

Hartree–Fock theory

5.1 The Hartree–Fock wave function

The main assumption of Hartree–Fock theory is that the wave functions for N electrons may be approximated with a single Slater determinant built from a set of “optimal” spin orbitals $\{\psi_1, \psi_2, \dots, \psi_N\}$:

$$|\Phi_0\rangle = |\psi_1, \psi_2, \dots, \psi_N\rangle. \quad (5.1)$$

The goal is to find the best set of spin orbitals in the variational sense. Let us define the functional $E[\Phi]$ of a trial wave function Φ :

$$E[\Phi] = \langle \Phi | \hat{H} | \Phi \rangle, \quad (5.2)$$

then the Hartree–Fock energy is given by:

$$E_0 = \min_{\Phi} E[\Phi] \quad \text{enforcing} \quad \langle \Phi | \Phi \rangle = 1, \quad (5.3)$$

where we assume that the determinant Φ is normalized.

In a spin orbital basis the Hartree–Fock energy is:

$$E_0[\{\psi_i\}] = \sum_i^{\text{occ}} \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle ij | ij \rangle. \quad (5.4)$$

5.1.1 Minimization of the Hartree–Fock energy functional

As we have seen in the previous section, if the spin orbitals are orthonormal then Φ is guaranteed to be normalized. How can we minimize the Hartree–Fock energy functional and at the same time impose the condition $\langle \psi_i | \psi_j \rangle = \delta_{ij}$? We will use the technique of Lagrange multipliers.

■ Example 5.1.1 | Minimization of a function of two variables with a constraint.

Let us consider the problem of minimizing the function $f(x, y) = x^2 + 2y^2$ with the constraint that $x + y = 2$. We rewrite the constraint as a function $g(x, y) = x + y - 2$ and we will try to impose $g(x, y) = 0$. Consider the Lagrangian function of three independent variables (x, y, λ) :

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda g(x, y) = x^2 + 2y^2 - \lambda(x + y - 2). \quad (5.5)$$

The stationary point of $\mathcal{L}(x, y, \lambda)$ with respect to the variables x, y , and λ corresponds to:

$$\frac{\partial \mathcal{L}(x, y, \lambda)}{\partial x} = 2x - \lambda = 0 \quad (5.6)$$

$$\frac{\partial \mathcal{L}(x, y, \lambda)}{\partial y} = 4y - \lambda = 0 \quad (5.7)$$

$$\frac{\partial \mathcal{L}(x, y, \lambda)}{\partial \lambda} = g(x, y) = 0. \quad (5.8)$$

From Eqs. (5.6) and (5.7) we obtain:

$$x = 2y, \quad (5.9)$$

which combined with Eq. (5.8) gives:

$$g(2y, y) = 2y + y - 2 = 0 \Rightarrow y = \frac{2}{3}, \quad (5.10)$$

and $x = \frac{4}{3}$. ■

To enforce orthonormality of the spin orbitals we need to enforce the condition:

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \quad \forall i, j \in \text{occupied}. \quad (5.11)$$

To enforce this constraint we consider the following Lagrangian:

$$\mathcal{L}[\{\psi_i\}] = E_0[\{\psi_i\}] - \sum_{ij}^{\text{occ}} \lambda_{ji} (\langle \psi_i | \psi_j \rangle - \delta_{ij}), \quad (5.12)$$

¹In writing the Lagrange multiplier we chose for convenience to put the index “q” before “p”. This notation will be convenient later in the derivation.

²That is $\lambda_{ji} = \lambda_{ji}^*$.

³A complex Lagrangian cannot be minimized because there is no such a thing as the minimum of a complex function.

where λ_{ji} is a matrix of Lagrange multipliers.¹ Note that λ_{ji} must be a Hermitian matrix², which is a consequence of requiring the Lagrangian to be a real function.³ To find the minimum of $\mathcal{L}[\{\psi_i\}]$ we require that the Lagrangian is stationary with respect to small variations of the spin orbitals $\{\delta\psi_i\}$. This means that for any arbitrary variation of the spin orbitals $\{\delta\psi_i\}$ the Lagrangian must satisfy:

$$\mathcal{L}[\{\psi_i + \delta\psi_i\}] = \mathcal{L}[\{\psi_i\}], \quad (5.13)$$

up to first order in $\{\delta\psi_i\}$. Let us first evaluate the variation of the energy:

$$\begin{aligned} E_0[\{\psi_i + \delta\psi_i\}] &= \sum_i^{\text{occ}} \langle \psi_i + \delta\psi_i | \hat{h} | \psi_i + \delta\psi_i \rangle \\ &\quad + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle (\psi_i + \delta\psi_i)(\psi_j + \delta\psi_j) | | (\psi_i + \delta\psi_i)(\psi_j + \delta\psi_j) \rangle \\ &= E_0[\{\psi_i\}] + \sum_i^{\text{occ}} \langle \psi_i | \hat{h} | \delta\psi_i \rangle + \sum_i^{\text{occ}} \langle \delta\psi_i | \hat{h} | \psi_i \rangle \\ &\quad + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle \delta\psi_i \psi_j | | \psi_i \psi_j \rangle + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle \psi_i \delta\psi_j | | \psi_i \psi_j \rangle \\ &\quad + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle \psi_i \psi_j | | \delta\psi_i \psi_j \rangle + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle \psi_i \psi_j | | \psi_i \delta\psi_j \rangle \\ &\quad + (\text{higher order}) \\ &= E_0[\{\psi_i\}] + \sum_i^{\text{occ}} \left[\langle \delta\psi_i | \hat{h} | \psi_i \rangle + \sum_j^{\text{occ}} \langle \delta\psi_i \psi_j | | \psi_i \psi_j \rangle \right] \\ &\quad + \text{h.c.} + (\text{higher order}), \end{aligned} \quad (5.14)$$

where “h.c.” stands for Hermitian conjugate.

The variation of the constraint $-\sum_{ij}^{\text{occ}} \lambda_{ji}(\langle \psi_p | \psi_q \rangle - \delta_{pq})$ is equal to:

$$\begin{aligned} -\sum_{ij}^{\text{occ}} \lambda_{ji}(\langle \psi_i + \delta \psi_i | \psi_j + \delta \psi_j \rangle - \delta_{ij}) &= -\sum_{ij}^{\text{occ}} \lambda_{ji}(\langle \psi_i | \psi_j \rangle - \delta_{ij}) \\ &\quad - \sum_{ij}^{\text{occ}} \lambda_{ji} \langle \delta \psi_i | \psi_j \rangle - \sum_{ij}^{\text{occ}} \lambda_{ji} \langle \psi_i | \delta \psi_j \rangle \\ &\quad + (\text{higher order}). \end{aligned} \quad (5.15)$$

Let us identify the change in the Lagrangian up to linear terms, that is:

$$\delta \mathcal{L} = \mathcal{L}[\{\psi_i + \delta \psi_i\}] - \mathcal{L}[\{\psi_i\}] \quad (5.16)$$

$$\delta \mathcal{L} = \sum_i^{\text{occ}} \left[\langle \delta \psi_i | \hat{h} | \psi_i \rangle + \sum_j^{\text{occ}} \langle \delta \psi_i \psi_j | | \psi_i \psi_j \rangle - \sum_{ij}^{\text{occ}} \lambda_{ji} \langle \delta \psi_i | \psi_j \rangle \right] + \text{h.c.} \quad (5.17)$$

We now rewrite this expression in terms of integrals:

$$\begin{aligned} \delta \mathcal{L} &= \sum_i^{\text{occ}} \int dx_1 \delta \psi_i^*(x_1) \times \\ &\quad \times \left[\hat{h} \psi_i(x_1) + \sum_j^{\text{occ}} \int dx_2 \frac{|\psi_j(x_2)|^2 \psi_i(x_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} - \sum_j^{\text{occ}} \int dx_2 \frac{\psi_j^*(x_2) \psi_i(x_2) \psi_j(x_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} - \sum_j^{\text{occ}} \lambda_{ji} \psi_j(x_1) \right]. \end{aligned} \quad (5.18)$$

If we require $\delta \mathcal{L}$ to be null for any variation of the orbitals then the quantity inside the square bracket must be identically null. We can write this condition in the following way, which is suggestive of an eigenvalue problem:

$$\hat{h} \psi_i(x_1) + \sum_j^{\text{occ}} \int dx_2 \frac{|\psi_j(x_2)|^2 \psi_i(x_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} - \sum_j^{\text{occ}} \int dx_2 \frac{\psi_j^*(x_2) \psi_i(x_2) \psi_j(x_1)}{|\mathbf{r}_1 - \mathbf{r}_2|} = \sum_j^{\text{occ}} \lambda_{ji} \psi_j(x_1). \quad (5.19)$$

