2.3 Hartree-Fock

The Hartree-Fock method assumes that the true N-body ground state wave function can be approximated by a single Slater determinant and minimizes the energy among all possible choices of a single Slater determinant wave function, $|\Psi_S\rangle = |\chi_1 \ldots \chi_N\rangle$. This is what we do next.

2.3.1 A derivation of the Fock operator in second quantization

In this section we derive the Fock operator for closed shell real Slater determinants in second quantization ([Helgaker 10]).

We showed on the last lecture that the creation operators $\{\tilde{a}_j^\dagger\}$ corresponding to MOs $\{\tilde{\psi}^{(j)}\}$ can be obtained from the creation operators $\{a_j^\dagger\}$ through a transformation $\tilde{a}_p^\dagger = \exp(-\hat{u}) a_p^\dagger \exp(\hat{u})$.

The operator $\hat{u}$ is defined as $\hat{u} = \sum_{pq} u_{pq} a_p^\dagger a_q$, $\hat{u}^\dagger = -\hat{u}$.

We also deduced that transformations between Slater determinants are always of the form

$$|\tilde{k}\rangle = \prod_p (\tilde{a}_p^\dagger)^{k_p} |\text{vac}\rangle = e^{-\hat{u}} \prod_p (a_p^\dagger)^{k_p} e^{\hat{u}} |\text{vac}\rangle = e^{-\hat{u}} \prod_p (a_p^\dagger)^{k_p} |\text{vac}\rangle = e^{-\hat{u}} |k\rangle,$$

where we have used $\hat{u} |\text{vac}\rangle = 0$.

We now introduce spin operators in the second quantization formalism. We simply define operators $a_{p\sigma}^\dagger$ with $\sigma \in \{\alpha, \beta\}$ that create a spin orbital with spin $\sigma$. It is clear that all the second quantization machinery still works. Because we are working with a spin independent Hamiltonian and most of the transformations we are interested in are also spin independent, spin operators will usually appear in pairs. For instance, the electronic Hamiltonian is

$$H = h_{pq} (\sum_{\sigma} a_{p\sigma}^\dagger a_{q\sigma}) + \frac{1}{2} \sum_{pqr} g_{pqrst} (\sum_{\sigma} a_{p\sigma}^\dagger a_{r\sigma}^\dagger a_{s\tau} a_{q\tau}),$$

with

$$h_{pq} = \int d\mathbf{r}_1 \psi^{(p)}(\mathbf{r}_1)^* \left( -\frac{1}{2} \nabla^2 - \sum_{J} \frac{Z_J}{r_{1J}} \right) \psi^{(q)}(\mathbf{r}_1)$$

and

$$g_{pqrst} = \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\psi^{(p)}(\mathbf{r}_1)^* \psi^{(r)}(\mathbf{r}_2)^* \psi^{(q)}(\mathbf{r}_1) \psi^{(s)}(\mathbf{r}_2)}{r_{12}}.$$
We introduce excitation operators
\[ E_{qp} = a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta} \].

We can write any closed-shell state as
\[ |\text{cs}\rangle = e^{-\hat{\mathbf{v}}}|1_{1\alpha}1_{1\beta}...1_{N/2\beta}0_{N+1}...0_M\rangle \],
with the rotation \( \hat{\mathbf{v}} \) restricted to have the form\(^1\)
\[ \hat{\mathbf{v}} = \sum_{pq} v_{pq} E_{pq} \].

We are also going to distinguish between inactive orbitals, which are doubly occupied in \( |\text{cs}\rangle \) and virtual orbitals, which are unoccupied. We will use the indexes \( i, j, k, l \) for inactive orbitals, \( a, b, c, e, d \) for virtual orbitals and \( m, n, o, p, q, r, s, t, u \) for arbitrary orbitals. Notice that
\[ |\text{cs}\rangle = \prod_i a_{i\alpha}^\dagger a_{i\beta}^\dagger |\text{vac}\rangle \].

Then it follows that
\[ E_{ab}|\text{cs}\rangle = 0 \],
because \( b \) denotes a virtual orbital and \( a_{b\sigma}|\text{cs}\rangle = 0 \). This implies \( E_{ab}|\text{cs}\rangle = 0 \), and we can set the corresponding parameter to 0 in \( \hat{\mathbf{v}} \). Also
\[ E_{kl}|\text{cs}\rangle = 0 \],
because \( k \neq l \) and these are inactive orbitals. That is, they are doubly occupied, so we can not create any further excitations there. In conclusion, the only parameters of \( \hat{\mathbf{v}} \) are \( v_{aj} \), which generate occupied-virtual rotations.

