

IV. The Hartree-Fock Theory

A. The Hartree-Fock Energy and Equations

The variational principle says that the lowest expectation value of the electronic Hamiltonian attained by the ground-state. This Hamiltonian, in the non-relativistic approximation, for N electrons is given by:

$$\hat{H} = \hat{h} + \hat{U} = \sum_{n=1}^N \hat{h}_n + \frac{e^2}{2} \sum_{n \neq m=1}^N \frac{1}{r_{nm}} \quad (4.1.1)$$

Where the “one body” operator is

$$\hat{h} = \sum_{m=1}^N \hat{h}_m = \sum_{m=1}^N -\frac{\hbar^2}{2m_e} \nabla_m^2 + v(\mathbf{r}_m) \quad (4.1.2)$$

And the two-body Coulomb repulsion operator is:

$$\hat{U} = \frac{e^2}{2} \sum_{n \neq m=1}^N \frac{1}{|\mathbf{r}_n - \mathbf{r}_m|} \quad (4.1.3)$$

Given a family of wave functions we can choose the “best” of them by finding that which minimizes the expectation value of \hat{H} . For the Slater wave functions $\Psi_S = |\phi_1 \dots \phi_N|$, the energy to be minimized is:

$$E[\Psi_S] = \langle \Psi_S | \hat{H} | \Psi_S \rangle \equiv E_{HF}[\phi_1 \dots \phi_N] \quad (4.1.4)$$

Because we want the orbitals to be orthonormal, we write a Lagrangian:

$$L[\Psi_S] = L_{HF}[\phi_1 \dots \phi_N] = E_{HF}[\phi_1 \dots \phi_N] - \sum_{i,j=1}^N \mu_{ij} [\langle \phi_i | \phi_j \rangle - \delta_{ij}] \quad (4.1.5)$$

The minimum is attained by:

$$0 = \frac{\delta L_{HF}[\phi_1 \dots \phi_N]}{\delta \phi_i(x)} = \frac{\delta E_{HF}[\phi_1 \dots \phi_N]}{\delta \phi_i(x)} - \sum_{j=1}^N (\mu_{ij} + \mu_{ji}) \phi_j(x) \quad (4.1.6)$$

Let us try a solution with

$$\mu_{ij} = \frac{1}{2} \epsilon_i \delta_{ij} \quad (4.1.7)$$

i.e.:

$$\frac{\delta E_{HF}[\phi_1 \dots \phi_N]}{\delta \phi_i(x)} = \epsilon_i \phi_i(x) \quad (4.1.8)$$

If we can solve this equation, and if the solutions are naturally orthogonal, then we have obtained the necessary conditions for a minimum. Now, we only need to estimate the left hand side of this equation. From the previous work, we know:

$$E_{HF}[\phi_1 \dots \phi_N] = \langle \Psi_S | \hat{h} | \Psi_S \rangle + \langle \Psi_S | \hat{U} | \Psi_S \rangle \quad (4.1.9)$$

where

$$\langle \Psi_S | \hat{h} | \Psi_S \rangle \equiv h[\phi_1 \dots \phi_N] = \sum_{m=1}^N \langle \phi_m | \hat{h}_1 | \phi_m \rangle \quad (4.1.10)$$

Is the "one body energy" defined the as the sum of kinetic energy and "external" energy (i.e. energy due to the frozen nuclei):

$$\hat{h}_1 = -\frac{\hbar^2}{2\mu_e} \nabla^2 + v(\mathbf{r}) \quad (4.1.11)$$

Furthermore, the 2-electron energy can be written as:

$$\begin{aligned} \langle \Psi_S | \hat{U} | \Psi_S \rangle &= \frac{e^2}{2} \sum_{\substack{n,m=1 \\ n \neq m}}^N \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_n \phi_m \right\rangle - \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_m \phi_n \right\rangle \\ &\equiv J[\phi_1 \dots \phi_N] + K[\phi_1 \dots \phi_N] \end{aligned} \quad (4.1.12)$$

Where the direct or Hartree energy is:

$$J[\phi_1 \dots \phi_N] = \frac{e^2}{2} \sum_{n,m=1}^N \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_n \phi_m \right\rangle, \quad (4.1.13)$$

and the energy exchange is

$$K[\phi_1 \dots \phi_N] = -\frac{e^2}{2} \sum_{n,m=1}^N \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_m \phi_n \right\rangle. \quad (4.1.14)$$

The direct energy is numerically equal to the Hartree energy $E_H[n]$ which is a functional of the Slater wave function electron density $n(\mathbf{r}) = \langle \Psi_S | \hat{n}(\mathbf{r}) | \Psi_S \rangle = \sum_{m=1}^{N_e} |\phi_m(\mathbf{r}, s)|^2$:

$$E_H[n] = \frac{e^2}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' \quad (4.1.15)$$

The exchange energy is numerically equal to the “exchange energy” $E_X[\rho]$ which is a functional of the *density matrix* defined by

$$\rho(\mathbf{x}, \mathbf{x}') = \sum_{m=1}^N \phi_m(\mathbf{x}) \phi_m(\mathbf{x}') \quad (4.1.16)$$

As:

$$E_X[\rho] = -\frac{e^2}{2} \iint \frac{|\rho(\mathbf{x}, \mathbf{x}')|^2}{|\mathbf{r} - \mathbf{r}'|} dx dx' \quad (4.1.17)$$

Notice that: $\int \rho(\mathbf{x}, \mathbf{x}) ds = n(\mathbf{r})$. The density matrix is *idempotent*:

$$\int \rho(\mathbf{x}, \mathbf{x}'') \rho(\mathbf{x}'', \mathbf{x}') d^3x'' = \rho(\mathbf{x}, \mathbf{x}') \quad (4.1.18)$$

This result shows that the DM is a projection operator, projecting onto the space of orbitals which defines the Slater wave function.

