad (1.34)$$

where $\Psi_H = \psi_1(\mathbf{q}_1)\psi_2(\mathbf{q}_2)\dots\psi_N(\mathbf{q}_N)$ is the well-known Hartree product wave function.

A more compact notation for the SD is

$$|\Psi_{SD}\rangle = |\psi_1\psi_2\psi_3\dots\psi_N\rangle \quad (1.35)$$

where the prefactor $(N!)^{-1/2}$ is implicit. Moreover, if two spin-orbitals, say ψ_1 and ψ_2 , have the same orbital function but different spin function, we can write

$$|\Psi_{SD}\rangle = |\phi_1^2\psi_3\dots\psi_N\rangle \quad (1.36)$$

where ϕ_1 represent a pure spatial function.

A nice property of determinants is that they change sign when any two rows (or columns) are interchanged, which is equivalent to interchanging the coordinates of two electrons. As a result, exchange of any two particle's coordinates changes the sign of the wave function of the system, which is

what we want. Moreover, if two spin-orbitals are equal, say $\psi_1 = \psi_2$, then the Slater determinant will vanish, which is a reflection of Pauli's exclusion principle; two electrons cannot occupy the same quantum mechanical state.

The most interesting consequence of the functional form of the Slater determinate, however, is that the electrons are indistinguishable, which is a critical quantum mechanical feature. This is reflected by the fact that each electron is associated with every spin-orbital! This is in contrast to the Hartree product where the indistinguishability of electrons is not respected. In the Hartree product wave function each electron is associated with only their own spin-orbital; electron i occupies spin-orbital i , electron $i+1$ occupies spin-orbital $i+1$ and so on.

Exchange hole

By using SD as wave function, we certainly improve the quality of our trial wave function, compared to the Hartree product and the electrons are no longer uncorrelated. To see how the correlation arises, we consider the probability density for finding one electron with coordinate \mathbf{q}_1 and another with \mathbf{q}_2 simultaneously. To make the algebra easier, we consider a two-electron SD given as

$$\Psi_{SD}(\mathbf{q}_1, \mathbf{q}_2) = |\psi_1\psi_2\rangle \quad (1.37)$$

The desired probability density with this wave function, is found by expanding the SD;

$$\begin{aligned} \rho(\mathbf{q}_1, \mathbf{q}_2) &= |\Psi_{SD}|^2 \\ &= \frac{1}{2} |\psi_1(\mathbf{q}_1)\psi_2(\mathbf{q}_2) - \psi_1(\mathbf{q}_2)\psi_2(\mathbf{q}_1)|^2 \\ &= \frac{1}{2} \left[|\psi_1(\mathbf{q}_1)|^2 |\psi_2(\mathbf{q}_2)|^2 + |\psi_1(\mathbf{q}_2)|^2 |\psi_2(\mathbf{q}_1)|^2 \right. \\ &\quad \left. - \psi_1^*(\mathbf{q}_1)\psi_2(\mathbf{q}_1)\psi_2^*(\mathbf{q}_2)\psi_1(\mathbf{q}_2) - \psi_1(\mathbf{q}_1)\psi_2^*(\mathbf{q}_1)\psi_2(\mathbf{q}_2)\psi_1^*(\mathbf{q}_2) \right] \end{aligned} \quad (1.38)$$

In the uncorrelated case, the same probability density is given as

$$\rho(\mathbf{q}_1, \mathbf{q}_2) = |\psi_1(\mathbf{q}_1)|^2 |\psi_2(\mathbf{q}_2)|^2. \quad (1.39)$$

We see already at this point the inclusion of correlation in a SD, compared to Hartree product, where the probability density is simplified to be the product

of the one-electron probability densities. However, if we are interested in knowing the probability density for finding one electron with coordinate \mathbf{r}_1 and another with \mathbf{r}_2 simultaneously, we will obtain (by integrating out the spin degrees of freedom):

$$\begin{aligned} \rho(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2} & \left[|\phi_1(\mathbf{r}_1)|^2 |\phi_2(\mathbf{r}_2)|^2 + |\phi_1(\mathbf{r}_2)|^2 |\phi_2(\mathbf{r}_1)|^2 \right. \\ & - \phi_1^*(\mathbf{r}_1) \phi_1(\mathbf{r}_2) \phi_2^*(\mathbf{r}_2) \phi_2(\mathbf{r}_1) \delta_{\xi_1 \xi_2} \delta_{\xi_1 \xi_2} \\ & \left. - \phi_1(\mathbf{r}_1) \phi_2^*(\mathbf{r}_1) \phi_2(\mathbf{r}_2) \phi_1^*(\mathbf{r}_2) \delta_{\xi_1 \xi_2} \delta_{\xi_1 \xi_2} \right] \end{aligned} \quad (1.40)$$

Now, if ψ_1 and ψ_2 have opposite spin, the two last terms on the right hand side vanishes and therefore opposite spin-orbitals are still uncorrelated. This becomes particularly apparent if we consider the case where $\phi_1 = \phi_2$ ³:

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = |\phi_1(\mathbf{r}_1)|^2 |\phi_1(\mathbf{r}_2)|^2, \quad (1.41)$$

which is exactly the result from the uncorrelated case⁴.

But if ψ_1 and ψ_2 have same spin ($\delta_{\xi_1 \xi_2} = 1$), the two last terms in Eq. (1.40) will not disappear, so that when $\mathbf{r}_1 = \mathbf{r}_2$, the terms on the right hand side will cancel and the probability density will be zero! This indicates that electron pairs with parallel spin are kept apart, which is a correlation effect. One can imagine that every electron is surrounded by an "exchange hole", where other electrons with the same spin are hardly found. Electrons with opposite spin, on the other hand, are not affected by the exchange hole. It is important to realize that the occurrence of exchange holes, has nothing to do with the electron-electron repulsion, but is rather a consequence of the antisymmetric nature of the wave function. This phenomena will, however, as we will see shortly, have a direct effect on the energy of the system.

Note that we have referred to exchange hole as an correlation effect, but usually the term "correlation effects" is reserved for all correlations apart from exchange, and it is in this sense we are going to use this term from now on. Also, note that, since the motion of electrons with opposite spin remains uncorrelated, it is customary to refer a single determinant wave function as an uncorrelated wave function. In fact, every trial wave function, which at most includes exchange effects, is referred to as an uncorrelated wave function.

³The electrons are allowed to have the same spatial orbital in this case, since they have opposite spin.

⁴Note that, when $\phi_1 \neq \phi_2$, the probability density is given as an average, in contract to Eq. (1.39). This is a consequence of the indistinguishability of the electrons, which is not respected in the Hartree product.

1.4 The Hartree-Fock Theory

Earlier in this chapter, we saw that optimal one-electron functions in a Hartree product wave function can be found as eigenfunctions of a set of one-electron operators. These operators (Eq. (1.24)) occurred as a result of applying variational calculus for minimizing the expectation value of the Hamiltonian, given in Eq. (1.11). We will in this section follow the same strategy, but use one single SD as wave function, instead of Hartree product. In other words, we wish to find optimal spin-orbitals in the SD, by invoking the variational method. This is the basic philosophy behind the Hartree-Fock (HF) method, which is a natural extension of the Hartree's SCF procedure.

In the following sections we will describe the HF method in detail and derive the so-called HF equation. We will first give a formal derivation of this equation, and thereafter discuss the physical meaning of it in its general form. The derivations in this section is based on the chapter 4 in [4], written by Jos Thijssen.

1.4.1 Derivation of the Hartree-Fock Equation

In this section we will derive the Hartree-Fock equation in its general spin-orbital form, by minimizing the energy expression for a single SD, and without making any assumptions on the spin-orbitals. Later in this chapter, we will discuss restricted and unrestricted spin-orbitals and derive the corresponding equations for each case.

The derivation of Eq. (1.24) is analogous, but in that case we use the Hartree product as wave function instead of SD.