We want to find the closed shell state \( |\text{cs}\rangle \) which minimizes the energy
\[ \langle \text{cs}|H|\text{cs}\rangle \].

A necessary condition is that, if \( |\text{cs}\rangle \) is the HF solution, the expectation value of the energy is stationary for small transformations
\[ \frac{\delta}{\delta \hat{\mathbf{u}}} \langle \text{cs}|e^{\hat{\mathbf{u}}H}e^{-\hat{\mathbf{u}}}|\text{cs}\rangle \bigg|_{\hat{\mathbf{u}}=0} = 0 \].

We can expand the energy as commutators using the formula introduced in the previous section
\[ \langle \text{cs}|e^{\hat{\mathbf{u}}H}e^{-\hat{\mathbf{u}}}|\text{cs}\rangle = \langle \text{cs}|H|\text{cs}\rangle + \langle \text{cs}|[\hat{\mathbf{u}}, H]|\text{cs}\rangle + \cdots \]

\(^1\)Further, because we can restrict our attention to real orbitals, we could work with real operators of the general form
\[ \hat{\mathbf{v}} = \sum_{p>q} v_{pq}(E_{pq} - E_{qp}) = \sum_{p>q} v_{pq}E_{pq}^\dagger \].

53
The stationary condition for $|cs\rangle$ then becomes
\[
\frac{\partial}{\partial u_{pq}} \langle cs| e^{\hat{H}} e^{-\hat{u}} |cs\rangle \bigg|_{u=0} = \langle cs| [E_{pq}, H] |cs\rangle = 0
\]
Furthermore, from our discussion on rotations of closed shells states, it is enough to enforce
\[
\langle cs|[E_{ai}, H]|cs\rangle = 0 .
\]
What this means is that rotations only among virtual or inactive orbitals also generate acceptable solutions.

The Fock operator is an effective one-electron operator
\[
\hat{f} = \sum_{pq} f_{pq} E_{pq} ,
\]
with corresponding Hermitian matrix $(F)_{pq} = f_{pq}$ such that this matrix is diagonal in a stationary point of the energy. That is, for the spin orbitals of a solution $|cs\rangle$ we have
\[
f_{pq} = \delta_{qp} \varepsilon_p .
\]
The following possible choice for $f$ now jumps to mind
\[
f_{ai} \propto \langle cs|[E_{ai}, H]|cs\rangle .
\]
The reason is that if this matrix is diagonal in some basis, then, by the previous discussion, the corresponding $|cs\rangle$ state is an stationary point of the energy. As it turns out, this formula does not give an operator which reduces to the one electron Hamiltonian when there are no interactions, and the operator is not Hermitian either. These are properties that we would like the Fock operator to have, and which are reasonable because indeed the solution of the non-interacting problem is given by a single Slater determinant. We will show later that this is the case for the operator that we end up defining.

The Fock operator is obtained from the following manipulations
\[
\langle cs|[E_{ai}, H]|cs\rangle = \sum_{\sigma} \langle cs|[a_{a\sigma}^\dagger a_{i\sigma}, H]|cs\rangle
\]
\[
= \sum_{\sigma} \langle cs|[a_{a\sigma}^\dagger, H], a_{i\sigma}|cs\rangle + \langle cs|[a_{a\sigma}, a_{i\sigma}, H]|cs\rangle
\]
\[
= \sum_{\sigma} \langle cs|[a_{a\sigma}^\dagger, H], a_{i\sigma}|cs\rangle .
\]
The second term is 0 because $a_{a\sigma}|cs\rangle = 0$, as the index $a$ denotes a virtual orbital. Similarly
\[
\sum_{\sigma} \langle cs|[a_{a\sigma}^\dagger, H], a_{i\sigma}|cs\rangle = \sum_{\sigma} \langle cs|[a_{a\sigma}^\dagger, H], a_{i\sigma}|cs\rangle + \sum_{\sigma} \langle cs|[a_{a\sigma}, [a_{a\sigma}^\dagger, H]|cs\rangle
\]
\[
= \sum_{\sigma} \langle cs|[a_{a\sigma}^\dagger, H], a_{i\sigma}|cs\rangle ,
\]

