Chapter 3

Basis functions in quantum chemistry

This chapter is adapted from Chapter 5 of Jensen’s book: F. Jensen, ’Introduction to Computational Chemistry’, Wiley.

In the derivation in the previous chapter, we have introduced the concept of basis function for the expansion of the one-electron molecular orbitals used for the generation of the many-electrons wave functions (Slater determinants or linear combination of Slater determinants). There we derived the following expansion (eq. 2.17):

$$\phi_m(r, s) = \sum_n D_{mn} \chi_n(r)$$

(where $\chi_n$ is an atom centered basis function and the spin dependent part of the wavefunctions is left out).

In this chapter, we introduce the different basis functions, $\chi_n$ commonly used in computational quantum chemistry.

Finiteness of Basis Sets: Approximations

One of the approximations inherent in essentially all ab initio methods is the introduction of a finite basis set. Expanding an unknown function, such as a molecular orbital, in a set of known functions is not an approximation, if the basis is complete. However, a complete basis means that an infinite number of functions must be used, which is impossible in actual calculations. An unknown MO can be thought of as a function in the infinite coordinate system spanned by the complete basis set. When a finite basis is used, only the components of the MO along those coordinate axes corresponding to the selected basis can be represented. The smaller the basis, the poorer the representation. The type of basis functions used also influences the accuracy. The better a single basis function is able to reproduce the unknown function,
the fewer basis functions necessary for achieving a given level of accuracy. Knowing that the computational effort of ab initio methods scales formally at least as $M^4$, it is of course of prime importance to make the basis set as small as possible without compromising the accuracy.

### 3.1 Slater and Gaussian Type Orbitals

There are two types of basis functions (also called Atomic Orbitals, AO, although in general they are not solutions to an atomic Schrödinger equation) commonly used in electronic structure calculations: Slater Type Orbitals (STO) and Gaussian Type Orbitals (GTO).

A procedure that has come into wide use is to fit a Slater-type orbital (STO) to a linear combination of $n = 1, 2, 3, \ldots$ primitive Gaussian functions. This is the STO-$nG$ procedure. In particular, STO-3G basis sets are often used in polyatomic calculations, in preference to evaluating integrals with Slater functions.

![Comparison of Slater function with Gaussian function: least squares fits of a 1s Slater function ($\zeta = 1.0$) by a n GTOs](image)

#### 1. Slater type orbitals

have the functional form

$$
\chi_{\zeta,n,l,m}(r, \theta, \varphi) = N Y_{l,m}(\theta, \varphi) r^{n-1} e^{-\zeta r} 
$$

(3.2)

$N$ is a normalization constant and $Y_{l,m}$ are the usual spherical harmonic functions. The exponential dependence on the distance between the nucleus and the electron mirrors the exact decay behavior of the orbitals for the hydrogen atom. However, since STOs do not have any radial nodes, nodes in the radial part are introduced by making linear combinations of STOs. The exponential dependence ensures a fairly rapid convergence with increasing
number of functions, however, the calculation of three- and four-centre two-electron integrals cannot be performed analytically. STOs are primarily used for atomic and diatomic systems where high accuracy is required, and in semi-empirical methods where all three- and four-center integrals are neglected.

2. Gaussian type orbitals can be written in terms of polar or cartesian coordinates

\[ \chi_{\zeta,n,l,m}(r,\theta,\phi) = N Y_{l,m}(\theta,\phi) r^{2n-2-l} e^{-\zeta r^2} \]  
\[ \chi_{\zeta,l_x,l_y,l_z}(x,y,z) = N x^{l_x} y^{l_y} z^{l_z} e^{-\zeta r^2} \]

where the sum of \( l_x, l_y \) and \( l_z \) determines the type of orbital (for example \( l_x + l_y + l_z = 1 \) is a p-orbital)\(^1\).

3. Comparison between STO and GTO

i. The \( r^2 \) dependence in the exponent makes the GTOs inferior to the STOs in two aspects. At the nucleus the GTO has zero slope, in contrast to the STO which has a "cusp" (discontinuous derivative), and GTOs have problems representing the proper behavior near the nucleus.

ii. The other problem is that the GTO falls off too rapidly far from the nucleus compared with an STO, and the "tail" of the wave function is consequently represented poorly.

iii. Both STOs and GTOs can be chosen to form a complete basis, but the above considerations indicate that more GTOs are necessary for achieving a certain accuracy compared with STOs. A rough guideline says that three times as many GTOs as STOs are required for reaching a given level of accuracy. The increase in number of basis functions, however, is more than compensated for by the ease by which the required