If we introduce the Coulomb (\hat{J}_j) and exchange (\hat{K}_j) operators:

$$\hat{J}_j(x_1) \psi_i(x_1) = \int dx_2 \frac{|\psi_j(x_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_i(x_1) \quad (5.20)$$

$$\hat{K}_j(x_1) \psi_i(x_1) = \int dx_2 \frac{\psi_j^*(x_2) \psi_i(x_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_j(x_1), \quad (5.21)$$

we may rewrite the stationary conditions as:

$$\left[\hat{h}(x_1) + \sum_j^{\text{occ}} \hat{J}_j(x_1) - \hat{K}_j(x_1) \right] \psi_i(x_1) = \sum_j^{\text{occ}} \lambda_{ji} \psi_j(x_1), \quad (5.22)$$

or

$$\hat{f}(x_1) \psi_i(x_1) = \sum_j^{\text{occ}} \lambda_{ji} \psi_j(x_1). \quad (5.23)$$

Eq. (5.23) is called the **Hartree-Fock equation in noncanonical form** and the operator $\hat{f}(x_1) = \hat{h}(x_1) + \sum_j^{\text{occ}} \hat{J}_j(x_1) - \hat{K}_j(x_1)$ is called the Fock operator. Note, that this is not an eigenvalue problem because on the r.h.s. we have a sum of functions. Essentially this equation says that when we apply the Fock operator to any occupied orbital, we get back a linear combination of occupied orbitals. This implies that if we apply the Fock operator to a unoccupied (virtual) orbital, then we get back a linear combination of virtual orbitals.

5.1.2 Hartree–Fock equations in the canonical form

Consider a unitary transformation of the occupied orbitals:

$$|\psi'_i\rangle = \sum_j^{\text{occ}} |\psi_j\rangle U_{ji}, \quad (5.24)$$

$${}^4\mathbf{U}^{-1} = \mathbf{U}^\dagger$$

where the matrix \mathbf{U} is unitary.⁴ It is possible to prove that a determinant built from the spin orbitals ψ'_i is related to the determinant in the original basis by a complex phase factor:

$$|\Phi'\rangle = |\psi'_1 \psi'_2 \cdots \psi'_N\rangle = e^{i\theta} |\Phi\rangle. \quad (5.25)$$

From this it can be seen that any property that is given by expectation value is invariant:

$$\langle \Phi' | \hat{A} | \Phi' \rangle = \langle \Phi | \hat{A} | \Phi \rangle e^{-i\theta} e^{i\theta} = \langle \Phi | \hat{A} | \Phi \rangle. \quad (5.26)$$

It is possible to prove that the Coulomb and exchange operators are invariant:

$$\sum_j \hat{J}'_j(x_1) = \sum_j \hat{J}_j(x_1), \quad (5.27)$$

$$\sum_j \hat{K}'_j(x_1) = \sum_j \hat{K}_j(x_1) \quad (5.28)$$

and consequently the Fock operator is invariant with respect to this orbital rotation, that is, $\hat{f}' = \hat{f}$. From Eq. (5.23) it is easy to show that:

$$\langle \psi_j | \hat{f} | \psi_i \rangle = \lambda_{ji}. \quad (5.29)$$

In the unitarily-transformed basis this condition reads:

$$\langle \psi'_j | \hat{f}' | \psi'_i \rangle = \lambda'_{ji} = \sum_{kl} U_{lj}^* \langle \psi_l | \hat{f}' | \psi_k \rangle U_{ki} = \sum_{kl} U_{lj}^* \lambda_{lk} U_{ki}, \quad (5.30)$$

which may be written in the equivalent form:

$$\boldsymbol{\lambda}' = \mathbf{U}^\dagger \boldsymbol{\lambda} \mathbf{U}. \quad (5.31)$$

This result tells us that the Lagrange multipliers in the transformed basis are a unitary transformation of the original Lagrange multipliers. If we choose the unitary transformation in such a way that $\boldsymbol{\lambda}'$ is diagonal, then we can rewrite the Hartree–Fock equation as:

$$\hat{f}'(x_1) \psi'_i(x_1) = \lambda'_{ii} \psi'_i(x_1). \quad (5.32)$$

Eq. (5.33) is the **Hartree–Fock equation in canonical form**. The orbitals that satisfy this equation are called canonical Hartree–Fock MOs. In writing this equation we may assume that we are using the set of canonical orbitals so we can drop the prime and replace λ'_{ii} with ϵ_i :

$$\hat{f}(x_1) \psi_i(x_1) = \epsilon_i \psi_i(x_1). \quad (5.33)$$

The quantity ϵ_i is called the canonical orbital energy.