54
using the fact that the index $i$ corresponds to an inactive (occupied) orbital. Now, because all the orbitals are real

$$\sum_{\sigma} \langle \text{cs}\{[a_{i\sigma}^\dagger, H], a_{i\sigma}\}\rangle = \sum_{\sigma} \langle \text{cs}\{[a_{i\sigma}^\dagger, H], a_{i\sigma}\}\rangle^*$$

$$= \sum_{\sigma} \langle \text{cs}\{[a_{i\sigma}^\dagger, H], a_{i\sigma}\}\rangle^\dagger$$

$$= -\sum_{\sigma} \langle \text{cs}|a_{i\sigma}^\dagger, [a_{i\sigma}, H]|\text{cs}\rangle .$$

Finally, we choose the normalization factor to define a candidate Fock operator as

$$f_{pq} = \frac{1}{2} \sum_{\sigma} \langle \text{cs}|{[a_{q\sigma}^\dagger, [a_{p\sigma}, H]]}\rangle .$$

We now only need to check that this gives the standard matrix elements of the usual Fock operator, but we already now that the stationary point spin orbitals are given by its eigenvectors.

### 2.3.2 Properties of the Fock operator

We first check that the operator is Hermitian. Because we have real wave functions

$$\langle \text{cs}|{[a_{q\sigma}^\dagger, [a_{p\sigma}, H]]}|\text{cs}\rangle = \langle \text{cs}|{[a_{p\sigma}, [H, a_{q\sigma}^\dagger]]}|\text{cs}\rangle .$$

Expanding the commutators

$$= a_{p\sigma} H a_{q\sigma}^\dagger - a_{p\sigma} a_{q\sigma}^\dagger H + H a_{q\sigma}^\dagger a_{p\sigma} - a_{q\sigma}^\dagger H a_{p\sigma}$$

$$= a_{p\sigma} H a_{q\sigma}^\dagger + a_{q\sigma}^\dagger a_{p\sigma} H - H a_{p\sigma} a_{q\sigma}^\dagger - a_{q\sigma}^\dagger H a_{p\sigma} ,$$

using the fermionic anti-commutation relation in the second line. And finally, grouping again

$$= \langle \text{cs}|{[a_{q\sigma}^\dagger, [a_{p\sigma}, H]]}|\text{cs}\rangle .$$

Now we check that it does indeed give the one-electron Hamiltonian in the non-interacting case. We use

$$[a_{p\sigma}, E_{rs}] = \sum_{\tau} \{a_{p\sigma}, a_{r\tau}^\dagger\} a_{s\tau} - a_{r\sigma}^\dagger \{a_{p\sigma}, a_{s\tau}\} = \delta_{pr} a_{s\sigma} .$$

We get

$$\frac{1}{2} \sum_{rs} h_{rs} \sum_{\sigma} \langle \text{cs}|{a_{r\sigma}^\dagger, [a_{p\sigma}, E_{rs]}]}|\text{cs}\rangle = \frac{1}{2} \sum_{rs} h_{rs} \sum_{\sigma} \langle \text{cs}|{a_{q\sigma}^\dagger, \delta_{pr} a_{s\sigma}}|\text{cs}\rangle$$

$$= \frac{1}{2} \sum_{rs} h_{rs} \sum_{\sigma} \delta_{pr} \delta_{qs} = h_{pq} .$$
For the electron-electron interaction part we use the general equation

\[
[A, B_1 \cdots B_n] = \sum_{k=1}^{n} (-1)^{k-1} B_1 \cdots \{A, B_k\} \cdots B_n \quad (n \text{ even}).
\]

We get, with the notation previously introduced for the corresponding operator,

\[
[a_{m\sigma}, g] = \frac{1}{3} \sum_{pqrs} \sum_{\tau\nu} g_{pqrs} \left[a_{m\sigma}, a^\dagger_{p\nu} a^\dagger_{r\tau} a_{s\nu} a_{q\sigma}\right]
= \frac{1}{3} \sum_{pqrs} \sum_{\tau\nu} g_{pqrs} \left(\delta_{mp} \delta_{\sigma\nu} a^\dagger_{r\tau} - \delta_{mr} \delta_{\sigma\tau} a^\dagger_{p\nu}\right) a_{s\nu} a_{q\sigma}.
\]