\(^1\)Although a GTO appears similar in the two sets of coordinates, there is a subtle difference. A d-type GTO written in terms of the spherical functions has five components \((Y_{2,2}, Y_{2,1}, Y_{2,0}, Y_{2,-1}, Y_{2,-2})\), but there appear to be six components in the Cartesian coordinates \((x_2, y_2, z_2, x y, x z, y z)\). The latter six functions, however, may be transformed to the five spherical d-functions and one additional s-function \((x^2 + y^2 + z^2)\). Similarly, there are 10 Cartesian "f-functions" which may be transformed into seven spherical f-functions and one set of spherical p-functions. Modern programs for evaluating two-electron integrals are geared to Cartesian coordinates, and they generate pure spherical d-functions by transforming the six Cartesian components to the five spherical functions. When only one d-function is present per atom the saving by removing the extra s-function is small, but if many d-functions and/or higher angular moment functions (f-, g-, h- etc. functions) are present, the saving can be substantial. Furthermore, the use of only the spherical components reduces the problems of linear dependence for large basis sets, as discussed below.
integrals can be calculated. In terms of computational efficiency, GTOs are therefore preferred, and used almost universally as basis functions in electronic structure calculations.

3.2 Classification of Basis Sets

Having decided on the type of basis function (STO/GTO) and their location (nuclei), the most important factor is the number of functions to be used. The smallest number of functions possible is a minimum basis set. Only enough atomic orbital functions are employed to contain all the electrons of the neutral atom(s).

3.2.1 Minimum basis sets. Examples

For hydrogen (and helium) this means a single s-function. For the first row in the periodic table it means two s-functions (1s and 2s) and one set of p-functions (2px, 2py, and 2pz). Lithium and beryllium formally only require two s-functions, but a set of p-functions is usually also added. For the second row elements, three s-functions (1s, 2s, and 3s) and two sets of p-functions (2p and 3p) are used.

3.2.2 Improvements

1. The first improvement in the basis sets is a doubling of all basis functions, producing a Double Zeta (DZ) type basis. The term zeta stems from the fact that the exponent of STO basis functions is often denoted by the greek letter $\zeta$.

   A DZ basis thus employs two s-functions for hydrogen (1s and 1s'), four s-functions (1s, 1s', 2s, and 2s') and two p-functions (2p and 2p') for first row elements, and six s-functions and four p-functions for second row elements. Doubling the number of basis functions allows for a much better description of the fact that the electron distribution in molecules can differ significantly from the one in the atoms and the chemical bond may introduce directionalities which can not be described by a minimal basis.

   The chemical bonding occurs between valence orbitals. Doubling the 1s-functions in for example carbon allows for a better description of the 1s-electrons. However, the 1s orbital is essentially independent of the chemical environment, being very close to the atomic case. A variation of the DZ type basis only doubles the number of valence orbitals, producing a split valence basis.  

   2In actual calculations a doubling of the core orbitals would rarely be considered, and the term DZ basis is also used for split valence basis sets (or sometimes VDZ, for valence double zeta)
2. The next step up in basis set size is a **Triple Zeta** (TZ) basis. Such a basis contains three times as many functions as the minimum basis, i.e. six s-functions and three p-functions for the first row elements. Some of the core orbitals may again be saved by only splitting the valence, producing a triple zeta split valence basis set. The names **Quadruple Zeta** (QZ) and **Quintuple Zeta** (5Z, not QZ) for the next levels of basis sets are also used, but large sets are often given explicitly in terms of the number of basis functions of each type.

3. In most cases higher angular momentum functions are also important, these are denoted **polarization functions**. Consider for example a C-H bond which is primarily described by the hydrogen s-orbital(s) and the carbon s- and p- orbitals. It is clear that the electron distribution along the bond will be different than that perpendicular to the bond. If only s-functions are present on the hydrogen, this cannot be described. However, if a set of p-orbitals is added to the hydrogen, the p component can be used for improving the description of the H-C bond. The p-orbital introduces a polarization of the s-orbital(s). Similarly, d-orbitals can be used for polarizing p-orbitals, f-orbitals for polarizing d-orbitals etc. Once a p-orbital has been added to a hydrogen s-orbital, it may be argued that the p-orbital now should be polarized by adding a d-orbital, which should be polarized by an f-orbital, etc. For single determinant wave functions, where electron correlation is not considered, the first set of polarization functions (i.e. p-functions for hydrogen and d-functions for heavy atoms) is by far the most important, and will in general describe all the important charge polarization effects.

Adding a single set of polarization functions (p-functions on hydrogens and d-functions on heavy atoms) to the DZ basis forms a **Double Zeta plus Polarization** (DZP) type basis. Similarly to the sp-basis sets, multiple sets of polarization functions with different exponents may be added. If two sets of polarization functions are added to a TZ sp-basis, a Triple Zeta plus Double Polarization (TZ2P) type basis is obtained. For larger basis sets with many polarization functions the explicit composition in terms of number and types of functions is usually given. At the HF level there is usually little gained by expanding the basis set beyond TZ2P, and even a DZP type basis set usually gives "good" results (compared to the HF limit).