5.1.3 Meaning of the Hartree–Fock orbital energies. Koopman's theorem

From the canonical Hartree–Fock equation [Eq. (5.33)] we can derive an expression for the energy of an orbital:

$$\epsilon_i = \langle \psi_i | \hat{f} | \psi_i \rangle = \langle i | \hat{h} | i \rangle + \sum_k^{\text{occ}} \langle ik || ik \rangle. \quad (5.34)$$

By manipulation of the energy of a determinant we can relate it to the orbital energies:

$$\begin{aligned}
 E_0[\{\psi_i\}] &= \sum_i^{\text{occ}} \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{ij}^{\text{occ}} \langle ij || ij \rangle \\
 &= \sum_i^{\text{occ}} \langle i | \hat{h} | i \rangle + \sum_{ij}^{\text{occ}} \langle ij || ij \rangle - \frac{1}{2} \sum_{ij}^{\text{occ}} \langle ij || ij \rangle \\
 &= \sum_i^{\text{occ}} \epsilon_i - \frac{1}{2} \sum_{ij}^{\text{occ}} \langle ij || ij \rangle.
 \end{aligned} \tag{5.35}$$

It is important to note that the HF energy is not the sum of orbital energies.

Now consider a Slater determinant built from the Hartree-Fock orbitals in which one of the spin orbitals, in this case the last one, has been removed:

$$|\Phi_0^{N-1}\rangle = |\psi_1, \psi_2, \dots, \psi_{N-1}\rangle. \tag{5.36}$$

The energy of Φ_0^{N-1} is:

$$E_0^{N-1}[\{\psi_i\}] = \sum_i^{N-1} \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{ij}^{N-1} \langle ij || ij \rangle. \tag{5.37}$$

and the difference $E_0 - E_0^{N-1}$ is given by:⁵

$$\begin{aligned}
 E_0 - E_0^{N-1} &= \langle N | \hat{h} | N \rangle + \frac{1}{2} \sum_i^{N-1} \langle iN || iN \rangle + \frac{1}{2} \sum_j^{N-1} \langle Nj || Nj \rangle \\
 &\quad + \langle NN || NN \rangle \\
 &= \langle N | \hat{h} | N \rangle + \sum_i^{N-1} \langle iN || iN \rangle = \epsilon_N.
 \end{aligned} \tag{5.38}$$

⁵Note that the difference between $\sum_{ij}^N A_{ij} - \sum_{ij}^{N-1} A_{ij}$ is equal to $A_{NN} + \sum_i^{N-1} A_{iN} + \sum_j^{N-1} A_{Nj}$.

This result shows that the energy of the last occupied orbital is the energy necessary to remove one electron (ionize) from the system. This result may be generalized to any occupied orbital. In general we find that for a general occupied spin orbital (ψ_i):

$$\text{Ionization potential} = \text{IP} = E_0^{N-1} - E_0^N = -\epsilon_i. \tag{5.39}$$

By a similar procedure it is possible to show that for a generic unoccupied orbital (ψ_a):

$$\text{Electron affinity} = \text{EA} = E_0^N - E_0^{N+1} = -\epsilon_a. \tag{5.40}$$

These results are called Koopman's theorem. Note that orbitals are assumed to be frozen, which means that we are ignoring relaxation effects to the addition or removal of electrons. Koopman's theorem also ignores electron correlation, so the IP and EA will deviate from experiment.

We can also investigate the energy of an electronically excited determinant. Consider for example the orbital substitution $\psi_i \rightarrow \psi_a$ to give the determinant Φ_i^a . The energy of this excited determinant is:

$$\begin{aligned}
 \langle \Phi_i^a | \hat{H} | \Phi_i^a \rangle &= \sum_j^N \langle j | \hat{h} | j \rangle + \langle a | \hat{h} | a \rangle - \langle i | \hat{h} | i \rangle \\
 &\quad + \frac{1}{2} \sum_{ij}^N \langle ij || ij \rangle + \sum_j^N \langle aj || aj \rangle - \sum_j^N \langle ij || ij \rangle \\
 &\quad - \langle ia || ia \rangle,
 \end{aligned} \tag{5.41}$$

where we subtract $\langle ia||ia \rangle$ to avoid counting the interaction of an electron in ψ_a with ψ_i . If we collect terms that contribute to E_0 , ϵ_a , and ϵ_i , we may rewrite the energy of an excited state as:

$$\langle \Phi_i^a | \hat{H} | \Phi_i^a \rangle = E_0 + \epsilon_a - \epsilon_i - \underbrace{\langle ia||ia \rangle}_{\text{hole/electron interaction}}. \quad (5.42)$$

This expression shows that the excitation energy is the difference between orbital energies minus a term that represents the Coulomb/exchange interaction of the hole created by removing an electron from ψ_i and the electron in ψ_a .