Expectation value of the energy

We start by rewriting the Hamiltonian in Eq. (1.11) as

$$\mathcal{H} = \sum_{i=1}^{N_e} \mathcal{h}_i + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} \mathcal{g}_{ij} \quad \text{with} \quad (1.42)$$

$$\mathcal{h}_i = -\frac{1}{2} \nabla_i^2 - \sum_{n=1}^{N_n} \frac{Z_n}{|\mathbf{r}_i - \mathbf{R}_n|} \quad (1.43)$$

$$\mathcal{g}_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (1.44)$$

where N_e and N_n are the number of electrons and nuclei (with charge Z_n), respectively. We see immediately that \mathcal{h}_i is an one-particle operator since it

depends only on \mathbf{r}_i , while g_{ij} is a two-particle operator and depends on both \mathbf{r}_i and \mathbf{r}_j . In order to find the expectation value of the energy, we need to know the expectation value of both h_i and g_{ij} .

The expectation value of the one-particle operator h_i is found by integrating over both spatial and spin degrees of freedom:

$$\begin{aligned} \langle \Psi_{SD} | \sum_{i=1}^{N_e} h_i | \Psi_{SD} \rangle &= N! \int [\mathcal{A}\Psi_H]^\dagger \left(\sum_{i=1}^{N_e} h_i \right) [\mathcal{A}\Psi_H] d\mathbf{q} \\ &= N! \sum_{i=1}^{N_e} \int \Psi_H^* (h_i \mathcal{A}) \Psi_H d\mathbf{q} \end{aligned} \quad (1.45)$$

where we have used that $\mathcal{A}^\dagger \mathcal{A} = \mathcal{A}^2 = \mathcal{A}$ and that \mathcal{A} commutes with \mathcal{H} and thereby with all h_i as well (see section A.2). Now, any permutation of electron coordinates will make the integral zero because of the orthonormality of the spin-orbitals, and therefore the only nonzero contribution is when there is no permutation of coordinates;

$$\begin{aligned} \langle \Psi_{SD} | \sum_{i=1}^{N_e} h_i | \Psi_{SD} \rangle &= \int \Psi_H^* \left(\sum_{i=1}^{N_e} h_i \right) \Psi_H d\mathbf{q} \\ &= \left(\int \psi_1^* h_1 \psi_1 d\mathbf{q}_1 \right) + \left(\int \psi_2^* h_2 \psi_1 d\mathbf{q}_2 \right) + \cdots + \left(\int \psi_N^* h_N \psi_N d\mathbf{q}_N \right) \\ &= \sum_{k=1}^{N_{so}} \langle \psi_k | h | \psi_k \rangle. \end{aligned} \quad (1.46)$$

Note that we have dropped the subscript on h , since this term has the same form for all i (i.e. for all electrons), which is not surprising since electrons are indistinguishable. We have also changed the summation index from i to k , to clearly indicate that we are summing over spin-orbitals and not electrons, although the number of each is the same ($N_e = N_{so}$).

The expectation value of the two-particle integral g_{ij} is found in the same way;

$$\begin{aligned}
\langle \Psi_{SD} | \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} g_{ij} | \Psi_{SD} \rangle &= N! \int [\mathcal{A}\Psi_H]^\dagger \left(\sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} g_{ij} \right) [\mathcal{A}\Psi_H] d\mathbf{q} \\
&= N! \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} \int \Psi_H^* (g_{ij} \mathcal{A}) \Psi_H d\mathbf{q} \quad (1.47)
\end{aligned}$$

In this case we will get nonzero contribution when i and j are interchanged on one side, in addition to the contribution from zero permutation;

$$\begin{aligned}
\langle \Psi_{SD} | \sum_{\substack{i,j=1 \\ i \neq j}}^{N_e} g_{ij} | \Psi_{SD} \rangle &= \left(\int \psi_1^*(\mathbf{q}_1) \psi_2^*(\mathbf{q}_2) g_{12} \psi_1(\mathbf{q}_1) \psi_2(\mathbf{q}_2) d\mathbf{q}_1 d\mathbf{q}_2 \right) + \dots \\
&+ \left(\int \psi_{N-1}^*(\mathbf{q}_{N-1}) \psi_N^*(\mathbf{q}_N) g_{N-1,N} \psi_{N-1}(\mathbf{q}_{N-1}) \psi_N(\mathbf{q}_N) d\mathbf{q}_{N-1} d\mathbf{q}_N \right) \\
&- \left(\int \psi_1^*(\mathbf{q}_1) \psi_2^*(\mathbf{q}_2) g_{12} \psi_2(\mathbf{q}_1) \psi_1(\mathbf{q}_2) d\mathbf{q}_1 d\mathbf{q}_2 \right) - \dots \\
&- \left(\int \psi_{N-1}^*(\mathbf{q}_{N-1}) \psi_N^*(\mathbf{q}_N) g_{N-1,N} \psi_N(\mathbf{q}_{N-1}) \psi_{N-1}(\mathbf{q}_N) d\mathbf{q}_{N-1} d\mathbf{q}_N \right) \\
&= \sum_{k,l=1}^{N_{so}} \langle \psi_k \psi_l | g | \psi_k \psi_l \rangle - \langle \psi_k \psi_l | g | \psi_l \psi_k \rangle. \quad (1.48)
\end{aligned}$$

where we have used the following notation in the last step:

$$\langle \psi_k \psi_l | g | \psi_m \psi_n \rangle = \int \psi_k^*(\mathbf{q}) \psi_l^*(\mathbf{q}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_m(\mathbf{q}) \psi_n(\mathbf{q}') d\mathbf{q} d\mathbf{q}'. \quad (1.49)$$

The negative sign in front of the second term in Eq. (1.48) arises because all permutations which yield a single interchange of electron coordinates will be generated by an odd power of the permutation operator (P_n in Eq. (1.33) is odd).

Note that, we have dropped the indices on g and changed the summation indices, of the same reason as before; electrons are indistinguishable, so every electron pair yield the same result. Moreover, the restriction in summation is ignored, since

$$\langle \psi_k \psi_l | g | \psi_k \psi_l \rangle - \langle \psi_k \psi_l | g | \psi_l \psi_k \rangle = 0,$$

when $k = l$. We are therefore allowed to sum freely.

By combining Eqs. (1.46) and (1.48), the expectation value of the energy can be expressed as

$$\begin{aligned} E &= \langle \Psi_{SD} | \mathcal{H} | \Psi_{SD} \rangle \\ &= \sum_{k=1}^{N_{so}} \langle \psi_k | \hat{h} | \psi_k \rangle + \frac{1}{2} \sum_{k,l=1}^{N_{so}} \langle \psi_k \psi_l | \mathcal{J} | \psi_k \psi_l \rangle - \langle \psi_k \psi_l | \mathcal{J} | \psi_l \psi_k \rangle, \end{aligned} \quad (1.50)$$

where the electron coordinate indices on \hat{h} and \mathcal{J} operator are not needed, due to the indistinguishability of electrons. We can express this expression more compact if we introduce the following operators;

$$J_l \psi(\mathbf{q}) = \int \frac{\psi_l^*(\mathbf{q}') \psi_l(\mathbf{q}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{q}' \psi(\mathbf{q}) \quad (1.51a)$$

$$K_l \psi(\mathbf{q}) = \int \frac{\psi_l^*(\mathbf{q}') \psi(\mathbf{q}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{q}' \psi_l(\mathbf{q}) \quad (1.51b)$$

and furthermore

$$\mathcal{J} = \sum_l J_l, \quad \mathcal{K} = \sum_l K_l, \quad (1.52)$$

which are known as the Coulomb and exchange operator, respectively. Using these operators we can rewrite Eq. (1.50) as

$$E = \sum_{k=1}^{N_{so}} \langle \psi_k | \hat{h} + \frac{1}{2} (\mathcal{J} - \mathcal{K}) | \psi_k \rangle. \quad (1.53)$$

If we were to use Hartree product as trial wave function instead of SD, the expectation value of the energy is almost the same as the expression above, but without the exchange term and with constraint $k \neq l$ in summation over l in \mathcal{J} . This is because the Coulomb part doesn't cancel without the exchange term, in the case $k = l$.