---

3There is a variation where polarization functions are only added to non-hydrogen atoms. This does not mean that polarization functions are not important on hydrogens. However, hydrogens often have a "passive" role, sitting at the end of bonds which does not take an active part in the property of interest. The errors introduced by not including hydrogen polarization functions are often rather constant and, as the interest is usually in energy differences, they tend to cancel out. As hydrogens often account for a large number of atoms in the system, a saving of three basis functions for each hydrogen is significant. If hydrogens play an important role in the property of interest, it is of course not a good idea to neglect polarization functions on hydrogens.
3.3 Basis set balance

In principle many sets of polarization functions may be added to a small sp-basis. This is not a good idea. If an insufficient number of sp-functions has been chosen for describing the fundamental electron distribution, the optimization procedure used in obtaining the wave function (and possibly also the geometry) may try to compensate for inadequacies in the sp-basis by using higher angular momentum functions, producing artefacts. A rule of thumb says that the number of functions of a given type should at most be one less than the type with one lower angular momentum. A 3s2p1d basis is balanced, but a 3s2p2d2f1g basis is too heavily polarized.

Another aspect of basis set balance is the occasional use of mixed basis sets, for example a DZP quality on the atoms in the "interesting" part of the molecule and a minimum basis for the "spectator" atoms. Another example would be addition of polarization functions for only a few hydrogens which are located "near" the reactive part of the system. For a large molecule this may lead to a substantial saving in the number of basis functions. It should be noted that this may bias the results and can create artefacts. For example, a calculation on the H₂ molecule with a minimum basis at one end and a DZ basis at the other end will predict that H₂ has a dipole moment, since the variational principle will preferentially place the electrons near the center with the most basis functions. The majority of calculations are therefore performed with basis sets of the same quality (minimum, DZP, TZ2P, . . .) on all atoms, possibly cutting polarization and/or diffuse (small exponent) functions on hydrogens.

Except for very small systems it is impractical to saturate the basis set so that the absolute error in the energy is reduced below chemical accuracy, for example 1 kcal/mol. The important point in choosing a balanced basis set is to keep the error as constant as possible. The use of mixed basis sets should therefore only be done after careful consideration. Furthermore, the use of small basis sets for systems containing elements with substantially different numbers of valence electrons (like LiF) may produce artefacts.

3.4 How do we choose the exponents in the basis functions?

The values for s- and p-functions are typically determined by performing variational HF calculations for atoms, using the exponents as variational parameters. The exponent values which give the lowest energy are the "best", at least for the atom. In some cases the optimum exponents are chosen on the basis of minimizing the energy of a wave function which includes electron correlation. The HF procedure cannot be used for determining exponents of polarization functions for atoms. By definition these functions are unoccupied in atoms, and therefore make no contribution to the energy. Suitable polarization exponents may be chosen by performing variational calculations on molecular systems (where the HF energy does depend on polarization functions) or on atoms with correlated wave functions. Since the main func-
tion of higher angular momentum functions is to recover electron correlation, the latter approach is usually preferred. Often only the optimum exponent is determined for a single polarization function, and multiple polarization functions are generated by splitting the exponents symmetrically around the optimum value for a single function. The splitting factor is typically taken in the range 2-4. For example if a single d-function for carbon has an exponent value of 0.8, two polarization functions may be assigned with exponents of 0.4 and 1.6 (splitting factor of 4).

3.5 Contracted Basis functions

One disadvantage of all energy optimized basis sets is the fact that they primarily depend on the wave function in the region of the inner shell electrons. The 1s-electrons account for a large part of the total energy, and minimizing the energy will tend to make the basis set optimal for the core electrons, and less than optimal for the valence electrons. However, chemistry is mainly dependent on the valence electrons. Furthermore, many properties (for example polarizability) depend mainly on the wave function "tail" (far from the nucleus), which energetically is unimportant. An energy optimized basis set which gives a good description of the outer part of the wave function needs therefore to be very large, with the majority of the functions being used to describe the 1s-electrons with an accuracy comparable to that for the outer electrons in an energetic sense. This is not the most efficient way of designing basis sets for describing the outer part of the wave function. Instead energy optimized basis sets are usually augmented explicitly with diffuse functions (basis functions with small exponents). Diffuse functions are needed whenever loosely bound electrons are present (for example in anions or excited states) or when the property of interest is dependent on the wave function tail (for example polarizability). The fact that many basis functions go into describing the energetically important, but chemically unimportant, core electrons is the foundation for contracted basis sets.