With the direct and exchange energies we also define their functional derivatives:

$$\begin{aligned}
\frac{\delta J}{\delta \phi_i(\mathbf{x})} &= \frac{1}{2} \sum_{n,m} \frac{\delta}{\delta \phi_i(\mathbf{x})} \left\langle \phi_n \phi_m \left| \frac{e^2}{r_{12}} \right| \phi_n \phi_m \right\rangle \\
&= \frac{e^2}{2} \sum_{n,m} \frac{\delta}{\delta \phi_i(\mathbf{x})} \iint \frac{\phi_n(\mathbf{x}_1)^2 \phi_m(\mathbf{x}_2)^2}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \frac{e^2}{2} \sum_{n,m} \left(2\delta_{in} \phi_n(\mathbf{x}) \int \frac{\phi_m(\mathbf{x}_2)^2}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{x}_2 \right. \\
&\quad \left. + 2\delta_{im} \phi_m(\mathbf{x}) \int \frac{\phi_n(\mathbf{x}_1)^2}{|\mathbf{r}_1 - \mathbf{r}|} d\mathbf{x}_1 \right) \\
&= 2e^2 \left(\int \frac{n(\mathbf{r}_2)}{|\mathbf{r} - \mathbf{r}_2|} d^3r_2 \right) \phi_i(\mathbf{x}) \equiv 2v_H(\mathbf{r})\phi_i(\mathbf{x}) \\
&\equiv 2\hat{v}_H\phi_i(\mathbf{x})
\end{aligned} \tag{4.1.19}$$

Where in the last line we defined the Hartree potential:

$$v_H(\mathbf{r}) = e^2 \sum_{m=1}^N \int \frac{|\phi_m(\mathbf{x})|^2}{|\mathbf{r} - \mathbf{r}'|} d^3r' = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \tag{4.1.20}$$

Then the direct energy can also be written as a functional of the orbitals:

$$J[\phi_1 \dots \phi_N] = \frac{1}{2} \sum_{m=1}^N \langle \phi_m | \hat{v}_H | \phi_m \rangle = \frac{1}{2} \int v_H(\mathbf{r}) n(\mathbf{r}) d^3r = \frac{1}{2} \langle \Psi_S | \hat{V}_H | \Psi_S \rangle \tag{4.1.21}$$

where $\hat{V}_H = \sum_{n=1}^N v_H(\mathbf{r}_n)$ is the total Hartree potential. A similar treatment exists for the exchange energy functional derivative:

$$\begin{aligned}
\frac{\delta K}{\delta \phi_i(\mathbf{x})} &= -\frac{e^2}{2} \sum_{n,m} \frac{\delta}{\delta \phi_i(\mathbf{x})} \left\langle \phi_n \phi_m \left| \frac{1}{r_{12}} \right| \phi_m \phi_n \right\rangle \\
&= -\frac{e^2}{2} \sum_{n,m} \left(2\delta_{in} \phi_m(\mathbf{x}) \int \frac{\phi_m(\mathbf{x}_2) \phi_n(\mathbf{x}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{x}_2 \right. \\
&\quad \left. + 2\delta_{im} \phi_n(\mathbf{x}) \int \frac{\phi_n(\mathbf{x}_2) \phi_m(\mathbf{x}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{x}_2 \right) = \\
&= -2e^2 \sum_n \int \frac{\phi_n(\mathbf{x}_2) \phi_i(\mathbf{x}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{r}_2 \phi_n(\mathbf{x}) = 2\hat{K}\phi_i(\mathbf{x})
\end{aligned} \tag{4.1.22}$$

Where the last equality is based on a definition of a one-particle exchange operator:

$$\begin{aligned}\hat{K}_1\psi(\mathbf{x}) &\equiv -e^2 \sum_{m=1}^N \phi_m(\mathbf{x}) \int \frac{\phi_m(\mathbf{x}')\psi(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}' \\ &= -e^2 \int \frac{\rho(\mathbf{x},\mathbf{x}')\psi(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}'\end{aligned}\quad (4.1.23)$$

Then the exchange energy is written as a functional of the orbitals:

$$K[\phi_1 \dots \phi_N] = \frac{1}{2} \sum_{n=1}^N \langle \phi_n | \hat{K}_1 | \phi_n \rangle = \frac{1}{2} \langle \Psi_S | \hat{Q}_H | \Psi_S \rangle \quad (4.1.24)$$

where $\hat{Q}_H = \sum_{n=1}^N \hat{K}_n$ is the total exchange operator. The other functional derivatives needed are:

$$\frac{\delta}{\delta\phi_i(\mathbf{x})} \int \phi_n(\mathbf{x}') \nabla^2 \phi_n(\mathbf{x}') d^3x' = 2\nabla^2 \phi_n(\mathbf{x}) \delta_{in} \quad (4.1.25)$$

and

$$\frac{\delta}{\delta\phi_i(\mathbf{x})} \int \phi_n(\mathbf{x}')^2 v(\mathbf{r}') d\mathbf{x}' = 2v(\mathbf{r}) \phi_n(\mathbf{x}) \delta_{in} \quad (4.1.26)$$

Thus:

$$\frac{\delta}{\delta\phi_i(\mathbf{x})} \int \phi_n(\mathbf{x}') \hat{h}(\mathbf{r}') \phi_n(\mathbf{x}') d\mathbf{x}' = 2\hat{h}\phi_n(\mathbf{x}) \delta_{in} \quad (4.1.27)$$