5.1.4 Generalized, unrestricted, and restricted Hartree–Fock theory

So far we have not imposed any restriction on the spin orbital basis. Within HF theory one may form a hierarchy of increasingly more constrained theory. The most general case is *generalized* Hartree–Fock (GHF). In this formalism spin orbitals are a linear combination of a alpha plus beta component:

$$\psi_p^{\text{GHF}}(\mathbf{r}, \omega) = \phi_{p,\alpha}(\mathbf{r})\alpha(\omega) + \phi_{p,\beta}(\mathbf{r})\beta(\omega). \quad (5.43)$$

In GHF electrons is also called a non-collinear HF because electron spins are not constrained to be eigenfunctions of \hat{S}_z .

Unrestricted Hartree–Fock (UHF) restricts spin orbitals to be alpha or beta. If we arrange spin orbitals in such a way that alpha and beta spin alternate we can write the unrestricted spin orbitals as:

$$\begin{aligned} \psi_{2p-1}^{\text{UHF}}(\mathbf{r}, \omega) &= \phi_{2p-1}(\mathbf{r})\alpha(\omega), \\ \psi_{2p}^{\text{UHF}}(\mathbf{r}, \omega) &= \phi_{2p}(\mathbf{r})\beta(\omega), \end{aligned} \quad (5.44)$$

with $p = 1, \dots, K$, where K is the size of the spatial orbital basis. We are free to arrange these functions and later we will use an indexing that puts alpha spin orbitals before beta spin orbitals:

$$\psi_p^{\text{UHF}}(\mathbf{r}, \omega) = \phi_p(\mathbf{r}) \begin{cases} \alpha(\omega) & \text{if } p < K \\ \beta(\omega) & \text{if } p \geq K \end{cases}, \quad (5.45)$$

with $p = 1, \dots, 2K$.

Restricted Hartree–Fock (RHF) theory constrains spin orbitals to have alpha or beta spin. Spin orbitals come in pairs with identical spatial components:

$$\begin{aligned} \psi_{2p-1}^{\text{RHF}}(\mathbf{r}, \omega) &= \phi_p(\mathbf{r})\alpha(\omega), \\ \psi_{2p}^{\text{RHF}}(\mathbf{r}, \omega) &= \phi_p(\mathbf{r})\beta(\omega), \end{aligned} \quad (5.46)$$

with $p = 1, \dots, K$. If the spin orbitals are arranged so that the alpha spin orbitals come before the beta spin orbitals then:

$$\psi_p^{\text{RHF}}(\mathbf{r}, \omega) = \begin{cases} \phi_p(\mathbf{r})\alpha(\omega) & \text{if } p < K \\ \phi_{p-K}(\mathbf{r})\beta(\omega) & \text{if } p \geq K \end{cases}, \quad (5.47)$$

with $p = 1, \dots, 2K$.

5.1.5 Closed-shell restricted Hartree–Fock theory

The Hartree–Fock equations can be further simplified in the case of closed-shell species with equal number of alpha and beta electrons. In this case, we assume electrons with alpha and beta spin are paired in orbitals with identical spatial component and the wave function is approximated with the following Slater determinant:

$$|\Phi_0^{\text{RHF}}\rangle = |\phi_1\alpha, \dots, \phi_{N/2}\alpha, \phi_1\beta, \dots, \phi_{N/2}\beta\rangle. \quad (5.48)$$

Expressions for the energy and Fock matrix in terms of spatial orbitals can be derived by expressing a sum over spin orbitals in terms of sums over spatial orbitals:

$$\sum_i^N \psi_i = \sum_i^{N/2} \phi_i \alpha + \sum_i^{N/2} \phi_i \beta. \quad (5.49)$$