An example. The carbon atom Consider for example a basis set consisting of 10 s-functions (and some p-functions) for carbon. Having optimized these 10 exponents by variational calculations on the carbon atom, maybe six of the 10 functions are found primarily to be used for describing the 1s orbital, and two of the four remaining describe the "inner" part of the 2s-orbital. The important chemical region is the outer valence. Out of the 10 functions, only two are actually used for describing the chemically interesting phenomena. Considering that the computational cost increases as the fourth power (or higher) of the number of basis functions, this is very inefficient. As the core orbitals change very little depending on the chemical bonding situation, the MO expansion coefficients in front of these inner basis functions also change very little. The majority of the computational effort is therefore spent describing the chemically uninteresting part of the wave function, which furthermore is almost constant. Consider now making the varia-
tional coefficients in front of the inner basis functions constant, i.e. they are no longer parameters to be determined by the variational principle. The 1s-orbital is thus described by a fixed linear combination of say six basis functions. Similarly the remaining four basis functions may be contracted into only two functions, for example by fixing the coefficient in front of the inner three functions. In doing this the number of basis functions to be handled by the variational procedure has been reduced from 10 to three.

Combining the full set of basis functions, known as the primitive GTOs (PGTOs), into a smaller set of functions by forming fixed linear combinations is known as basis set contraction, and the resulting functions are called contracted GTOs (CGTOs)

\[ \chi(\text{CGTO}) = \sum_{i} a_i \chi_i(\text{PGTO}) \]  

The previously introduced acronyms DZP, TZ2P etc., refer to the number of contracted basis functions. Contraction is especially useful for orbitals describing the inner (core) electrons, since they require a relatively large number of functions for representing the wave function cusp near the nucleus, and furthermore are largely independent of the environment. Contracting a basis set will always increase the energy, since it is a restriction of the number of variational parameters, and makes the basis set less flexible, but will also reduce the computational cost significantly. The decision is thus how much loss in accuracy is acceptable compared to the gain in computational efficiency.

### 3.5.1 The degree of contraction

The degree of contraction is the number of PGTOs entering the CGTO, typically varying between 1 and 10. The specification of a basis set in terms of primitive and contracted functions is given by the notation

\[ (10s4p1d/4s1p) \rightarrow [3s2p1d/2s1p] \]  

The basis in parentheses is the number of primitives with heavy atoms (first row elements) before the slash and hydrogen after. The basis in the square brackets is the number of contracted functions. Note that this does not tell how the contraction is done, it only indicates the size of the final basis (and thereby the size of the variational problem in HF calculations).

### 3.6 Example of Contracted Basis Sets; Pople Style Basis Sets

There are many different contracted basis sets available in the literature or built into programs, and the average user usually only needs to select a suitable quality basis for the calculation. For short description of some basis
sets which often are used in routine calculations (see for instance the book of Frank Jensen, *Introduction to Computational Chemistry*, Wiley, 2002. Chapter 5).

**STO-nG basis sets**  $n$ PGTOs fitted to a 1 STO. This is a minimum type basis where the exponents of the PGTO are determined by fitting to the STO, rather than optimizing them by a variational procedure. Although basis sets with $n = 2 - 6$ have been derived, it has been found that using more than three PGTOs to represent the STO gives little improvement, and the STO-3G basis is a widely used minimum basis. This type of basis set has been determined for many elements of the periodic table. The designation of the carbon/hydrogen STO-3G basis is $(6s^3p/3s) \rightarrow [2s1p/1s]$.

**k-nlmG basis sets**  These basis sets have been designed by Pople and coworkers, and are of the split valence type, with the $k$ in front of the dash indicating how many PGTOs are used for representing the core orbitals. The $nlm$ after the dash indicate both how many functions the valence orbitals are split into, and how many PGTOs are used for their representation. Two values (e.g. $nl$) indicate a split valence, while three values (e.g. $nlm$) indicate a triple split valence. The values before the G (for Gaussian) indicate the s- and p-functions in the basis; the polarization functions are placed after the G. This type of basis sets has the further restriction that the same exponent is used for both the s- and p-functions in the valence. This increases the computational efficiency, but of course decreases the flexibility of the basis set. The exponents in the PGTO have been optimized by variational procedures.

**3-21G**  This is a split valence basis, where the core orbitals are a contraction of three PGTOs, the inner part of the valence orbitals is a contraction of two PGTOs and the outer part of the valence is represented by one PGTO. The designation of the carbon/hydrogen 3-21G basis is $(6s^3p/3s) \rightarrow [3s^2p/2s]$. Note that the 3-21G basis contains the same number of primitive GTOs as the STO-3G, however, it is much more flexible as there are twice as many valence functions which can combine freely to make MOs.