Plugging all these terms into Eq. (4.1.7), we obtain the **Hartree-Fock equations**:

$$\hat{F}\phi_i(\mathbf{x}) = \epsilon_i \phi_i(\mathbf{x}) \quad (4.1.28)$$

where:

$$\hat{F} = \hat{h} + v_H(\mathbf{r}) + \hat{K}_1 \quad (4.1.29)$$

Equations (4.1.28) seem very much like 1-electron eigenvalue equations of the Schrödinger equation, except that instead of a regular Hamiltonian, we have a Fock operator including the non-local exchange. We showed in the exercise that \hat{K} is Hermitean and thus so is \hat{F} , i.e. $\langle \phi | \hat{F} | \psi \rangle = \langle \psi | \hat{F} | \phi \rangle^*$. We can thus choose the orbital solutions of the eigenvalue equation (4.1.28) orthonormal. This shows that the choice Eq. (4.1.7) is indeed acceptable.

Because $v_H(\mathbf{r})$ and \hat{K} themselves depend on ϕ_m , the Hartree-Fock equations are fundamentally different from the Schrödinger Equation: they are *nonlinear equations*.

Now that the sum of orbital energies is:

$$\begin{aligned} \sum_{m=1}^N \epsilon_m &= \sum_{m=1}^N \langle \phi_m | \hat{F} | \phi_m \rangle = \sum_{m=1}^N [\langle \phi_m | \hat{h}_1 | \phi_m \rangle + \langle \phi_m | \hat{v}_H + \hat{K}_1 | \phi_m \rangle] \\ &= h[\phi_1, \dots, \phi_N] + 2J[\phi_1, \dots, \phi_N] + 2K[\phi_1, \dots, \phi_N] \end{aligned} \quad (4.1.30)$$

This shows that the orbital sum is not equal to the energy of the wave function, since it involves double counting of the direct and exchange energies. The HF energy is thus:

$$E_{HF}[\phi_1, \dots, \phi_N] = \sum_{m=1}^N \epsilon_m - (J[\phi_1, \dots, \phi_N] + K[\phi_1, \dots, \phi_N]) \quad (4.1.31)$$

B. Restricted closed-shell Hartree-Fock

For molecules with even number N of electrons in a spin-singlet state, we can impose the following structure on the Slater wave function. We can assume that the $2N$ spin-orbitals come in pairs: $\phi_{2j-1}(\mathbf{x}) = \psi_j(\mathbf{r})\alpha(\omega)$ and $\phi_{2j}(\mathbf{x}) = \psi_j(\mathbf{r})\beta(\omega)$. Thus each orbital ψ_j , $j = 1, \dots, N/2$ is “doubly occupied” by electrons of both spins. By imposing this constraint we obtain the “restricted” Hartree-Fock ground state. It will sometime be of higher energy than the fully

unrestricted case. However, the wave function has a well-defined spin which may be advantageous in some applications.

The Restricted Hartree-Fock (RHF) energy in the closed shell case remain essentially the same except for counting business. We can formulate all expressions using only the spatial orbitals. Indeed, the RHF energy is given by:

$$E_{RHF}[\psi_1, \dots, \psi_{N/2}] = 2 \sum_{m=1}^{N/2} \langle \psi_m | \hat{h} | \psi_m \rangle + 4J[\psi_1, \dots, \psi_{N/2}] + 2K[\psi_1, \dots, \psi_{N/2}] \quad (4.2.1)$$

Where $J[\psi_1, \dots, \psi_{N/2}]$ and $K[\psi_1, \dots, \psi_{N/2}]$ are the orbital functionals defined in Eqs. (4.1.13) and (4.1.14) respectively. The reason for multiplying the one body part by two is evident: each orbitals is double occupied so has double contribution. The direct part is multiplied by 4 since the density is multiplied by two and the direct part depends on the density multiplied by itself. Finally, the exchange part is multiplied by 2 and not 4 since only $\alpha - \alpha$ and $\beta - \beta$ contribute, while $\alpha - \beta$ and $\beta - \alpha$ do not (so only half the contribution of direct).

The RHF equations are then:

$$\left(-\frac{\hbar^2}{2\mu_e} \nabla^2 + v(\mathbf{r}) + v_H(\mathbf{r}) \right) \psi_m(\mathbf{r}) + \hat{K}_1 \psi_m(\mathbf{r}) = \epsilon_m \psi_m(\mathbf{r}) \quad (4.2.2)$$

Where:

$$v_H(\mathbf{r}) = 2e^2 \sum_{m=1}^{N/2} \int \frac{|\psi_m(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (RHF) \quad (4.2.3)$$

and

$$\hat{K}_1\psi(\mathbf{r}) \equiv -e^2 \sum_{m=1}^{N/2} \psi_m(\mathbf{r}) \int \frac{\psi_m(\mathbf{r}')\psi(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \quad (RHF) \quad (4.2.4)$$

Note that the Hartree interaction is between each electron and all other electrons regardless of their spin while the exchange interaction involves each electrons with all other electrons *of the same spin*.