Variational method

In order to find optimal spin-orbitals, we can apply the standard techniques of calculus of variations. This is, seek for an optimal set of single-particle states that makes the energy functional in Eq. (1.53) stationary under infinitesimal

changes; $|\psi_k\rangle \rightarrow |\psi_k\rangle + |\delta\psi_k\rangle$. The variation in spin-orbitals is, however, not completely arbitrary and has to satisfy the orthonormality requirement;

$$\langle\psi_k|\psi_l\rangle = \int \psi_k^* \psi_l \, d\mathbf{r} = \delta_{kl} \quad (1.54)$$

By introducing the Lagrange multipliers Λ_{kl} ([6]), we can set up the Lagrange function, which we wish to minimize;

$$\mathcal{L} = E - \sum_{kl} \Lambda_{kl} (\langle\psi_k|\psi_l\rangle - \delta_{kl}) \quad (1.55)$$

Since the Lagrangian has to be real and $\langle\psi_k|\psi_l\rangle = \langle\psi_l|\psi_k\rangle^*$ we must have $\Lambda_{kl} = \Lambda_{lk}^*$, i.e. Lagrange multipliers are elements of an hermitian matrix. By applying an arbitrary change $|\psi_k\rangle \rightarrow |\psi_k\rangle + |\delta\psi_k\rangle$ and require $\delta\mathcal{L} = 0$, we obtain

$$\begin{aligned} \delta E - \sum_{kl} \Lambda_{kl} (\langle\delta\psi_k|\psi_l\rangle + \langle\psi_k|\delta\psi_l\rangle) \\ = \delta E - \sum_{kl} (\Lambda_{kl} \langle\delta\psi_k|\psi_l\rangle + \Lambda_{lk}^* \langle\psi_l|\delta\psi_k\rangle) \\ = \delta E - \sum_{kl} (\Lambda_{kl} \langle\delta\psi_k|\psi_l\rangle + \Lambda_{kl}^* \langle\delta\psi_k|\psi_l\rangle^*) \\ = \delta E - \sum_{kl} \Lambda_{kl} \langle\delta\psi_k|\psi_l\rangle + \text{complex conj.} = 0 \end{aligned} \quad (1.56)$$

with

$$\begin{aligned} \delta E &= \sum_k \langle\delta\psi_k| \hat{h} |\psi_k\rangle + \text{complex conj.} \\ &+ \frac{1}{2} \sum_{k,l} \left(\langle\delta\psi_k \psi_l| \mathcal{G} |\psi_k \psi_l\rangle + \langle\psi_k \delta\psi_l| \mathcal{G} |\psi_k \psi_l\rangle \right. \\ &\quad \left. - \langle\delta\psi_k \psi_l| \mathcal{G} |\psi_l \psi_k\rangle - \langle\psi_k \delta\psi_l| \mathcal{G} |\psi_l \psi_k\rangle \right) + \text{complex conj.} \\ &= \sum_k \langle\delta\psi_k| \hat{h} |\psi_k\rangle + \text{complex conj.} \\ &+ \sum_{k,l} \left(\langle\delta\psi_k \psi_l| \mathcal{G} |\psi_k \psi_l\rangle - \langle\delta\psi_k \psi_l| \mathcal{G} |\psi_l \psi_k\rangle \right) + \text{complex conj.}, \end{aligned} \quad (1.57)$$

where we have used

$$\langle \psi_k | \hat{h} | \delta \psi_k \rangle = \langle \delta \psi_k | \hat{h} | \psi_k \rangle^* \quad (1.58a)$$

$$\langle \psi_k \psi_l | \mathcal{G} | \delta \psi_k \psi_l \rangle = \langle \delta \psi_k \psi_l | \mathcal{G} | \psi_k \psi_l \rangle^* \quad (1.58b)$$

etc., which follow from the definition of the integrals. In the last step, the dummy indices k and l are interchanged in the second term in second and third line, and the symmetry property of the two-electron element used;

$$\langle \psi_k \psi_l | \mathcal{G} | \psi_m \psi_n \rangle = \langle \psi_l \psi_k | \mathcal{G} | \psi_n \psi_m \rangle. \quad (1.59)$$

We can rewrite Eq. (1.56) as

$$\sum_k \langle \delta \psi_k | \mathcal{F} | \psi_k \rangle + \text{complex conj.} = \sum_{kl} \Lambda_{kl} \langle \delta \psi_k | \psi_l \rangle + \text{complex conj.} \quad (1.60)$$

with

$$\mathcal{F} = \hat{h} + \mathcal{J} - \mathcal{K}, \quad (1.61)$$

known as the Fock operator. Since the variations $\delta \psi^*$ and $\delta \psi$ are arbitrary and independent, we can formulate the minimization condition as⁵

$$\mathcal{F} \psi_k = \sum_{kl} \Lambda_{kl} \psi_k. \quad (1.62)$$

Now, the Lagrange multipliers Λ_{kl} must be chosen so that the set of solutions $\{\psi_k\}$ form an orthonormal set. The case where $\Lambda_{kl} = \delta_{kl} \epsilon_k$, leads to

$$\mathcal{F} \psi_k = \epsilon_k \psi_k, \quad (1.63)$$

which is an eigenvalue equation, with ψ_k as an eigenfunction of \mathcal{F} . Therefore will solutions of this equation form an orthonormal set and thereby is the constraint equation (1.54) automatically satisfied.

Eq. (1.63) is known as the Hartree-Fock equation, and is used to find optimal spin-orbitals that minimize the energy expression for a single SD. This equation is nonlinear and must be solved by a self-consistency (SCF) iterative procedure. The philosophy is the same as before; make a first guess on all of the one-electron wave functions $\{\psi_k\}$, to construct the Fock operator and thereafter solve Eq. (1.63) to obtain a new set of one-electron functions

⁵To see this, replace $|\delta \psi\rangle$ by $i |\delta \psi\rangle$ and derive similar equation to the one in Eq. (1.60) and combine the two equations.

$\{\psi_k\}$, which we expect to be closer to the true optimal set. We solve the same equation with the new set and repeat this procedure until self-consistency is reached, i.e. until spin-orbitals used to construct \mathcal{F} are the same as its eigenfunctions.

1.4.2 Hartree-Fock Equation - Physical Picture

By taking antisymmetry into account, the one-electron Hamiltonian in Eq. (1.24) can be extended to

$$\mathcal{F} = -\frac{1}{2}\nabla^2 - \sum_{n=1}^{N_n} \frac{Z_n}{|\mathbf{r} - \mathbf{R}_n|} + V^{HF}, \quad (1.64)$$

where the last term is the interaction potential operator, and can be expressed in terms of its action on an arbitrary state $\psi_k(\mathbf{q})$;

$$V^{HF}\psi_k(\mathbf{q}) = \sum_l \left[\int \frac{\psi_l^*(\mathbf{q}')\psi_l(\mathbf{q}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{q} \psi_k(\mathbf{q}) - \int \frac{\psi_l^*(\mathbf{q}')\psi_k(\mathbf{q}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{q}' \psi_l(\mathbf{q}) \right] \quad (1.65)$$

These two terms are known as the direct term (or Coulomb term) and exchange term, respectively. If we exclude $l = k$ in summation, the direct term is nothing else than the Hartree potential given in Eq. (1.25), and describes the total averaged potential acting on an electron in spin-orbital ψ_k , arising from other electrons in other spin-orbitals. However unlike Eq. (1.25), the direct term contains coupling between orbital k and itself, since we don't have any restriction in summation over l . This is of course unphysical, since an electron does not interact with itself. But, fortunately, this term is canceled by the exchange term, so we can nevertheless sum freely over l .