**6-31G**  This is also a split valence basis, where the core orbitals are a contraction of six PGTOs, the inner part of the valence orbitals is a contraction of three PGTOs and the outer part of the valence represented by one PGTO. The designation of the carbon/hydrogen 6-31G basis is $(10s^4p/4s) \rightarrow [3s^2p/2s]$. In terms of contracted basis functions it contains the same number as 3-21G, but the representation of each functions is better since more PGTOs are used.

**6-311G**  This is a triple zeta split valence basis, where the core orbitals are a contraction of six PGTOs and the valence split into three functions, represented by three, one, and one PGTOs, respectively.
To each of these basis sets one can add diffuse and/or polarization functions.

- **Diffuse functions** are normally s- and p-functions and consequently go before the G. They are denoted by + or ++, with the first + indicating one set of diffuse s- and p-functions on heavy atoms, and the second + indicating that a diffuse s-function is also added to hydrogens. The arguments for adding only diffuse functions on non-hydrogen atoms is the same as that for adding only polarization functions on non-hydrogens.

- **Polarization functions** are indicated after the G, with a separate designation for heavy atoms and hydrogens. The 6-31+G(d) is a split valence basis with one set of diffuse sp-functions on heavy atoms only and a single d-type polarization function on heavy atoms. A 6-311++G(2df,2pd) is similarly a triple zeta split valence with additional diffuse sp-functions, and two d- and one f-functions on heavy atoms and diffuse s- and two p- and one d-functions on hydrogens. The largest standard Pople style basis set is 6-311++G(3df, 3pd). These types of basis sets have been derived for hydrogen and the first row elements, and same of the basis sets have also been derived for second and higher row elements.

If only one set of polarization functions is used, an alternative notation in terms of * is also widely used. The 6-31G* basis is identical to 6-31G(d), and 6-31G** is identical to 6-31G(d,p). A special note should be made for the 3-21G* basis. The 3-21G basis is basically too small to support polarization functions (it becomes unbalanced). However, the 3-21G basis by itself performs poorly for hypervalent molecules, such as sulfides and sulfoxides. This can be substantially improved by adding a set of d-functions. The 3-21G* basis has only d-functions on second row elements (it is sometimes denoted 3-21G(*) to indicate this), and should not be considered a polarized basis. Rather, the addition of a set of d-functions should be considered an ad hoc repair of a known flaw.
Chapter 4

An Introduction to Hartree Fock Theory

Adapted from C. D. Sherrill’s notes: "An Introduction to Hartree-Fock Molecular Orbital Theory".

4.1 Introduction

Hartree-Fock theory is fundamental to much of electronic structure theory. It is the basis of molecular orbital (MO) theory, which posits that each electron’s motion can be described by a single-particle function (orbital) which does not depend explicitly on the instantaneous motions of the other electrons. Some of you have probably learned about (and maybe even solved problems with) Hückel MO theory, which takes Hartree-Fock MO theory as an implicit foundation and throws away most of the terms to make it tractable for simple calculations. The ubiquity of orbital concepts in chemistry is a testimony to the predictive power and intuitive appeal of Hartree-Fock MO theory. However, it is important to remember that these orbitals are mathematical constructs which only approximate reality. Only for the hydrogen atom (or other one-electron systems, like He⁺) are orbitals exact eigenfunctions of the full electronic Hamiltonian. As long as we are content to consider molecules near their equilibrium geometry, Hartree-Fock theory often provides a good starting point for more elaborate theoretical methods which are better approximations to the electronic Schrödinger equation (e.g., many-body perturbation theory, single-reference configuration interaction). So, how do we calculate molecular orbitals using Hartree-Fock theory? This is the subject of these notes; we will explain Hartree-Fock theory at an introductory level.
4.2 What Problem Are We Solving?

It is always important to remember the context of a theory. Hartree-Fock theory was developed to solve the electronic Schrödinger equation that results from the time-independent Schrödinger equation after invoking the Born-Oppenheimer approximation. In atomic units, and with $r$ denoting electronic and $R$ denoting nuclear degrees of freedom, the electronic Schrödinger equation is

$$\left[ -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{I} \frac{Z_I}{|R_I - r_i|} + \sum_{I>j} \frac{Z_I Z_J}{|R_I - R_J|} + \sum_{i>j} \frac{1}{|r_i - r_j|} \right] \Psi(r, R) = E_{el} \Psi(r, R)$$

(4.1)

or, in a more compact notation,

$$\left[ \hat{T}_e(r) + \hat{V}_{eN}(r, R) + \hat{V}_{NN}(R) + \hat{V}_{ee}(r) \right] \Psi(r, R) = E_{el} \Psi(r, R)$$

(4.2)

Recall from the Born-Oppenheimer approximation that $E_{el}$ (plus or minus $\hat{V}_{NN}(R)$, which we include here) will give us the potential energy experienced by the nuclei. In other words, $E_{el}(R)$ the potential energy surface (from which we can get, for example, the equilibrium geometry and the vibrational frequencies). That’s one good reason why we want to solve the electronic Schrödinger equation. The other is that the electronic wavefunction $\Psi(r, R)$ contains lots of useful information about molecular properties such as dipole (and multipole) moments, polarizability, etc.