Example: The H₂ molecule

We apply the RHF theory for the H₂, having a pair of electrons. The 2-electron wave function includes just one spatial orbital populated by spin-paired electrons:

$$\Phi(\mathbf{x}_1, \mathbf{x}_2) = \begin{vmatrix} \psi(\mathbf{r}_1) & \psi(\mathbf{r}_2) \\ \bar{\psi}(\mathbf{r}_1) & \bar{\psi}(\mathbf{r}_2) \end{vmatrix} = \psi(\mathbf{r}_1)\psi(\mathbf{r}_2)[\alpha(1)\beta(2) - \alpha(2)\beta(1)] \quad (4.2.5)$$

Since there is only one orbital the exchange K and direct J involve just one and the same integral. Thus the RHF energy is in this case: $E_{RHF}[\psi] = 2\langle\psi_m|\hat{h}|\psi_m\rangle + 2J[\psi]$ and the RHF equation is

$$\left(-\frac{\hbar^2}{2\mu_e}\nabla^2 + v(\mathbf{r}) + \frac{1}{2}v_H(\mathbf{r}) \right) \psi(\mathbf{r}) = \epsilon\psi(\mathbf{r}) \quad (4.2.6)$$

where $v(\mathbf{r}) = -\frac{e^2}{|\mathbf{r}-\mathbf{R}_A|} - \frac{e^2}{|\mathbf{r}-\mathbf{R}_B|}$ and $v_H(\mathbf{r}) = 2e^2 \int \frac{|\psi(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d^3r'$. The effect of exchange here is to annihilate the Coulomb repulsion of the α (β) electron with itself, leaving only the interaction of the $\alpha - \beta$ electrons.

This RHF approach works nicely for the case that the distance between the nuclei $|\mathbf{R}_A - \mathbf{R}_B|$ is close to the typical bond length of H₂ (which is close to $1.4a_0$). The energy of the molecule at this configuration can be calculated numerically and results in $E_{RHF} = -1.134E_h$. Compared to the energy of 2 H atoms ($-1E_h$) this results indicates that the atomization energy of H₂ is $0.134E_h = 3.65\text{eV}$. The atomization energy based on experimental results is

about 4.75eV. This shows that the RHF approximation does not give high quality atomization energies, since a deficit of 1 eV is very substantial in Chemical energy terms.

A more severe problem arises when we place the two nuclei far from each other. We expect the resulting energy and wave function to resemble that of two separated H atoms. I.e the exact wave-function should approach something like $\Psi = |\psi_{1s}(\mathbf{r}_1 - \mathbf{R}_A)\bar{\psi}_{1s}(\mathbf{r}_2 - \mathbf{R}_B)| - |\psi_{1s}(\mathbf{r}_2 - \mathbf{R}_A)\bar{\psi}_{1s}(\mathbf{r}_1 - \mathbf{R}_B)|$. This form however is not supported by the RHF ansatz of Eq. (4.2.5). Indeed, if we think of the solution of the RHF equation as being approximately given by $\psi \approx \psi_A(\mathbf{r}) + \psi_B(\mathbf{r})$, where $\psi_X(\mathbf{r}) = \psi(\mathbf{r} - \mathbf{R}_X)$, $X = A, B$, then the RHF wave function is:

$$\begin{aligned} \Phi(\mathbf{x}_1, \mathbf{x}_2) &\approx \left| (\psi_A(\mathbf{r}_1) + \psi_B(\mathbf{r}_1)) (\bar{\psi}_A(\mathbf{r}_2) + \bar{\psi}_B(\mathbf{r}_2)) \right| \\ &= \left[|\psi_A(\mathbf{r}_1)\bar{\psi}_A(\mathbf{r}_2)| + |\psi_B(\mathbf{r}_1)\bar{\psi}_B(\mathbf{r}_2)| \right] \\ &\quad + \left[|\psi_A(\mathbf{r}_1)\bar{\psi}_B(\mathbf{r}_2)| + |\psi_B(\mathbf{r}_1)\bar{\psi}_A(\mathbf{r}_2)| \right] \end{aligned} \quad (4.2.7)$$

The first term is an ionic term, where both electrons are on the same atom (either A or B) while the second term places one electron on each atom – “neutral” term. The problem of RHF theory is the ionic term. It may be important when the atoms are close but it should go to zero when they are far.

C. Atomic Orbitals and Gaussian Basis sets

Where do we get "good" basis functions? What is "good"?

We want a small basis that can still describe the electrons. On natural source are the atomic orbitals of the atoms. These are of the form resembling exponentials time polynomials. Thus, one choice is:

$$\chi_{lm}^A(\mathbf{r}) = \xi_{lm}^A(\mathbf{r} - \mathbf{R}_A) \quad (4.2.8)$$

Where:

$$\xi_{lm}^{\zeta}(\mathbf{r}) = r^l e^{-\zeta r} Y_{lm}(\theta, \phi) \quad (4.2.9)$$

Where Y_{lm} are the spherical harmonics. One can also take appropriate combinations of these functions to make them all real. These functions have a desired analytical property: their l 's derivative exhibits a cusp of the correct order and structure at $\mathbf{r} = \mathbf{R}_A$. There exist analytical formula for doing the overlap and "one-body" integrals. But there are no convenient formulae for the 2-body integrals, although some progress was made in recent years (see articles by Handy).

A more convenient, although less natural choice (no cusp). Is the use of Gaussian functions, for example:

$$\xi_{lm}^{\zeta}(\mathbf{r}) = r^l \left(\sum_{\mu} a_{\mu,l} e^{-\beta_{\mu} r^2} \right) Y_{lm}(\theta, \phi) \quad (4.2.10)$$

Where $a_{l,\mu}$ are called "contraction coefficients. These are chosen so that $(\sum_{\mu} a_{\mu,l} e^{-\beta_{\mu} r^2})$ resembles $e^{-\zeta r}$. With Gaussian functions very effective and rapid algorithms were published allowing extremely fast 2-electron integrals.