We transform the dummy indexes in the second part to get

\[
- \sum_{qrs} g_{rsmq} \sum_{\tau} a^\dagger_{r\tau} a_{q\sigma} a_{s\tau} = \sum_{qrs} g_{mqs} \sum_{\tau} a^\dagger_{r\tau} a_{s\tau} a_{q\sigma},
\]

using the symmetry of \(g\) and the fermionic anti-commutation relation. This gives

\[
[a_{m\sigma}, g] = \sum_{qrs} \sum_{\tau} g_{mqs} a^\dagger_{r\tau} a_{s\tau} a_{q\sigma}.
\]

To continue we need

\[
\{A, B_1 \cdots B_n\} = \sum_{k=1}^{n} (-1)^{k-1} B_1 \cdots \{A, B_k\} \cdots B_n \quad (n \text{ odd}).
\]

We get

\[
\sum_{\sigma} \left\{ a^\dagger_{n\sigma}, [a_{m\sigma}, g] \right\} = \sum_{qrs} g_{mqs} \sum_{\sigma\tau} \left\{ a^\dagger_{n\sigma}, a^\dagger_{r\tau} a_{s\tau} a_{q\sigma}\right\}
= \sum_{qrs} g_{mqs} \sum_{\sigma\tau} \left(\delta_{nq} a_{r\tau} a_{s\tau} - \delta_{ns} \delta_{\sigma\tau} a^\dagger_{r\tau} a_{q\sigma}\right)
= \sum_{rs} \left(2g_{mns} - g_{msn}\right) E_{rs}.
\]

We evaluate this last element

\[
D_{pq} = \langle cs|E_{pq}|cs \rangle = \begin{cases} 2\delta_{pq} & p \text{ occupied} \\ 0 & p \text{ virtual} \end{cases}
\]

Putting it all together, we finally get

\[
F_{pq} = h_{pq} + \sum_i \left(2g_{pqii} - g_{piiq}\right).
\]
2.3.3 Restricted closed shell Fock operator in first quantization

We can work out an expression in first quantization for the Fock operator in the simple restricted closed shell case, following the result of the last section. Remember that in this case, we necessarily have an even number of electrons, and each MO appears twice, once for each spin value.

The easy part is the core operator, the single electron part, \( h_{pq} \), which is just \( \langle \psi(p) | h | \psi(q) \rangle \). That is, for that part the corresponding operator is just the single \( h \) operator introduced before. As advertised.

Now we look at the electron-electron interaction. We have given expressions for the integrals corresponding to \( g_{pqrs} \) a couple of times already. Using “chemist” notation, the first part is

\[
\langle r_1 | J_i | \psi(q) \rangle = \int d r_2 \left| \psi^{(i)}(r_2) \right|^2 \frac{1}{r_{12}} \psi^{(q)}(r_1) .
\]

This is the Coulomb term \( J_i \) (from a previous lecture) acting on \(|\psi^{(q)}\rangle\)

\[
\langle r_1 | J_i | \psi(q) \rangle = \int d r_2 \left| \psi^{(i)}(r_2) \right|^2 \frac{1}{r_{12}} \psi^{(q)}(r_1) .
\]

The second part is

\[
g_{piiq} = \langle pi | iq \rangle = \int d r_1 d r_2 \psi^{(p)}(r_1)^* \frac{\psi^{(i)}(r_2)^* \psi^{(q)}(r_2)}{r_{12}} \psi^{(i)}(r_1) ,
\]

which we recognize as the exchange operator \( K_i \) acting on \(|\psi^{(q)}\rangle\)

\[
\langle r_1 | K_i | \psi(q) \rangle = \int d r_2 \psi^{(i)}(r_2)^* \frac{\psi^{(q)}(r_2)}{r_{12}} \psi^{(q)}(r_1) .
\]

All together, the Fock matrix are the elements of the Fock operator \([\text{Szabo } 3.4]\]

\[
f(r_1) = h(r_1) + 2 \sum_i J_i(r_1) - \sum_i K_i(r_1) .
\]

Notice that this version of the Fock operator acts on MOs, and therefore there is no spin coordinate (we are in the closed-shell model). Another interesting point is the factor of 2 in front of the Coulomb operator. This is because each MO now stands for the two original spin orbitals and this factor of 2 shows up for the Coulomb field, which depends upon all other spin orbitals.