The exchange term looks much like the direct term, except that it is nonlocal, which means that when acting on ψ_k its value at \mathbf{q} is determined by the value assumed by ψ_k at all possible positions \mathbf{q}' . The occurrence of this term is a direct consequence of the antisymmetric form of the wave function, which lowers the Coulomb interaction between the electrons with same spin. This is because these electrons are kept apart (due to exchange hole) and therefore their interaction is reduced. As mentioned earlier, it is important to realize that this behavior has nothing to do with the electron-electron repulsion, and is a direct consequence of the antisymmetric nature of the wave function. But it does keep the electrons with parallel spin apart

and therefore reduce their Coulombic interaction, which certainly effect the energy of the system.

The HF equation (Eq. (1.63)) has the form of an ordinary Schrödinger equation, although $\{\epsilon_k\}$ are primarily identified as Lagrange multipliers, and not energies. But they are, however, related to the total energy by

$$E = \sum_{k=1}^{N_{so}} \langle \psi_k | \hat{h} + \frac{1}{2}(\mathcal{J} - \mathcal{K}) | \psi_k \rangle = \sum_{k=1}^{N_{so}} \epsilon_k - \frac{1}{2} \langle \psi_k | \mathcal{J} - \mathcal{K} | \psi_k \rangle. \quad (1.66)$$

It is thus apparent that the total energy is not simply the sum of all ϵ_k , which is due to the double counting of electron-electron interaction between pairs of electrons in $\sum_k \epsilon_k$. We must therefore compensate for this by subtracting $\frac{1}{2} \langle \psi_k | \mathcal{J} - \mathcal{K} | \psi_k \rangle$. In literature $\{\epsilon_k\}$ are often referred to as orbital energies and attached physical significance through the Koopman's theorem[4].

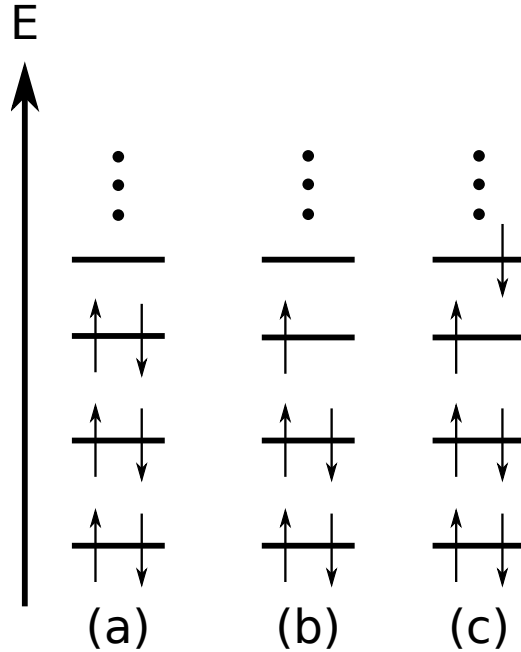


Figure 1.1: HF spectrum: schematic representation on how the levels are filled for (a): the ground state of an even number of electrons, (b): the ground state of an odd number of electrons and (c): an excited spectrum of an even number of electrons.

Solving HF equation yields a set $\{\psi_k\}$ of orthonormal spin-orbitals with orbital energies $\{\epsilon_k\}$. In principle, the set consist of infinite number of solutions, leading to an infinite HF spectrum. For a system consisting of N_e

electrons, the ground state is approximated by taking the N_e lowest eigenstates of this spectrum as spin-orbitals of the electrons (see Fig. 1.1). It is, however, not a clear a priori that the ground state is found by filling the lowest eigenstates, since the energy is not simply the sum of spin-orbital energies. But in practical applications this turns out to be the case anyway[4].

1.5 Closed- and Open-shell Systems

We have so far discussed the HF equation in a formal way, without specifying the explicit form of the spin-orbitals. In order to do actual calculations using the HF method, we must be more specific about the form of the spin-orbitals. Depending on the choice of spin-orbitals we can formulate two different versions of HF; restricted HF (RHF) and unrestricted HF (UHF). In RHF, the spin-orbitals have the same spatial part for different spin function, and is usually used to describe closed-shell systems, where all levels are doubly occupied. To describe open-shell systems, where there are partially filled levels containing only one electron, it is more common to use the UHF formalism⁶. In this case the spin-orbitals have different spatial function for different spin functions. In figure 1.2 both restricted and unrestricted configuration is shown.

Note that, a system with even number of electrons is not necessarily a closed-shell system, because of the possibility of existence of degenerate levels. In addition, we may also consider an excited state where an electron is excited to a higher level. In this case two levels will be partially filled and we are therefore dealing with an open-shell system, although the number of electrons is even. A system with odd number of electrons will always be an open-shell system.

We will in the following sections describe both RHF and UHF, where we restrict the discussion about RHF to just closed-shell systems, although the formalism also can be applied to open-shell systems. To describe open-shell systems the UHF formalism is usually used.

1.5.1 Restricted Hartree-Fock

In RHF, the spin-orbitals be grouped in pairs with the same spatial wave function, but opposite spin;

⁶The unrestricted formulation is sometimes also applied on systems that are normally thought as closed-shell systems, since it gives a better description than the restricted formalism does. One example is the dissociation problem discussed in [5].

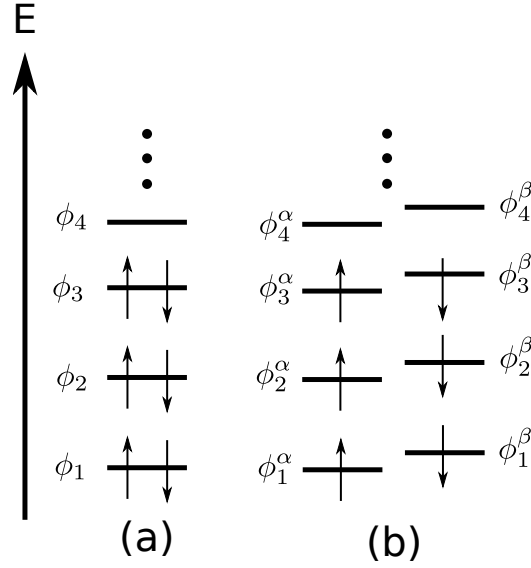


Figure 1.2: Restricted and unrestricted spectrum: schematic representation of (a): spin restricted configuration and (b): spin unrestricted configuration.

$$\begin{aligned}\psi_{2k}(\mathbf{q}) &= \phi_k(\mathbf{r})\alpha(s) \\ \psi_{2k-1}(\mathbf{q}) &= \phi_k(\mathbf{r})\beta(s), \quad k = 1, \dots, N/2\end{aligned}\quad (1.67)$$

where N is the total number of spin-orbitals. Using these spin orbitals we can express the ground state as

$$|\Psi_{RHF}\rangle = |\psi_1\psi_2\psi_3\dots\psi_N\rangle = |\phi_1^2\phi_2^2\dots\phi_{N/2}^2\rangle, \quad (1.68)$$

where the subscript "2" indicates doubly occupied spatial orbitals. We now want to replace the general spin-orbital HF equation in Eq. (1.63), with a pure spatial eigenvalue equation, by integrating out the spin degrees of freedom. Our starting point is naturally, the general spin-orbital HF equation:

$$\mathcal{F}\psi_k(\mathbf{q}) = \epsilon_k\psi_k(\mathbf{q}), \quad (1.69)$$

where ψ_k can either have α or β spin function. We will assume that ψ_k has α spin function, but identical results will be obtained with β spin function;

$$\mathcal{F}\phi_k(\mathbf{r})\alpha(s) = \epsilon_k\phi_k(\mathbf{r})\alpha(s). \quad (1.70)$$