D. Variational-Algebraic approach Hartree-Fock

We have seen that the Hartree-Fock equations can be derived by searching for that the most general Slater wave function that minimizes the Hartree-Fock functional. However implementing a solution to such equations is usually very difficult, if not impossible in practice. A more practical approach, that keeps the spirit of the Hartree_Fock approach was developed by Roothan and Hall. In this approach we find the optimal Slater wave function of orbitals constrained to lie in a finite dimensional vector space spanned by *basis*

functions, usually called atomic orbitals (although, up to a point, we need not assume this) $\chi_\sigma(\mathbf{x})$, $\sigma = 1, \dots, M$. Thus, a set of N molecular orbitals $\phi_n(\mathbf{x})$ ($n = 1, \dots, N$) in a determinant $\Psi_S[C]$ of this form must all be of the following form:

$$\phi_n(\mathbf{x}) = \sum_{\sigma=1}^M \chi_\sigma(\mathbf{x}) C_{\sigma n} \quad (4.3.1)$$

The $C_{\sigma n}$ coefficients form an $M \times N$ matrix, called the MO coefficient matrix for the determinant. Note that for this to make sense we must demand $M > N$. The HF energy functional now becomes a function of these coefficients $C_{\sigma n}$. The constraint that the MO's are orthonormal, $\langle \phi_n | \phi_m \rangle = \delta_{nm}$, becomes:

$$\delta_{nm} = \left\langle \sum_{\sigma=1}^M \chi_\sigma C_{\sigma n} \left| \sum_{\sigma'=1}^M \chi_{\sigma'} C_{\sigma' m} \right. \right\rangle = (C^T S C)_{nm} \quad (4.3.2)$$

Where we use matrix algebra notation and the $M \times M$ matrix S is defined by:

$$S_{\sigma\sigma'} = \langle \chi_\sigma | \chi_{\sigma'} \rangle \quad (4.3.3)$$

Thus, the orthonormality condition is:

$$C^T S C = I_N \quad (4.3.4)$$

Where I_N is the $N \times N$ unit matrix. Let us now derive an expression for the expectation value of a one-body operator in a Slater wave function of these MO's, by Eq. **Error! Reference source not found.**:

$$\langle \Psi_S | \hat{O} | \Psi_S \rangle = \sum_{n=1}^N \langle \phi_n | \hat{O}_1 | \phi_n \rangle = \text{Tr}[C^T O C] \quad (4.3.5)$$

Where O is the $M \times M$ matrix in the AO basis:

$$O_{\sigma\sigma'} = \langle \chi_\sigma | \hat{O}_1 | \chi_{\sigma'} \rangle \quad (4.3.6)$$

It is customary to define the $M \times M$ density matrix:

$$P = CC^T \quad (4.3.7)$$

And with it we can write:

$$\langle \Psi_S | \hat{O} | \Psi_S \rangle = Tr[C^T O C] = Tr[PO] \quad (4.3.8)$$

In the last step we used the fact that the trace of the product of two matrices is invariant to their order of multiplication:

$$Tr[AB] = \sum_{\sigma=1}^M (AB)_{\sigma\sigma} = \sum_{\sigma=1}^M \sum_{n=1}^N A_{\sigma n} B_{n\sigma} = \sum_{n=1}^N \sum_{\sigma=1}^M B_{n\sigma} A_{\sigma n} = Tr[BA] \quad (4.3.9)$$

Notice that the DM has the generalized idempotency property:

$$PSP = CC^T S C C^T = CC^T = P \quad (4.3.10)$$

One can see that P is a symmetric matrix. Furthermore, one can see that it is positive semi-definite, i.e. for any vector: $v^T P v = v^T C C^T v = (C^T v)^T (C^T v) \geq 0$.

Furthermore:

$$Tr[PS] = Tr[CC^T S] = Tr[C^T S C] = N$$

This last step is a result of Eq. (4.3.4). Finally, the 2-body operator, by Eq. (4.1.12), we need the direct and exchange. We use:

$$[kl|mn] = \sum_{\sigma, \sigma', \tau, \tau'=1}^N C_{k\sigma}^T C_{\sigma'l} [\sigma\sigma' | \tau\tau'] C_{m\tau}^T C_{\tau'n} \quad (4.3.11)$$

Then:

$$\begin{aligned} J[C] &= \frac{e^2}{2} C_{n\sigma}^T C_{\sigma'n} [\sigma\sigma' | \tau\tau'] C_{m\tau}^T C_{\tau'm} \\ K[C] &= -\frac{e^2}{2} C_{n\sigma}^T C_{\sigma'm} [\sigma\sigma' | \tau\tau'] C_{m\tau}^T C_{\tau'n} \end{aligned} \quad (4.3.12)$$

Where we use the convention that all repeated indices are summed over: latin indices are summed between 1 and N and greek indices between 1 and M . The summations on n, m can be done first and we obtain:

$$\begin{aligned} J[C] &= \frac{e^2}{2} P_{\sigma'\sigma} [\sigma\sigma' | \tau\tau'] P_{\tau\tau'} \\ K[C] &= -\frac{e^2}{2} P_{\sigma\tau'} [\sigma\sigma' | \tau\tau'] P_{\sigma'\tau} \end{aligned} \quad (4.3.13)$$

The last expression can be reindexed (assuming P is a symmetric matrix) as:

$$K[C] = -\frac{e^2}{2} P_{\sigma'\sigma} [\sigma\tau' | \tau\sigma'] P_{\tau\tau'} \quad (4.3.14)$$

Thus:

$$J[C] + K[C] = \frac{e^2}{2} P_{\sigma'\sigma} ([\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \tau\sigma']) P_{\tau\tau'} \quad (4.3.15)$$