In this case the energy expression then takes the simplified form:

$$\begin{aligned} E_0 = & \sum_{i=1}^{N/2} \langle i | \hat{h} | i \rangle + \sum_{i=1}^{N/2} \langle \bar{i} | \hat{h} | \bar{i} \rangle \\ & \frac{1}{2} \sum_{ij}^{N/2} \langle ij | ij \rangle + \frac{1}{2} \sum_{i\bar{j}}^{N/2} \langle i\bar{j} | i\bar{j} \rangle + \frac{1}{2} \sum_{\bar{i}j}^{N/2} \langle \bar{i}j | \bar{i}j \rangle + \frac{1}{2} \sum_{\bar{i}\bar{j}}^{N/2} \langle \bar{i}\bar{j} | \bar{i}\bar{j} \rangle. \end{aligned} \quad (5.50)$$

We can use the fact that in the restricted closed-shell case:

$$\langle i | \hat{h} | i \rangle = \langle \bar{i} | \hat{h} | \bar{i} \rangle, \quad (5.51)$$

while for the two-electron integrals:

$$\langle ij | ij \rangle = \langle \bar{i}\bar{j} | \bar{i}\bar{j} \rangle = \langle ij | ij \rangle - \langle ij | ji \rangle, \quad (5.52)$$

and

$$\langle i\bar{j} | i\bar{j} \rangle = \langle \bar{i}j | \bar{i}j \rangle = \langle ij | ij \rangle, \quad (5.53)$$

to write the energy expression as:

$$E_0 = 2 \sum_{i=1}^{N/2} \langle i | \hat{h} | i \rangle + \sum_{ij}^{N/2} [2 \langle ij | ij \rangle - \langle ij | ji \rangle]. \quad (5.54)$$

Similarly, we can express the Fock matrix as:

$$\langle i | \hat{f} | j \rangle = \langle i | \hat{h} | j \rangle + \sum_k^{N/2} [2 \langle ik | jk \rangle - \langle ik | kj \rangle]. \quad (5.55)$$

5.1.6 The Roothan equations

In this section we will derive the Roothan equations. These equations correspond to the Hartree-Fock equations expressed in a finite computational basis $\{\chi_\mu\}$. The main idea is to express the spatial molecular orbitals $\phi_i(\mathbf{r})$ as a linear combination of the computational basis:

$$\phi_i(\mathbf{r}) = \sum_{\mu} \chi_{\mu}(\mathbf{r}) C_{\mu i}. \quad (5.56)$$

Let us consider the Hartree-Fock equation:

$$\hat{f} |\psi_i\rangle = |\psi_i\rangle \epsilon_i, \quad (5.57)$$

next, insert the expansion on the computational basis and project on the left with $\langle \chi_\nu |$ to get:

$$\sum_{\mu} \langle \chi_\nu | \hat{f} | \chi_\mu \rangle C_{\mu i} = \sum_{\mu} \langle \chi_\nu | \chi_\mu \rangle C_{\mu i} \epsilon_i. \quad (5.58)$$

In general, the computational basis may not be orthonormal. Therefore, we write $\langle \chi_\nu | \chi_\mu \rangle = S_{\nu\mu}$, where $S_{\nu\mu}$ is the *overlap matrix*. If we denote the matrix elements of the Fock operator as $F_{\nu\mu} = \langle \chi_\nu | \hat{f} | \chi_\mu \rangle$, we may write:

$$\sum_{\mu} F_{\nu\mu} C_{\mu i} = \sum_{\mu} S_{\nu\mu} C_{\mu i} \epsilon_i, \quad (5.59)$$

or in matrix notation

$$\mathbf{F}\mathbf{C} = \mathbf{S}\mathbf{C}\epsilon, \quad (5.60)$$

where ϵ is a diagonal matrix. This is Roothan's form of the Hartree-Fock equations. Note that due to the presence of \mathbf{S} , this is a generalized eigenvalue problem.

The matrix \mathbf{F} is the Fock operator expressed in the computational basis and is given by:

$$\begin{aligned} (\mathbf{F})_{\mu\nu} &= f_{\mu\nu} = \langle \chi_\mu | \hat{h} | \chi_\nu \rangle + \sum_k^{N/2} [2 \langle \chi_\mu \phi_k | \chi_\nu \phi_k \rangle - \langle \chi_\mu \phi_k | \phi_k \chi_\nu \rangle] \\ &= h_{\mu\nu} + \sum_k^{N/2} [2 \langle \chi_\mu \chi_\rho | \chi_\nu \chi_\sigma \rangle - \langle \chi_\mu \chi_\rho | \chi_\sigma \chi_\nu \rangle] C_{\rho k}^* C_{\sigma k}. \end{aligned} \quad (5.61)$$