By multiplying $\alpha^*(s)$ from left and integrating over spin, we obtain

$$\hat{h}\phi_k(\mathbf{r}) + \sum_{l=1}^N \int \alpha^*(s) (J_l - K_l) \phi_k(\mathbf{r}) \alpha(s) ds = \epsilon_k \phi_k(\mathbf{r}), \quad (1.71)$$

where \hat{h} remains the same since it has no spin-dependency. Now, in a closed-shell system the sum over spin-orbitals includes a sum over those with α spin function and a sum over those with β spin function;

$$\sum_{l=1}^N \rightarrow \sum_{l_\alpha=1}^{N/2} + \sum_{l_\beta=1}^{N/2}$$

so we can rewrite the integral in last equation as

$$\begin{aligned} \sum_{l=1}^N \int \alpha^*(s) (J_l - K_l) \phi_k(\mathbf{r}) \alpha(s) ds &= \sum_{l_\alpha=1}^{N/2} \int \alpha^*(s) (J_{l_\alpha} - K_{l_\alpha}) \phi_k(\mathbf{r}) \alpha(s) ds \\ &+ \sum_{l_\beta=1}^{N/2} \int \alpha^*(s) (J_{l_\beta} - K_{l_\beta}) \phi_k(\mathbf{r}) \alpha(s) ds \\ &= \sum_{l=1}^{N/2} 2J_l(\mathbf{r}) - K_l(\mathbf{r}), \end{aligned} \quad (1.72)$$

where $J_l(\mathbf{r})$ and $K_l(\mathbf{r})$ are analogous to definitions in Eq. (1.51), but in terms of spatial orbitals only;

$$J_l\phi(\mathbf{r}) = \int \frac{\phi_l^*(\mathbf{r}')\phi_l(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \phi(\mathbf{r}) \quad (1.73a)$$

$$K_l\phi(\mathbf{r}) = \int \frac{\phi_l^*(\mathbf{r}')\phi(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \phi_l(\mathbf{r}) \quad (1.73b)$$

Note that, the factor 2 arises because the Coulomb terms are equal for α and β spin, while the exchange term for β spin vanishes due to spin orthogonality.

The closed-shell Fock operator can thus be written as

$$\mathcal{F}_{spatial} = \hat{h} + 2\mathcal{J}_{spatial} - \mathcal{K}_{spatial} \quad (1.74)$$

where

$$\mathcal{J}_{spatial} = \sum_{l=1}^{N/2} J_l(\mathbf{r}), \quad \mathcal{K}_{spatial} = \sum_{l=1}^{N/2} K_l(\mathbf{r}), \quad (1.75)$$

The corresponding energy functional is given by

$$E = \sum_{k=1}^{N/2} \langle \phi_k | 2\hat{h} + 2\mathcal{J}_{spatial} - \mathcal{K}_{spatial} | \phi_k \rangle. \quad (1.76)$$

Note that, the energy is a functional of spatial orbitals $\{\phi_k\}$, and not spin-orbitals $\{\psi_k\}$, as it was in Eq. (1.53). From now on, we will drop the subscript "spatial" on operators in Eqs. (1.74) and (1.76), and instead specify the spatial form of the Fock operator by $\mathcal{F}(\mathbf{r})$. The original form of the Fock operator including spin will be specified as $\mathcal{F}(\mathbf{q})$. Operators \mathcal{J} and \mathcal{K} , will be specified in the same way.

1.5.2 LCAO Approach: The Roothaan Equation

By eliminating the spin, we are left with an spatial eigenvalue problem;

$$\mathcal{F}(\mathbf{r}) \phi_k(\mathbf{r}) = \epsilon_k \phi_k(\mathbf{r}), \quad k = 1, \dots, N/2, \quad (1.77)$$

where N is the total number of spin-orbitals. This equation is in practice solved by introducing a set of known basis functions $\{\varphi_p\}$, which expand the unknown spatial wave functions (MO);

$$\phi_k(\mathbf{r}) = \sum_{p=1}^M C_{pk} \varphi_p(\mathbf{r}), \quad (1.78)$$

where the expansion coefficients C_{pk} are not yet known. This is analogous to what we did earlier in this chapter when we discussed the LCAO basis set approach (see section 1.2.1). Note that although the basis functions are assumed to be normalized and linearly independent, they are not in general orthogonal to each other. Orthogonality is required among the MOs, but not among the basis functions representing the MOs. If the set $\{\varphi_p\}$ was complete, we could represent the MO exact, but in practice we are limited to use a finite set of M basis functions, because of computational reasons. Since we are limited to finite basis sets, it is important to choose a basis that describe the MO efficient. In the next chapter we will discuss the questions involved in the choice of a basis set in detail, but for now we will assume that the set $\{\varphi_p\}$ is some known basis set.

By inserting the expansion in Eq. (1.78), into the HF equation, we obtain

$$\mathcal{F}(\mathbf{r}) \sum_{q=1}^M C_{qk} \varphi_q(\mathbf{r}) = \epsilon_k \sum_{q=1}^M C_{qk} \varphi_q(\mathbf{r}), \quad (1.79)$$

which can be converted to a matrix equation by multiplying $\varphi_p^*(\mathbf{r})$ from left and integrating;

$$\sum_{q=1}^M C_{qk} \int \varphi_p^*(\mathbf{r}) \mathcal{F}(\mathbf{r}) \varphi_q(\mathbf{r}) d\mathbf{r} = \epsilon_k \sum_{q=1}^M C_{qk} \int \varphi_p^*(\mathbf{r}) \varphi_q(\mathbf{r}) d\mathbf{r}. \quad (1.80)$$

This equation can be rewritten if we introduce the overlap and Fock matrix with elements;

$$S_{pq} = \int \varphi_p^*(\mathbf{r}) \varphi_q(\mathbf{r}) d\mathbf{r} \quad (1.81a)$$

$$F_{pq} = \int \varphi_p^*(\mathbf{r}) \mathcal{F}(\mathbf{r}) \varphi_q(\mathbf{r}) d\mathbf{r}. \quad (1.81b)$$

Using these, we can rewrite the HF equation as

$$\sum_{q=1}^M F_{pq} C_{qk} = \epsilon_k \sum_{q=1}^M S_{pq} C_{qk} \quad (1.82)$$

which also is known as the Roothaan equation. This equation can be written more compactly as

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon} \quad (1.83)$$

where \mathbf{C} is a $M \times N/2$ matrix;

$$\mathbf{C} = \begin{pmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,N/2} \\ C_{2,1} & C_{2,2} & \cdots & C_{2,N/2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{M,1} & C_{M,2} & \cdots & C_{M,N/2} \end{pmatrix} \quad (1.84)$$

and $\boldsymbol{\epsilon}$ is a rectangular diagonal matrix of the orbital energies ϵ_k ;

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 & & & \\ & \epsilon_2 & & \mathbf{0} \\ & & \ddots & \\ \mathbf{0} & & & \end{pmatrix} \quad (1.85)$$

Note that, it is the columns of \mathbf{C} that describe the MO, i.e. the first column in \mathbf{C} are the coefficients of ϕ_1 , the second column are the the coefficients of ϕ_2 and so on.