Finally, defining:

$$V_{IJ} \equiv V_{\sigma\sigma', \tau\tau'} = [\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \tau\sigma'] = [\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \sigma'\tau] \quad (4.3.16)$$

Using the double indexing $I \equiv (\sigma\sigma')$ and $J = (\sigma'\sigma''')$. Note that $V_{IJ} = V_{JI}$ as can be seen from:

$$\begin{aligned} V_{IJ} = V_{\sigma\sigma', \tau\tau'} &= [\sigma\sigma' | \tau\tau'] - [\sigma\tau' | \sigma'\tau] = [\tau\tau' | \sigma\sigma'] - [\tau\sigma' | \tau'\sigma] = V_{\tau\tau' \sigma\sigma'} \\ &= V_{JI} \end{aligned} \quad (4.3.17)$$

We may thus write:

$$J[C] + K[C] = \frac{e^2}{2} P^t V P \quad (4.3.18)$$

Where now, we consider P not as a $M \times M$ matrix, but as a column vector of M^2 elements. P^t is the corresponding row-vector. Similarly, V is not a $M \times M \times M \times M$ tensor but as $M^2 \times M^2$ matrix.

The HF energy is thus compactly written as:

$$E_{HF}[P] = P^t h + \frac{1}{2} P^t V P \quad (4.3.19)$$

Note, that our unknown variable is now the DM P . We want to minimize E_{HF} with respect to P , however, we need to impose 2 types of constraints. First, we need to specify the subject to the constraints:

$$\begin{aligned} Tr[PS] &= N \\ G &= P S P S - P S = 0 \end{aligned} \quad (4.3.20)$$

In order to minimize the energy we introduce the Lagrangian:

$$L[P] = E_{HF}[P] - \mu(P^t S - N) - \Lambda^t G \quad (4.3.21)$$

The number μ and the $M \times M$ matrix Λ are Lagrange multipliers. The algebraic Hartree-Fock equations are now $\frac{\delta L}{\delta P_{\sigma\tau}} = 0$. In order to obtain working expressions we derive:

$$\frac{\delta E_{HF}}{\delta P_I} = h_I + \frac{1}{2} (P_J V_{JI} + V_{IJ} P_J) = h_I + V_{IJ} P_J \quad (4.3.22)$$

Where again, we use the convention that when a super index appears twice we sum over it. This can be written more compactly as:

$$\frac{\delta E_{HF}}{\delta P} = h + V P \equiv F \quad (4.3.23)$$

This gradient is what we call the ‘‘Fock matrix’’ F . In our present notation F is a M^2 vector F_I with index I . But soon we will consider it as a matrix with two indices $F_{\sigma\sigma'}$. The constraints can easily be derived in a similar way, leading to the following Lagrangian gradient:

$$\frac{\delta L}{\delta P} = F - \mu S - (S P S \Lambda + \Lambda P S - S \Lambda) \quad (4.3.24)$$

It is easy to convince one’s self that F is always a symmetric matrix, for any P .

The condition for minimum is:

$$F - \mu S - (SPS\Lambda + S\Lambda PS - S\Lambda) = 0 \quad (4.3.25)$$

Multiplying by SP from the left we find:

$$SPF - \mu SPS - SPS\Lambda PS = 0 \quad (4.3.26)$$

Multiplying by PS from the right we find:

$$FPS - \mu SPS - SPS\Lambda PS = 0 \quad (4.3.27)$$

Subtracting, we obtain:

$$SPF - FPS = 0 \quad (4.3.28)$$

The set of equations that need to be solved simultaneously is:

$$\begin{aligned} SPF - FPS = 0 & & F = h + VP \\ PSP = P & & Tr[PS] = N \end{aligned} \quad (4.3.29)$$

One practical way of doing this is to go back to the matrix C . In terms of these, the equations become:

$$SCC^T F - FCC^T S = 0 \quad F = h + VCC^T \quad C^T SC = I_N \quad (4.3.30)$$

These equations can all be met if we demand that:

$$FC = SCE \quad F = h + VCC^T \quad (4.3.31)$$

Where E is a $M \times M$ diagonal matrix. Indeed, from this equation we also have, from the symmetry of F and S : $C^T F = EC^T S$. Left-multiplying by SC we find $SCC^T F = SCEC^T S$ and using the first equation in (4.3.31) on the right hand side we obtain the first equation in (4.3.30). Furthermore, multiplying the first equation in (4.3.31) by C^T we find: $C^T FC = C^T SCE$. On the left we replace $C^T F$ by $EC^T S$ and obtain:

$$[E, C^T SC] = 0$$

We find that C^TSC is commutative with a diagonal matrix. If no two elements on the diagonal of E are equal then C^TSC is diagonal. We know that the diagonal entries must be positive since S is positive definite. Furthermore, we can choose the norm of the columns of C so that all diagonal elements of C^TSC are equal to 1. In this case then $C^TSC = I$. When there are several elements on the diagonal of E which are exactly equal, then one can take linear combinations of the corresponding columns of the C -matrix, without disturbing their eigenstatishness. Once can then always create a situation which again allows $C^TSC = I$. We thus find that the procedure of finding the generalized eigenstates and eigenvalues of F is indeed a procedure for finding the minimum.

Thus Eq. (4.3.31) is the algebraic Hartree Fock equation. In actual calculations, it is very common that programs solve self consistently the algebraic HF equation. This procedure is appropriate for small to medium sized systems. But for larger system it may be beneficial to directly the minimize of the Lagrangian, using the gradient in Eq. (4.3.24). Of course, iterations are still needed because a search must be made for the Lagrange multiplier Λ .