If we introduce the density matrix:

$$P_{\rho\sigma} = \sum_k^{N/2} C_{\rho k}^* C_{\sigma k}, \quad (5.62)$$

the Fock matrix may be expressed as a function of $P_{\rho\sigma}$:

$$f_{\mu\nu}(\mathbf{P}) = h_{\mu\nu} + \underbrace{\sum_{\rho\sigma} [2 \langle \chi_\mu \chi_\rho | \chi_\nu \chi_\sigma \rangle - \langle \chi_\mu \chi_\rho | \chi_\sigma \chi_\nu \rangle] P_{\rho\sigma}}_{G_{\mu\nu}}. \quad (5.63)$$

Since the Fock matrix depends on \mathbf{P} and ultimately on \mathbf{C} , the Hartree-Fock equations are nonlinear and require an iterative solution:

$$\mathbf{F}(\mathbf{C})\mathbf{C} = \mathbf{S}\mathbf{C}\epsilon, \quad (5.64)$$

How can we reduce this generalized eigenvalue problem to a normal eigenvalue problem? Consider a transformation of the computational basis:

$$\chi_\mu \rightarrow \chi'_\mu = \sum_\nu \chi_\nu X_{\nu\mu}. \quad (5.65)$$

The molecular orbital represented in the $\{\chi_\mu\}$ and $\{\chi'_\mu\}$ basis are related via:

$$\phi_i = \sum_\mu \chi_\mu C_{\mu i} = \sum_\mu \chi'_\mu C'_{\mu i} = \sum_{\mu\nu} \chi_\nu X_{\nu\mu} C'_{\mu i}, \quad (5.66)$$

from which we get

$$\mathbf{C} = \mathbf{X}\mathbf{C}'. \quad (5.67)$$

Suppose we choose \mathbf{X} to satisfy the condition:

$$\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbf{1}, \quad (5.68)$$

then we can take Roothan's equation, multiply on the left with \mathbf{X}^\dagger , and substitute $\mathbf{C} = \mathbf{X}\mathbf{C}'$ to get:

$$\underbrace{\mathbf{X}^\dagger \mathbf{F} \mathbf{X}}_{\mathbf{F}'} \mathbf{C}' = \underbrace{\mathbf{X}^\dagger \mathbf{S} \mathbf{X}}_{\mathbf{1}} \mathbf{C}' \epsilon \quad (5.69)$$

$$\mathbf{F}' \mathbf{C}' = \mathbf{C}' \epsilon.$$

Thanks to this transformation we were able to transform Roothan's equation (a generalized eigenvalue problem) to a normal eigenvalue problem. Note that the definition of the Fock matrix and eigenvector have changed, but the eigenvalues of this equation are identical to

those of the generalized eigenvalue problem. To solve Roothan's equation in the orthonormal basis we first compute $\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X}$, then solve the eigenvalue equation, $\mathbf{F}' \mathbf{C}' = \mathbf{C}' \epsilon$, and then obtain the eigenvector from $\mathbf{C} = \mathbf{X} \mathbf{C}'$.

How can we pick a transformation \mathbf{X} such that $\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbf{1}$? It is easy to see that there are several ways to choose \mathbf{X} . In the symmetric orthogonalization procedure \mathbf{X} is chosen to be the (matrix) square root of \mathbf{S} :

$$\mathbf{X} = \mathbf{S}^{-\frac{1}{2}}. \quad (5.70)$$

This form of \mathbf{X} fits the role since:

$$\mathbf{S}^{-\frac{1}{2}} \mathbf{S} \mathbf{S}^{-\frac{1}{2}} = \mathbf{S}^{-\frac{1}{2}} \mathbf{S}^{\frac{1}{2}} = \mathbf{1}. \quad (5.71)$$

The matrix $\mathbf{S}^{-\frac{1}{2}}$ may be built by first diagonalizing \mathbf{S} with the unitary matrix \mathbf{U} :

$$\mathbf{S} \mathbf{U} = \mathbf{U} \mathbf{s}, \quad (5.72)$$

and then building

$$\mathbf{S}^{-\frac{1}{2}} = \mathbf{U} \mathbf{s}^{-\frac{1}{2}} \mathbf{U}^\dagger. \quad (5.73)$$