4.3 The many-electron wavefunction: the Slater determinant

The basic idea of Hartree-Fock theory is as follows. We know how to solve the electronic problem for the simplest atom, hydrogen, which has only one electron. We imagine that perhaps if we added another electron to hydrogen, to obtain $H^-$, then maybe it might be reasonable to start off pretending that the electrons don’t interact with each other (i.e., that $\hat{V}_{ee} = 0$). If that was true, then the Hamiltonian would be separable, and the total electronic wavefunction $\Psi(r_1, r_2)$ describing the motions of the two electrons would just be the product of two hydrogen atom wavefunctions (orbitals), $\Psi_H(r_1) \Psi_H(r_2)$ (you should be able to prove this easily).

However, we have already shown in the previous chapter that a correct many electron wavefunction must fulfill the antisymmetry principle together with the principle of indistinguishability of the electrons. The solution of this problem is the Slater determinant made of one-electron molecular orbitals (MOs). In the following we will derive a theory, the Hartree Fock theory, for the calculation of such single electron MOs.
4.4 Simplified Notation for the Hamiltonian

Hartree-Fock theory is based on the assumption that the many-electron wavefunction of the system can be described by a single Slater determinant made of one-electron molecular orbitals. Let’s re-examine the Hamiltonian to make it look as simple as possible. In the process, we will bury some complexity that would have to be taken care of later (in the evaluation of integrals).

We will define a one-electron operator $\hat{h}$ as follows

$$\hat{h}(i) = -\frac{1}{2} \nabla_i^2 - \sum_l \frac{Z_l}{|r_i - R_l|}$$

(4.3)

and a two-electron operator $\hat{v}(i, j)$ as

$$\hat{v}(i, j) = \frac{1}{|r_i - r_j|}$$

(4.4)

Now we can write the electronic Hamiltonian much more simply, as

$$\hat{H}_{el} = \sum_i \hat{h}(i) + \sum_{i<j} \hat{v}(i, j) + V_{NN}$$

(4.5)

Since $V_{NN}$ is just a constant for the fixed set of nuclear coordinates $\{R\}$, it doesn’t change the eigenfunctions, and only shifts the eigenvalues.

4.5 Energy Expression

Now that we have a form for the wavefunction and a simplified notation for the Hamiltonian, we have a good starting point to tackle the problem. Still, how do we obtain the molecular orbitals?

We state that the Hartree-Fock wavefunction will have the form of a Slater determinant, and that the energy will be given by the usual quantum mechanical expression (assuming the wavefunction is normalized):

$$E_{el} = \langle \Psi | \hat{H}_{el} | \Psi \rangle.$$

(4.6)

where $\Psi = \Psi(r_1, r_2, \ldots, r_N)$ is a Slater determinant, $|\phi_1, \phi_2, \ldots, \phi_N|$. 

We can employ the variational theorem, which states that the energy is always an upper bound to the true energy. Hence, we can obtain better approximate wavefunctions $\Psi$ by varying their parameters until we minimize the energy within the given functional space. Hence, the correct molecular orbitals in the Slater determinant, $\{\phi_i\}_{i=1}^N$, are those which minimize the electronic energy $E_{el}$! The molecular orbitals can be obtained numerically using integration over a grid, or (much more commonly) as a linear combination of a set of given basis functions (so-called "atomic orbital" basis functions, usually atom-centered Gaussian type functions).
Now, using some tricks we don’t have time to get into, we can re-write the Hartree-Fock energy $E_{el}$ in terms of integrals of the one- and two-electron operators:

$$E_{HF} = \sum_i \langle i|h|i \rangle + \frac{1}{2} \sum_{ij} [ii|jj] - [ij|ji], \quad (4.7)$$

where the one electron integral is $(x_i = (r_i, s_i))$ \(^1\)

$$\langle i|j \rangle = \int d\mathbf{x}_1 \phi_i^*(\mathbf{x}_1) h(\mathbf{r}_1) \phi_j(\mathbf{x}_1) \quad (4.8)$$

and a two-electron integral (Chemists’ notation) is

$$[ij|kl] = \int d\mathbf{x}_1 d\mathbf{x}_2 \phi_i^*(\mathbf{x}_1) \phi_j(\mathbf{x}_1) \frac{1}{r_{12}} \phi_k^*(\mathbf{x}_2) \phi_l(\mathbf{x}_2). \quad (4.9)$$

There exist efficient computer algorithms for computing such one- and two-electron integrals.