The matrix representation of the Fock operator

$$\mathcal{F}(\mathbf{r}) = \hbar + \sum_{l=1}^{N/2} 2J_l(\mathbf{r}) - K_l(\mathbf{r}), \quad (1.86)$$

in the basis $\{\varphi_p\}$ is given by

$$\begin{aligned} F_{pq} &= \int \varphi_p^*(\mathbf{r}) \left[\hbar + \sum_{l=1}^{N/2} 2J_l(\mathbf{r}) - K_l(\mathbf{r}) \right] \varphi_q(\mathbf{r}) \\ &= h_{pq} + \sum_{l=1}^{N/2} \sum_{r,s=1}^M C_{rl}^* C_{sl} \left(2g_{prqs} - g_{prsq} \right) \end{aligned} \quad (1.87)$$

where

$$h_{pq} = \langle p | \hbar | q \rangle = \int \varphi_p^*(\mathbf{r}) \hbar \varphi_q(\mathbf{r}) d\mathbf{r} \quad (1.88)$$

and

$$g_{prqs} = \langle pr | \mathcal{G} | qs \rangle = \int \varphi_p^*(\mathbf{r}) \varphi_r^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \varphi_q(\mathbf{r}) \varphi_s(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \quad (1.89)$$

Note that, l labels the MOs ϕ , while p, q, r and s label the basis functions φ . Defining the density matrix;

$$P_{pq} = 2 \sum_l^{N/2} C_{pl} C_{ql}^*, \quad (1.90)$$

we can rewrite the Fock matrix as

$$F_{pq} = h_{pq} + \frac{1}{2} \sum_{rs}^M P_{rs} \left(2g_{prqs} - g_{prsq} \right). \quad (1.91)$$

By inserting the expansion in Eq. (1.78) into the energy expression in Eq. (1.76), we can express the energy in terms of the density matrix as well;

$$\begin{aligned}
E &= \sum_{k=1}^{N/2} \langle \phi_k | 2\hat{h} + 2\mathcal{J}(\mathbf{r}) - \mathcal{K}(\mathbf{r}) | \phi_k \rangle \\
&= \sum_{pq}^M P_{pq} h_{pq} + \frac{1}{2} \sum_{pqrs}^M P_{pq} P_{rs} \left(g_{prqs} - \frac{1}{2} g_{prsq} \right) \\
&= \frac{1}{2} \sum_{pq}^M P_{pq} (h_{pq} + F_{pq}) \tag{1.92}
\end{aligned}$$

The density matrix is directly related to the charge density, as it is shown in [5], and can be used to characterize the HF results. In practice when using the SCF procedure, we make a guess on the density matrix (or equivalently on the coefficients C_{pk}), which is equivalent to make a guess on the charge density. With our guess we construct the Fock operator and the Fock matrix and solve Eq. (1.83), to obtain a new and better set of coefficients, which can be used to calculate a new density matrix. This procedure can be repeated until self-consistency, as described earlier in this chapter.

1.5.3 Orthogonalization of the Basis

As mentioned in the last section, a basis set used to represent MOs is not required to be orthonormal. The only requirement is that the basis functions are normalized. The consequence of using a non-orthonormal basis set, is the occurrence of the overlap matrix \mathbf{S} in Eq. (1.83), which is a generalized eigenvalue equation. Of course, if the basis set is an orthonormal set, then the overlap matrix is just the identity matrix, and we are left with an ordinary eigenvalue equation, which can be solved by standard methods in linear algebra.

The generalized eigenvalue problem in Eq. (1.83) can be transformed to an ordinary eigenvalue problem, by performing a basis transformation that orthogonalize the basis. This basis transformation involves finding a transformation matrix \mathbf{V} that makes a transformed set of functions $\{\varphi'_q\}$ given by

$$\varphi'_q = \sum_p V_{pq} \varphi_p, \tag{1.93}$$

orthonormal, i.e.

$$\int d\mathbf{r} \varphi_p'^*(\mathbf{r}) \varphi'_q(\mathbf{r}) = \delta_{pq}. \tag{1.94}$$

To investigate the properties of the transformation matrix \mathbf{V} , we can insert the transformation in Eq. (1.93) into the last equation to get

$$\begin{aligned} \int d\mathbf{r} \varphi_p'^*(\mathbf{r}) \varphi_q'(\mathbf{r}) &= \int d\mathbf{r} \left[\sum_r V_{rp}^* \varphi_r^*(\mathbf{r}) \right] \left[\sum_s V_{sq} \varphi_s(\mathbf{r}) \right] \\ &= \sum_{rs} V_{rp}^* S_{rs} V_{sq} = \delta_{pq}, \end{aligned} \quad (1.95)$$

where the last line can be written as

$$\mathbf{V}^\dagger \mathbf{S} \mathbf{V} = \mathbf{1}. \quad (1.96)$$

By using the relation above we can rewrite the generalized eigenvalue equation

$$\begin{aligned} \mathbf{F} \mathbf{C} &= \mathbf{S} \mathbf{C} \epsilon \\ \mathbf{V}^\dagger \mathbf{F} \mathbf{V} \mathbf{V}^{-1} \mathbf{C} &= \mathbf{V}^\dagger \mathbf{S} \mathbf{V} \mathbf{V}^{-1} \mathbf{C} \epsilon \\ \mathbf{F}' \mathbf{C}' &= \mathbf{C}' \epsilon \end{aligned} \quad (1.97)$$

where $\mathbf{F}' = \mathbf{V}^\dagger \mathbf{F} \mathbf{V}$ and $\mathbf{C}' = \mathbf{V}^{-1} \mathbf{C}$. This equation is an ordinary eigenvalue equation which can be solved for \mathbf{C}' and thereafter transformed back to the original coefficient matrix \mathbf{C} by $\mathbf{V} \mathbf{C}'$.

The remaining problem is to find the transformation matrix \mathbf{V} which bring \mathbf{S} to unit form according to Eq. (1.96). One alternative, known as symmetric orthogonalization, is to define the transformation matrix as

$$\mathbf{V} \equiv \mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger \quad (1.98)$$

where \mathbf{s} is a diagonal matrix of the eigenvalues of the overlap matrix and \mathbf{U} is a unitary matrix that diagonalize the overlap matrix, i.e.

$$\mathbf{U}^\dagger \mathbf{S} \mathbf{U} = \mathbf{s}. \quad (1.99)$$

This unitary matrix do exist because of the Hermitian nature of the overlap matrix \mathbf{S} . Moreover, the eigenvalues of the overlap matrix are all positive, which follows directly from the definition of the overlap matrix. Therefore there is no difficulty in Eq. (1.98) of taking the square roots. Inserting the definition of the transformation matrix into Eq. (1.96), we obtain

$$\begin{aligned} (\mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger)^\dagger \mathbf{S} (\mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger) &= (\mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger) (\mathbf{U} \mathbf{s} \mathbf{U}^\dagger) (\mathbf{U} \mathbf{s}^{-1/2} \mathbf{U}^\dagger) \\ &= \mathbf{U} \mathbf{s}^{-1/2} \mathbf{s} \mathbf{s}^{-1/2} \mathbf{U}^\dagger = \mathbf{1}, \end{aligned} \quad (1.100)$$

which shows that \mathbf{V} indeed has the desired property.

A second alternative to orthogonalize the basis set, known as canonical orthogonalization, is to define the transformation matrix as

$$\mathbf{V} \equiv \mathbf{U}\mathbf{s}^{-1/2} \quad (1.101)$$

which also is an orthogonalizing transformation matrix;

$$(\mathbf{U}\mathbf{s}^{-1/2})^\dagger \mathbf{S} \mathbf{U} \mathbf{s}^{-1/2} = \mathbf{s}^{-1/2} \mathbf{U}^\dagger \mathbf{S} \mathbf{U} \mathbf{s}^{-1/2} = \mathbf{s}^{-1/2} \mathbf{s} \mathbf{s}^{1/2} = \mathbf{1}. \quad (1.102)$$

1.5.4 Unrestricted Hartree-Fock

Addition of an electron to a closed-shell system will turn the system into an open-shell system, where the new electron will interact differently with spin-up and spin-down electrons present in the system, since exchange is felt by parallel spin only. In such case using restricted spin-orbitals is limiting, since there is no guarantee that pairs of electrons prefer to be described by the same spatial orbital for different spins. In fact, it turns out that the energy is lowered when the spin-orbitals have different spatial orbital for different spins. This motivates the introduction of unrestricted spin-orbitals defined as

$$\psi_k(\mathbf{q}) = \begin{cases} \phi_k^\alpha(\mathbf{r})\alpha(s) \\ \text{or} \\ \phi_k^\beta(\mathbf{r})\beta(s) \end{cases} \quad (1.103)$$

That is, electrons of α spin are described by a set of spatial orbitals $\{\phi_k^\alpha\}$, while electrons of β spin are described by a different set $\{\phi_k^\beta\}$.