Alternatively, in the canonical orthogonalization procedure, \mathbf{X} is defined as:

$$\mathbf{X} = \mathbf{U} \mathbf{s}^{-\frac{1}{2}}. \quad (5.74)$$

This form of \mathbf{X} also does the job:

$$\mathbf{s}^{-\frac{1}{2}} \mathbf{U}^\dagger \mathbf{S} \mathbf{U} \mathbf{s}^{-\frac{1}{2}} = \mathbf{s}^{-\frac{1}{2}} \mathbf{s} \mathbf{s}^{-\frac{1}{2}} = \mathbf{1}. \quad (5.75)$$

5.1.7 Outline of the Hartree-Fock-Roothan procedure

In this section we will discuss the computational steps required to solve the Hartree-Fock equations projected onto a finite computational basis. For now we will assume that this computational basis is an atomic basis but we will neglect the details of how it is constructed and what type of functions are used. Note that the computational basis might also be a set of delocalized functions that does not have a well defined atomic character.

The input to a HF computation is the molecule, which is specified by the coordinates ($\{\mathbf{R}_A\}$) and charges ($\{\mathbf{Z}_A\}$) of the nuclei, the number of electrons (N), and the projection of spin onto the z axis (M_S). We also have to provide a computation basis $\{\chi_\mu\}$.

The major steps in the Hartree-Fock procedure are:

1. **Evaluation of the integrals over basis functions.** The following integrals are needed:

$$S_{\mu\nu} = \int d\mathbf{r} \chi_\mu^*(\mathbf{r}) \chi_\nu(\mathbf{r}), \quad (5.76)$$

$$V_{\mu\nu} = \int d\mathbf{r} \chi_\mu^*(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + V_{eN}(\mathbf{r}) \right] \chi_\nu(\mathbf{r}), \quad (5.77)$$

and the two-electron integrals $\langle \mu\nu | \rho\sigma \rangle = \langle \chi_\mu \chi_\nu | \chi_\rho \chi_\sigma \rangle$.

2. **Diagonalization of the overlap matrix.** From \mathbf{S} we solve the eigenvalue problem:

$$\mathbf{S} \mathbf{U} = \mathbf{U} \mathbf{s}. \quad (5.78)$$

3. **Formation of the matrix \mathbf{X} .** The matrix \mathbf{X} is built as

$$\mathbf{X} = \mathbf{U} \mathbf{s}^{-\frac{1}{2}}. \quad (5.79)$$

4. **Initialization of the density matrix.** The density matrix \mathbf{P} is initialized to zero:

$$\mathbf{P}^{(0)} = \mathbf{0}. \quad (5.80)$$

5. **Computation of the two-electron contribution to the Fock matrix.** The two matrix \mathbf{G} is built from the current density matrix ($\mathbf{P}^{(k)}$) and the two-electron integrals:

$$\mathbf{G} = \mathbf{G}[\mathbf{P}^{(k)}, \langle \mu\nu | \rho\sigma \rangle]. \quad (5.81)$$

6. **Formation the Fock matrix.**

$$\mathbf{F} = \mathbf{H} + \mathbf{G}. \quad (5.82)$$

7. **Transformation the atomic orthonormal basis.**

$$\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X}. \quad (5.83)$$

8. **Diagonalization of \mathbf{F}' .** We solve the eigenvalue problem:

$$\mathbf{F}' \mathbf{C}' = \mathbf{F}' \boldsymbol{\epsilon}. \quad (5.84)$$

9. **Backtransformation to the nonorthogonal basis.**

$$\mathbf{C} = \mathbf{X} \mathbf{C}'. \quad (5.85)$$

10. **Update of the density matrix and energy.** Using the updated MO coefficient (\mathbf{C}) we compute an updated density matrix

$$P_{\mu\nu}^{(k+1)} = \sum_i^N C_{\mu i}^* C_{\nu i}. \quad (5.86)$$

Using the updated density matrix we recompute the energy. The change in energy:

$$\Delta E^{(k+1)} = E^{(k+1)} - E^{(k)}, \quad (5.87)$$

and the change in the density matrix:

$$\Delta \mathbf{P}^{(k+1)} = \mathbf{P}^{(k+1)} - \mathbf{P}^{(k)}, \quad (5.88)$$

are used to determine if the Hartree-Fock equation have converged. If $|\Delta E^{(k+1)}| \leq \eta_E$ and $\|\Delta \mathbf{P}^{(k+1)}\| \leq \eta_P$, where η_E and η_P are user-provided convergence thresholds, then the HF procedure is terminated. Otherwise, we go back to step 5 and build the matrix \mathbf{G} using the updated density matrix. This procedure is iterated until convergence.