### 4.6 The Hartree-Fock Equations

Again, the Hartree-Fock method seeks to approximately solve the electronic Schrödinger equation, and it assumes that the wavefunction can be approximated by a single Slater determinant made up of one spin orbital per electron. According to the variational theorem we know that the Slater determinant with the lowest energy is as close as we can get to the true wavefunction for the assumed functional form of a single Slater determinant. The Hartree-Fock method determines the set of spin orbitals which minimize the energy and give us this "best single determinant."

So, we need to minimize the Hartree-Fock energy expression with respect to changes in the orbitals $\phi_i \rightarrow \phi_i + \delta \phi_i$. We have also been assuming that the orbitals $\phi$ are orthonormal, and we want to ensure that our variational procedure leaves them orthonormal. We can accomplish this by Lagrange's method of undetermined multipliers, where we employ a functional $\mathcal{L}$ defined as

$$\mathcal{L}[\{\phi_i\}] = E_{HF}[\{\phi_i\}] - \sum_{ij} \epsilon_{ij}(\langle i|j \rangle - \delta_{ij}) \quad (4.10)$$

where $\epsilon_{ij}$ are the undetermined Lagrange multipliers and $\langle i|j \rangle$ is the overlap between spin orbitals $i$ and $j$, i.e.,

$$\langle i|j \rangle = \int \phi_i^*(\mathbf{x}) \phi_j(\mathbf{x}) d\mathbf{x}. \quad (4.11)$$

\(^1\int d\mathbf{x}_i = \sum_{s \in \{\alpha, \beta\}} \int d\mathbf{r}_s\)
Setting the first variation $\delta \mathcal{L} = 0$, and working through some algebra, we eventually arrive at the Hartree-Fock equations defining the orbitals:

$$h(x_1)\phi_i(x_1) + \sum_{j \neq i} \left[ \int dx_2 |\phi_j(x_2)|^2 r_{12}^{-1} \right] \phi_i(x_1) - \sum_{j \neq i} \left[ \int dx_2 \phi_j^*(x_2)\phi_i(x_2)r_{12}^{-1} \right] \phi_j(x_1) = \epsilon_i \phi_i(x_1),$$

(4.12)

where $\epsilon_i$ is the energy eigenvalue associated with orbital $\phi_i$.

The Hartree-Fock equations can be solved numerically (exact Hartree-Fock), or they can be solved in the space spanned by a set of basis functions (Hartree-Fock-Roothan equations). In either case, note that the solutions depend on the orbitals. Hence, we need to guess some initial orbitals and then refine our guesses iteratively. For this reason, Hartree-Fock is called a self-consistent-field (SCF) approach.

The first term above in square brackets,

$$\sum_{j \neq i} \left[ \int dx_2 |\phi_j(x_2)|^2 r_{12}^{-1} \right] \phi_i(x_1),$$

(4.13)

$(r_{12} = |r_1 - r_2|)$ gives the Coulomb interaction of an electron in spin orbital $\phi_i$ with the average charge distribution of the other electrons. Here we see in what sense Hartree-Fock is a "mean field" theory. This is called the Coulomb term, and it is convenient to define a Coulomb operator as

$$\mathcal{J}_j(x_1) = \int dx_2 |\phi_j(x_2)|^2 r_{12}^{-1},$$

(4.14)

which gives the average local potential at point $x_1$ due to the charge distribution from the electron in orbital $\phi_j$.

The other term in brackets in eq. (20) is harder to explain and does not have a simple classical analog. It arises from the antisymmetry requirement of the wavefunction. It looks much like the Coulomb term, except that it switches or exchanges spin orbitals $\phi_i$ and $\phi_j$. Hence, it is called the exchange term:

$$\sum_{j \neq i} \left[ \int dx_2 \phi_j^*(x_2)\phi_i(x_2)r_{12}^{-1} \right] \phi_j(x_1).$$

(4.15)

We can define an exchange operator in terms of its action on an arbitrary spin orbital $\phi_i$

$$\mathcal{K}_j(x_1) \phi_i(x_1) = \left[ \int dx_2 \phi_j^*(x_2)r_{12}^{-1} \phi_i(x_2) \right] \phi_j(x_1).$$

(4.16)

In terms of these Coulomb and exchange operators, the Hartree-Fock equations become considerably more compact.

$$\left[ h(x_1) + \sum_{j \neq i} \mathcal{J}_j(x_1) - \sum_{j \neq i} \mathcal{K}_j(x_1) \right] \phi_i(x_1) = \epsilon_i \phi_i(x_1).$$

(4.17)
Perhaps now it is more clear that the Hartree-Fock equations are eigenvalue equations. If we realize that

\[ [\mathcal{J}_i(x_1) - K_i(x_1)] \phi_i(x_1) = 0, \]  