To derive the spatial form of the HF equation for unrestricted spin-orbitals we need to insert the definition in Eq. (1.103) into the general spin-orbital HF equation:

$$\begin{aligned} \mathcal{F}(\mathbf{q})\psi_k(\mathbf{q}) = \epsilon_k\psi_k(\mathbf{q}) \quad \Rightarrow \quad \begin{aligned} \mathcal{F}(\mathbf{r})\phi_k^\alpha(\mathbf{r})\alpha(s) &= \epsilon_k^\alpha\phi_k^\alpha(\mathbf{r})\alpha(s) \\ \mathcal{F}(\mathbf{r})\phi_k^\beta(\mathbf{r})\beta(s) &= \epsilon_k^\beta\phi_k^\beta(\mathbf{r})\beta(s) \end{aligned} \end{aligned} \quad (1.104)$$

where we have made distinction between the orbital energies of ϵ_k^α and ϵ_k^β , since the spatial orbitals are different. The procedure is the same as for the restricted spin-orbitals; multiplying from left by $\alpha^*(s)$ or $\beta^*(s)$, depending on if we are considering the equation for ϕ_k^α or ϕ_k^β , and integrating out the spin degrees of freedom,

$$\int \alpha^*(s) \mathcal{F}(\mathbf{r}) \alpha(s) ds \phi_k^\alpha(\mathbf{r}) = \epsilon_k^\alpha \phi_k^\alpha(\mathbf{r})$$

$$\int \beta^*(s) \mathcal{F}(\mathbf{r}) \beta(s) ds \phi_k^\beta(\mathbf{r}) = \epsilon_k^\beta \phi_k^\beta(\mathbf{r})$$
(1.105)

Now, an electron of α spin will have an effective interaction, consisting of Coulomb and exchange interaction with all α electrons, and only Coulomb interaction with electrons of β spin. An analogous argument can be used for electrons of β spin. Thus, the integrals on the right hand side can be written as

$$\mathcal{F}^\alpha(\mathbf{r}) = \int \alpha^*(s) \mathcal{F}(\mathbf{r}) \alpha(s) ds = \hbar + \sum_{l=1}^{N_\alpha} J_l^\alpha(\mathbf{r}) - K_l^\alpha(\mathbf{r}) + \sum_{l=1}^{N_\beta} J_l^\beta(\mathbf{r})$$

$$\mathcal{F}^\beta(\mathbf{r}) = \int \beta^*(s) \mathcal{F}(\mathbf{r}) \beta(s) ds = \hbar + \sum_{l=1}^{N_\beta} J_l^\beta(\mathbf{r}) - K_l^\beta(\mathbf{r}) + \sum_{l=1}^{N_\alpha} J_l^\alpha(\mathbf{r})$$
(1.106)

where the exchange and Coulomb operator are defined in analogy to our previous definitions in Eq. (1.73)

$$J_l^\alpha \phi^\alpha(\mathbf{r}) = \int \frac{\phi_l^{\alpha*}(\mathbf{r}') \phi_l^\alpha(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \phi^\alpha(\mathbf{r})$$
(1.107a)

$$K_l^\alpha \phi^\alpha(\mathbf{r}) = \int \frac{\phi_l^{\alpha*}(\mathbf{r}') \phi^\alpha(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \phi_l^\alpha(\mathbf{r})$$
(1.107b)

The definitions of J_l^β and K_l^β are analogous. Note that, the self-interaction is eliminated in both \mathcal{F}^α and \mathcal{F}^β , since

$$[J_l^\alpha - K_l^\alpha] \phi_l^\alpha(\mathbf{r}) = [J_l^\beta - K_l^\beta] \phi_l^\beta(\mathbf{r}) = 0.$$
(1.108)

We can now define the HF equations for unrestricted spin-orbitals as

$$\mathcal{F}^\alpha(\mathbf{r}) \phi_k^\alpha(\mathbf{r}) = \epsilon_k^\alpha \phi_k^\alpha(\mathbf{r})$$
(1.109a)

$$\mathcal{F}^\beta(\mathbf{r}) \phi_k^\beta(\mathbf{r}) = \epsilon_k^\beta \phi_k^\beta(\mathbf{r})$$
(1.109b)

These equations are, however, coupled and cannot be solved independently. This is because \mathcal{F}^α depends on β orbitals through J_l^β , while \mathcal{F}^β depends on α orbitals through J_l^α .

1.5.5 LCAO Approach: Pople-Nesbet Equations

In order to solve the unrestricted HF equations (1.109), we use the same philosophy as we used to derive the Roothaan equation; we introduce some known basis set and convert the unrestricted HF equations to matrix equations. More specific, we expand ϕ^α and ϕ^β in terms of a set $\{\varphi_p\}$;

$$\phi_k^\alpha(\mathbf{r}) = \sum_{q=1}^M C_{qk}^\alpha \varphi_q(\mathbf{r}), \quad k = 1, \dots, N_\alpha \quad (1.110a)$$

$$\phi_k^\beta(\mathbf{r}) = \sum_{q=1}^M C_{qk}^\beta \varphi_q(\mathbf{r}), \quad k = 1, \dots, N_\beta \quad (1.110b)$$

Substituting these expansion, in corresponding equations and multiplying with $\varphi_p^*(\mathbf{r})$ and integrating over coordinate \mathbf{r} , gives us the so-called Pople-Nesbet equations;

$$\sum_{q=1}^M F_{pq}^\alpha C_{qk}^\alpha = \epsilon_k^\alpha \sum_{q=1}^M S_{pq} C_{qk}^\alpha \quad (1.111a)$$

$$\sum_{q=1}^M F_{pq}^\beta C_{qk}^\beta = \epsilon_k^\beta \sum_{q=1}^M S_{pq} C_{qk}^\beta \quad (1.111b)$$

or more compactly

$$\mathbf{F}^\alpha \mathbf{C}^\alpha = \mathbf{S} \mathbf{C}^\alpha \boldsymbol{\epsilon}^\alpha \quad (1.112a)$$

$$\mathbf{F}^\beta \mathbf{C}^\beta = \mathbf{S} \mathbf{C}^\beta \boldsymbol{\epsilon}^\beta \quad (1.112b)$$

where \mathbf{F}^α and \mathbf{F}^β are the matrix representation of \mathcal{F}^α and \mathcal{F}^β operators, \mathbf{C}^α and \mathbf{C}^β are the coefficient matrices with dimensions $M \times N_\alpha$ and $M \times N_\beta$ respectively, \mathbf{S} is the overlap matrix, and ϵ^α and ϵ^β are the rectangular diagonal matrices of orbital energies for the α and β orbitals.

By defining the density matrices

$$P_{pq}^\alpha = \sum_l^{N_\alpha} C_{pl}^\alpha (C_{ql}^\alpha)^* \quad (1.113a)$$

$$P_{pq}^\beta = \sum_l^{N_\beta} C_{pl}^\beta (C_{ql}^\beta)^* \quad (1.113b)$$

we can express the Fock matrix elements as⁷

$$F_{pq}^{\alpha} = h_{pq} + \sum_{rs}^M P_{rs}^{\alpha} \left(g_{prqs} - g_{prsq} \right) + \sum_{rs}^M P_{rs}^{\beta} g_{prqs} \quad (1.114a)$$

$$F_{pq}^{\beta} = h_{pq} + \sum_{rs}^M P_{rs}^{\beta} \left(g_{prqs} - g_{prsq} \right) + \sum_{rs}^M P_{rs}^{\alpha} g_{prqs} \quad (1.114b)$$

and the energy functional as

$$E = \frac{1}{2} \sum_{pq}^M \left[(P_{pq}^{\alpha} + P_{pq}^{\beta}) h_{pq} + P_{pq}^{\alpha} F_{pq}^{\alpha} + P_{pq}^{\beta} F_{pq}^{\beta} \right] \quad (1.115)$$

In order to solve the matrix eigenvalue problems in Eq. (1.112), we must orthonormalize the basis set and diagonalize the Fock matrix in the new orthonormal basis, and thereafter transform the resulting coefficient matrices back to the old basis, just as described in section 1.5.3.