E. The Algebraic Density Matrix and Charge Analysis

We have seen that the density matrix is defined by the relation $P = CC^T$ where C_{an} is the coefficient of the AO $\chi_\alpha(\mathbf{x})$ in the expansion of the MO $\phi_n(\mathbf{x})$. The relation of P the the real space density matrix defined in XXX is (we use the convention that repeated roman indices are summed from 1 to N and repeated greek indices are summed from 1 to M):

$$\rho(\mathbf{x}, \mathbf{x}') = \phi_n(\mathbf{x})\phi_n(\mathbf{x}') = C_{an}\chi_\alpha(\mathbf{x})\chi_\beta(\mathbf{x}')C_{\beta n} = P_{\alpha\beta}\chi_\alpha(\mathbf{x})\chi_\beta(\mathbf{x}')$$

Thus we see that $P_{\alpha\beta}$ determines ρ . Hence the common name for these two quantities. Note also that $P_{\alpha\beta}$ determines the density $n(\mathbf{r})$ since that is easily obtained by place $\mathbf{x} = \mathbf{x}'$ in the DM and integrating over spin:

$$n(\mathbf{r}) = \lim_{\eta \rightarrow 0^+} P_{\alpha\beta} S_{\alpha\beta} \frac{\chi_\alpha(\mathbf{r})\chi_\beta(\mathbf{r})}{S_{\alpha\beta} + i\eta} = q_\alpha n_\alpha(\mathbf{r}) + \sum_{\alpha \neq \beta} q_{\alpha\beta} n_{\alpha\beta}(\mathbf{r})$$

Where:

$$n_\alpha(\mathbf{r}) = \frac{\chi_\alpha(\mathbf{r})^2}{S_{\alpha\alpha}}$$

Are the “atomic electron number densities” (each integrates to 1) and

$$n_{\alpha\beta}(\mathbf{r}) = \frac{\chi_\alpha(\mathbf{r})\chi_\beta(\mathbf{r})}{S_{\alpha\beta} + i\eta}$$

Is the bond electron number density (again, integrating to 1, or zero). The atomic charges are then $q_\alpha = P_{\alpha\alpha} S_{\alpha\alpha}$ and the bond charges are $q_{\alpha\beta} = P_{\alpha\beta} S_{\alpha\beta}$.

This form of charge analysis is very popular and allows to obtain “intuitive” pictures for the charge distribution in the molecule. While useful to many, the user should be warned that this analysis is “basis-form” dependent. What we mean by this is that if we take different linear combinations of the *same* set of basis functions (i.e. we stay in the same Hilbert space), our charge analysis will yield totally different results. This is because when we take linear combinations: $\chi' = \mathbf{T}\chi$ then the density matrix changes by $P' = \mathbf{T}^T P \mathbf{T}$. Thus in general the charges on each atom can change by this procedure.

Note that when one integrates over \mathbf{r} , on the left hand one gets N . On the right hand the first gives (assuming the basis functions are normalized $P_{\alpha\alpha}$

F. Solving the Hartree-Fock Equations

A plausible “algorithm” for solving the Hartree-Fock equations is as follows:

Guess $\phi_m(\mathbf{r})$, $m = 1, \dots, N_e$.

Build $v_H(\mathbf{r})$ (Eq.(4.1.20)), \hat{K} (Eq.(4.1.23)) thus determining \hat{F} (eq. (4.1.29)).

Solve the eigenvalue equations (Eq. (4.1.28)) to get a new set of orbitals corresponding to the lowest energy orbitals.

Redo from step 2 using the new orbitals, until you converge – i.e. until the orbitals change no more.

While this algorithm seems reasonable, in practice it rarely converges. There are several ways to make an algorithm “practical”.

i. Direct inversion in iterative space (DIIS)

This method, devised by Pulay (P. Pulay, Chem. Phys. Lett. 73, 393 (1980)) is designed to speed up the convergence. Suppose the iterative process has produced M iterants \vec{v}_m $m = 1, \dots, M$ (Fockians, density matrices or sets of N_e orbitals). We can define residuals by:

$$\delta \vec{v}_m = \vec{v}_m - \vec{v}_{m-1} \quad (4.4.1)$$

We want to produce a new iterant by interpolation:

$$\vec{v} = \sum_{m=1}^M w_m \vec{v}_m \quad (4.4.2)$$

where w_m are the weights and they sum to unity:

$$\sum_{m=1}^M w_m = 1 \quad (4.4.3)$$

These weights are obtained by minimizing the residual, assuming linearity:

$$\delta \vec{v} = \sum_{m=1}^M w_m \delta \vec{v}_m \quad (4.4.4)$$

The function to be minimized is:

$$\begin{aligned}
J[w] &= \frac{1}{2} \delta \vec{v}^T \delta \vec{v} - \lambda \left[\sum_{m=1}^M w_m - 1 \right] \\
&= \frac{1}{2} \sum_{m,m'=1}^M w_m B_{m,m'} w_{m'} - \lambda \left[\sum_{m=1}^M w_m - 1 \right]
\end{aligned} \tag{4.4.5}$$

Where:

$$B_{m,m'} = \delta \vec{v}_m^T \delta \vec{v}_{m'} \tag{4.4.6}$$

Differentiating with respect to w_k gives:

$$0 = \frac{\partial J}{\partial w_k} = B_{km} w_m - \lambda w_k \tag{4.4.7}$$

The solution of these equations, together with the constraints Eq. (4.4.3) gives:

$$\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1M} & -1 \\ B_{21} & B_{22} & \cdots & B_{2M} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ B_{M1} & B_{M2} & \cdots & B_{MM} & -1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{4.4.8}$$