(4.19)

then it becomes clear that we can remove the restrictions \( j \neq i \) in the summations, and we can introduce a new operator, the \textbf{Fock operator}, as

\[ f(x_1) = h(x_1) + \sum_j \mathcal{J}_j(x_1) - K_j(x_1). \]  

(4.20)

And now the Hartree-Fock equations are just

\[ f(x_1) \phi_i(x_1) = \epsilon_i \phi_i(x_1). \]  

(4.21)

### 4.6.1 Matrix representation of the Hartree-Fock equation: the Roothaan equation

Introducing a basis set transforms the Hartree-Fock equations into the Roothaan equations. Denoting the atomic orbital basis functions as \( \chi \), we have the expansion

\[ \phi_i = \sum_{\mu=1}^K D_{\mu i} \chi_\mu \]  

(4.22)

for each spin orbital \( i \). This leads to

\[ f(x_1) \sum_\nu D_{\nu i} \chi_\nu(x_1) = \epsilon_i \sum_\nu D_{\nu i} \chi_\nu(x_1). \]  

(4.23)

Left multiplying by \( \chi_\nu(x_1) \) and integrating yields a matrix equation

\[ \sum_\nu D_{\nu i} \int dx_1 \chi^*_\mu(x_1) f(x_1) \chi_\nu(x_1) = \epsilon_i \sum_\nu D_{\nu i} \int dx_1 \chi^*_\mu(x_1) \chi_\nu(x_1). \]  

(4.24)

This can be simplified by introducing the matrix element notation

\[ S_{\mu\nu} = \int dx_1 \chi^*_\nu(x_1) \chi_\mu(x_1), \]  

(4.25)

\[ F_{\mu\nu} = \int dx_1 \chi^*_\nu(x_1) f(x_1) \chi_\mu(x_1). \]  

(4.26)

Now the Hartree-Fock-Roothaan equations can be written in matrix form as

\[ \sum_\nu F_{\mu\nu} D_{\nu i} = \epsilon_i \sum_\nu S_{\mu\nu} D_{\nu i} \]  

(4.27)
or even more simply as matrices

$$\mathbf{F} \mathbf{D} = \mathbf{S} \mathbf{D} \epsilon$$

(4.28)

where $\epsilon$ is a diagonal matrix of the orbital energies $\epsilon_i$. This is like an eigenvalue equation except for the overlap matrix $\mathbf{S}$. One performs a transformation of basis to go to an orthogonal basis to make $S$ vanish. Then it's just a matter of solving an eigenvalue equation (or, equivalently, diagonalizing $F$!). Well, not quite. Since $F$ depends on its own solution (through the orbitals), the process must be done iteratively. This is why the solution of the Hartree-Fock-Roothaan equations are often called the self-consistent-field procedure.

**Computational Aspects: Variational Optimization of Orbitals**

The variational theorem states that for a time-independent Hamiltonian operator, any trial wavefunction will have an energy expectation value that is greater than or equal to the true ground state wavefunction corresponding to the given Hamiltonian.

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0$$

(4.29)

Because of this, the Hartree-Fock energy is an upper bound to the true ground state energy of a given molecule. In the context of the Hartree-Fock method, the best possible solution is at the Hartree-Fock limit, i.e. the limit of the Hartree-Fock energy as the basis set approaches completeness. (The other is the full-CI limit, where the last two approximations of the Hartree-Fock theory as described above are completely undone. It is only when both limits are attained that the exact solution is obtained.)

The starting point for the Hartree-Fock method is a set of approximate one-electron wavefunctions known as orbitals. For an atomic calculation, these are typically the orbitals for a hydrogenic atom (an atom with only one electron, but the appropriate nuclear charge). For a molecular or crystalline calculation, the initial approximate one-electron wavefunctions are typically a linear combination of atomic orbitals (LCAO).

The orbitals above only account for the presence of other electrons in an average manner. In the Hartree-Fock method, the effect of other electrons are accounted for in a mean-field theory context. The orbitals are optimized by requiring them to minimize the energy of the respective Slater determinant. This operation leads to the Hartree-Fock equation described above.
Since the Fock operator depends on the orbitals used to construct the corresponding Fock matrix, the eigenfunctions of the Fock operator are in turn new orbitals which can be used to construct a new Fock operator. In this way, the Hartree-Fock orbitals are optimized iteratively until the change in total electronic energy falls below a predefined threshold. In this way, a set of self-consistent one-electron orbitals are calculated. The Hartree-Fock electronic wavefunction is then the Slater determinant constructed out of these orbitals. Following the basic postulated of quantum mechanics, the Hartree-Fock wavefunction can then be used to compute any desired chemical or physical property within the framework of the Hartree-Fock method and the approximations employed.

Flow diagram of a Hartree-Fock calculation (from wikipedia.)