⁷For more detailed derivation see chapter 3.8 in [5].

Appendices

Appendix A

Mathematics

A.1 Atomic units

Atomic units are designed to simplify the appearance of fundamental equations in quantum mechanics, which in their original form consist of many small valued constants. Beside the aesthetic aspect, use of atomic units is critical from a programming point of view as well, to avoid the effects of round off errors.

The origin of atomic units is the requirement on four of the most fundamental physical constants:

- Electron mass: $m_e = 1$
- Elementary charge : $e = 1$
- Reduced Planck's constant: $\hbar = 1$
- Coulomb's constant: $\frac{1}{4\pi\epsilon_0} = 1$

By using these constraints, we can derive the atomic unit of some of the most important physical quantities [7]:

- **Unit of length:** $a_0 \equiv \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.5291772192(17)\text{\AA}$
- **Unit of energy:** $E_h \equiv \frac{m_e e^4}{(4\pi\epsilon_0\hbar)^2} = 27.211\text{eV}$
- **Unit of time:** $\tau_0 \equiv \frac{\hbar}{E_h} = 2.418884326505(16) \cdot 10^{-17}\text{s}$
- **Unit of velocity:** $v_0 \equiv \frac{a_0 E_h}{\hbar} = 2.1876912633(73) \cdot 10^6\text{m/s}$

These quantities are the conversion factor for their respective physical quantity, such that the SI value, Q , of each of them is related to their value in atomic units Q' by

$$Q = XQ' \quad (\text{A.1})$$

where X is the conversion factor.

The electronic Hamiltonian is in its original form (in SI units) given as

$$\mathcal{H} = - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m_e} \nabla_i^2 + \frac{1}{4\pi\epsilon_0} \sum_{i<j}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \frac{1}{4\pi\epsilon_0} \sum_{n=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_n e^2}{|\mathbf{r}_i - \mathbf{R}_n|}. \quad (\text{A.2})$$

This expression can be brought to dimensionless form, by introducing the following relations;

$$\mathcal{H}' = \mathcal{H}/E_h, \quad \nabla' = \nabla/a_0, \quad \mathbf{r}' = \mathbf{r}/a_0, \quad \mathbf{R}' = \mathbf{R}/a_0. \quad (\text{A.3})$$

By using these relations in Eq. (A.2), we obtain

$$\begin{aligned} \mathcal{H} &= -\frac{\hbar^2}{2m_e a_0^2} \sum_i^N (\nabla'_i)^2 - \frac{e^2}{4\pi\epsilon_0 a_0} \left[\sum_i^N \frac{Z_n}{|\mathbf{r}'_i - \mathbf{R}'_n|} + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{1}{|\mathbf{r}'_i - \mathbf{r}'_j|} \right] \\ &= E_h \left(-\frac{1}{2} \sum_i^N (\nabla'_i)^2 - \sum_i^N \frac{Z_n}{|\mathbf{r}'_i - \mathbf{R}'_n|} + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{1}{|\mathbf{r}'_i - \mathbf{r}'_j|} \right) \\ &= E_h \mathcal{H}' \end{aligned} \quad (\text{A.4})$$

The electronic Schrödinger equation can therefor be written as

$$\mathcal{H}' \Psi(\{\mathbf{r}'_i\}; \{\mathbf{R}'_n\}) = E' \Psi(\{\mathbf{r}'_i\}; \{\mathbf{R}'_n\}) \quad \text{where} \quad E' = \frac{E}{E_h} \quad (\text{A.5})$$

In this dimensionless form of Schrödinger equation, the energy will be measured in the units of Hartree (or a.u.), where $1E_h = 2 \cdot 13.6\text{eV}$ is twice the ground state energy of hydrogen atom.

A.2 Antisymmetrizer

The antisymmetrizer operator \mathcal{A} is a linear operator, defined as

$$\mathcal{A} \equiv \frac{1}{N!} \sum_P (-1)^{P_n} \mathcal{P}, \quad (\text{A.6})$$

where \mathcal{P} is the permutation operator and P_n is the parity of the permutation. By applying this operator on a wave function on the form

$$\Psi_H = \psi_1(\mathbf{q}_1) \psi_2(\mathbf{q}_2) \dots \psi_N(\mathbf{q}_N), \quad (\text{A.7})$$

the wave function will be antisymmetric with respect to exchange of coordinates of any pair of electrons, i.e.

$$\begin{aligned} & \mathcal{P}_{ij} \Psi(\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N) \\ &= \mathcal{P}_{ij} \left[\sqrt{N!} \mathcal{A} \Psi_H(\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N) \right] \\ &= -\sqrt{N!} \mathcal{A} \Psi_H(\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N) \\ &= -\Psi(\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_j, \dots, \mathbf{q}_N). \end{aligned} \quad (\text{A.8})$$

where $\sqrt{N!}$ is included to ensure normalization of the wave function. The antisymmetrizer operator has several interesting properties¹:

- **Idempotent:**

The square of the operator is itself, i.e. $\mathcal{A}^2 = \mathcal{A}$.

Proof:

$$\mathcal{A}\mathcal{A} = \left(\frac{1}{N!} \right)^2 \sum_P \sum_Q (-1)^{P_n+Q_n} \mathcal{P}\mathcal{Q} \quad (\text{A.9})$$

For a given \mathcal{P} , the product $\mathcal{R} = \mathcal{P}\mathcal{Q}$ runs over all $N!$ permutations, so that

$$\mathcal{A}\mathcal{A} = \left(\frac{1}{N!} \right)^2 \sum_P \underbrace{\left[\sum_{\mathcal{R}} (-1)^{R_n} \mathcal{R} \right]}_{N!\mathcal{A}} = \left(\frac{1}{N!} \right)^2 N!^2 \mathcal{A} = \mathcal{A} \quad (\text{A.10})$$

¹The proofs are based on the discussion in chapter 7 in [8].

- **Hermitian:**

Antisymmetrizer is self-adjoint, i.e. $\mathcal{A}^\dagger = \mathcal{A}$.

Proof:

Permutations of identical particles are unitary, i.e. $\mathcal{P}^\dagger = \mathcal{P}^{-1}$. By taking the adjoint of \mathcal{A} , we obtain

$$\mathcal{A}^\dagger = \frac{1}{N!} \sum_{P^\dagger} (-1)^{P_n} \mathcal{P}^\dagger = \frac{1}{N!} \sum_{P^{-1}} (-1)^{P_n} \mathcal{P}^{-1} \quad (\text{A.11})$$

Now, \mathcal{P} and \mathcal{P}^{-1} perform the same operations, only in inverse order, and therefore will their parity be the same. This means that \mathcal{A} and \mathcal{A}^\dagger have the same action, so $\mathcal{A} = \mathcal{A}^\dagger$.

- **Commutation relation with \mathcal{H} :**

The antisymmetrizer commutes with the Hamiltonian, i.e. $[\mathcal{A}, \mathcal{H}] = 0$

Proof:

Since \mathcal{H} is symmetric in the system coordinates, it will commute with each individual permutation, and therefore also with sum of such permutations. So $[\mathcal{A}, \mathcal{H}] = 0$.

Bibliography

- [1] D. Griffiths, *Introduction to Quantum Mechanics*. Pearson, 2nd edition ed., 2005.
- [2] R. Shankar, *Principles of Quantum Mechanics*. Springer, 1994.
- [3] C. Cramer, *Essentials of Computational Chemistry: Theories and Models*. Wiley, 2005.
- [4] J. Thijssen, *Computational Physics*. Cambridge University Press, 2007.
- [5] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Books on Chemistry Series, Dover Publications, 1996.
- [6] M. Hjorth-Jensen, “Computational Physics.” 2010.
- [7] Wikipedia, “Atomic units — wikipedia, the free encyclopedia,” 2013. [Online; accessed 15-January-2014].
- [8] F. Pilar, *Elementary quantum chemistry*. No. v. 1 in Schaum’s outline series in science, McGraw-Hill, 1990.