The solution of this equation gives the desired weights. The use of this algorithm can be done in the following way:

1. Get new \vec{v} as output from the iterative procedure. Add it to the list i.e. designate it as \vec{v}_M
2. Find weights from which get interpolant $\vec{v}' = \sum_{m=1}^M w_m \vec{v}_m$.
3. Use \vec{v}' as input to the iterative procedure and redo from 1.

ii. Direct Minimization

Sometimes the DIIS procedure is not effective and other methods are tried. One of the most useful methods is to use numerical minimization techniques, such as the conjugate gradients algorithm to directly minimize the energy of the Slater wave function under the constraints. This methods is especially

useful when the molecule being studied is very large. Special tricks are used to formulate the minimization problem is an “unconstraint minimization” (see for example, Nunes et al, Phys. Rev. B 17611, 50 (1994)).

G. Performance of the Hartree-Fock approximation

We examine the performance of Hartree-Fock approximation on, for example Formaldehyde. There are 2 sources of error. One is in the application, since we use finite basis sets. Then there is the intrinsic error.

In the table below, we see the prediction of various properties of formaldehyde, calculated with increasing quality of basis set and compared to experiment.

Basis	R(CO)Å	R(CH)Å	A(O-C-H)	Energy(au)
sto-3g	1.2169	1.1014	122.73	-112.35435
3-21g	1.2071	1.0833	122.51	-113.22182
sto-6g	1.2163	1.0981	122.61	-113.44078
6-31g	1.2103	1.0816	121.69	-113.80837
D95	1.2170	1.0843	121.57	-113.83071
D95v*	1.1887	1.0935	121.96	-113.89173
6-311g**	1.1787	1.0949	122.09	-113.89915
6-311++g**	1.1797	1.0943	121.97	-113.90287
apvtz	1.1786	1.0927	121.94	-113.91534
experimental	1.210	1.1020	121.1	

We see Hartree-Fock converges when basis set quality increases. However the converged quantity deviates somewhat from experimental values.

This deviance exists because Hartree-Fock theory is only an approximation. What it assumes is that the electrons act as if they are independent particles (since it imposes a single determinant). The real ground-state is composed from a huge series of determinants. The "independent" particles interact with the mean field of all other particles, while in essence each electrons has to

interact with each other electron, trying as much as possible to avoid it, without paying too much in kinetic or electron-nuclear potential energy.

H. Beyond Hartree-Fock

The Hartree-Fock method is very successful, since it typically accounts for over 99% of the electronic energy of molecules. Yet it is not accurate enough for most applications in chemistry. The reason is that most quantities of chemical significance are *energy differences* - not absolute energies. When differences are considered the errors in the Hartree-Fock approach are not small.

One way to improve the situation is to approximate the groundstate wave function by a series of determinants:

$$\Psi_{gs} = \sum_{[n_1, \dots, n_{N_e}]} c_{[n_1, \dots, n_{N_e}]} |\phi_{n_1} \dots \phi_{n_{N_e}}| \quad (4.6.1)$$

where ϕ_n are an infinite orthonormal set of orbitals. Such an expansion can always be made, with any such set. We can thus take the orbitals produced by the Hartree-Fock process. This has the added nicety that the first determinant is already a good approximation to the ground-state.

In this case we can classify the determinants in the following way. We divide the orbitals into two sets: one is the set of N_e HF orbitals, called the occupied orbitals, the rest of orbitals named *virtuals*. We then classify the determinants by the number of *occupieds* missing. Thus we speak of all single substitutions, double substitutions etc.

A commonly employed method is the configuration-interaction (CI) method. In a CI one takes a number of determinants D_α $\alpha = 1, \dots, M$ and uses them to minimize the energy:

$$E[\vec{c}] = \sum_{m,n=1}^M c_m \langle D_m | \hat{H} | D_n \rangle c_n \quad (4.6.2)$$

under the constraint of normalization: $\sum_{m,n=1}^M c_m \langle D_m | D_n \rangle c_n = 1$. One common way of choosing the determinants that go into this expansion is by collecting all single, double, triple etc excitations. A determinant is singly excited if when compared to the HF determinant it has one occupied molecular orbital replaced by some virtual orbital. Virtual orbitals are excited eigenfunctions of the Fock operator. One can show that the singles alone do not allow a reduction of energy. However, singles and doubles give sometimes good results. Such a method is called singles-doubles CI (SDCI). One problem with this theory is that it is not "size consistent". For example, calculating the energy of 2 distant Helium atoms will not give the twice the energy of one Helium atom under the same order of theory.

Another approach is many-body perturbation theory, called Moller-Plesser theory. In this approach, one writes the many-body Hamiltonian as:

$$\hat{H} = \sum_{n=1}^M \hat{F}_n + \hat{W} \quad (4.6.3)$$

Where $\hat{W} = \hat{H} - \sum_{n=1}^N \hat{F}_n$ is considered a "small" perturbation. This quantity is not small enough and high order MP theory does not converge. However, second order MP theory, called MP2, is sometimes a useful approach. It is size consistent. However, it relies heavily on the quality of the Hartree-Fock solution: Hartree-Fock must be a good reference in which to base a perturbation theory.

All wave-function methods beyond Hartree-Fock theory become quickly very expensive as system size grows. In fact, the numerical cost of good methods

typically scales as $O(N^7)$ for the coupled cluster method, which is a size-consistent non-variational method (variational methods are derived from the variational theorem), not discussed above. Thus, every enlargement of number of electrons by a factor of two makes the calculation a factor 100 more expensive!