But why is there not a factor of 2 in front of the exchange term? While this factors of 2 come out naturally in the derivation of the Fock operator for MOs, we can see the underlying reason by considering the expectation value

\[
\langle \psi^{(j)} \bar{\psi}^{(k)} \rangle \frac{1}{r_{12}} \left| \psi^{(j)} \bar{\psi}^{(k)} \right\rangle .
\]
We are going to get a Coulomb term like
\[
\langle \psi^{(j)} | \langle \alpha | \langle \psi^{(k)} | \langle \beta | \frac{1}{r_{12}} | \psi^{(j)} \rangle | \alpha \rangle | \psi^{(k)} \rangle | \beta \rangle = \langle \psi^{(j)} | \langle \psi^{(k)} | \frac{1}{r_{12}} | \psi^{(j)} \rangle | \psi^{(k)} \rangle.
\]
For the exchange term, though, we get
\[
-\langle \psi^{(j)} | \langle \alpha | \langle \psi^{(k)} | \langle \beta | \frac{1}{r_{12}} | \psi^{(j)} \rangle | \beta \rangle | \psi^{(j)} \rangle | \alpha \rangle \propto \langle \alpha | \beta \rangle = 0.
\]
This explains the different factors for the Coulomb and exchange terms.

Notice that both the Coulomb and exchange term depend on the basis of orbitals chosen for \( |cs \rangle \). That is, diagonalizing the Fock operator is not by itself a solution of the Fock problem if we use the wrong orbitals for the electron-electron term. We will see how to solve this problem iteratively in a following section using the so called self consistent field method.

We have seen that a basis of molecular orbitals which diagonalizes the Fock operator is an stationary point of the energy in the restricted space of closed-shell states. The corresponding eigenvalues \( \varepsilon_j \) are called orbital energies, and do not give directly the energy of \( |\Psi_S\rangle \). This is because the Hartree-Fock operator is not the initial Hamiltonian. They are related to ionization potentials and electron affinities by the so-called Koopmans’ Theorem. For the closed-shell case, though, the following relation for the ground state energy holds [Szabo Eq. 3.183]
\[
E_0 = \sum_a^{N/2} (h_{aa} + \varepsilon_a).
\]

2.3.4 The general Hartree-Fock equation

We obtained the Fock operator by finding stationary points of the energy
\[
\langle cs | H | cs \rangle.
\]
for the N-body electron problem in the restricted closed-shell case. It is possible to perform a similar analysis using a single Slater determinant with no further constrains. This can be done, for instance, using the variational principle [Szabo 3.2].

Let’s write the single Slater determinant as
\[
|\Psi_S\rangle = |\chi^{(j_1)} \cdots \chi^{(j_N)}\rangle.
\]
The Hartree-Fock equation for spin orbital \( \chi^{(k)} \) is
\[
\left[ h(x_1) + \sum_k J_k(x_1) - \sum_k K_k(x_1) \right] \chi^{(j)}(x_1) = \varepsilon_j \chi^{(j)}(x_1),
\]
where \( h \) is the single electron Hamiltonian, \( J_k \) is the Coulomb operator obtained from orbital \( \chi_k \) and \( K_k \) is the exchange term obtained from orbital \( \chi_k \).
2.3.5 Roothaan equations

Let’s write the MO of the restricted closed shell HF as LCAOs

\[ \psi_j = \sum_{\mu=1}^{K} C_{\mu j} \phi_\mu, \]

where \( K \) is the number of AO of the basis set. If we write this into the HF equation we get

\[ f(\mathbf{r}_1) \sum_{\mu} C_{\mu j} \phi_\mu(\mathbf{r}_1) = \varepsilon_j \sum_{\mu} C_{\mu j} \phi_\mu(\mathbf{r}_1) \]

Left multiplying by \( \phi_\nu^*(\mathbf{r}_1) \) and integrating yields a matrix equation:

\[ \sum_{\mu} C_{\mu j} \int d\mathbf{r}_1 \phi_\nu^*(\mathbf{r}_1) f(\mathbf{r}_1) \phi_\mu(\mathbf{r}_1) = \varepsilon_j \sum_{\mu} C_{\mu j} \int d\mathbf{r}_1 \phi_\nu^*(\mathbf{r}_1) \phi_\mu(\mathbf{r}_1) \]

This equation is simplified by introducing the following matrix elements:

\[ F_{\nu \mu} = \int d\mathbf{r}_1 \phi_\nu^*(\mathbf{r}_1) f(\mathbf{r}_1) \phi_\mu(\mathbf{r}_1) \]
\[ S_{\nu \mu} = \int d\mathbf{r}_1 \phi_\nu^*(\mathbf{r}_1) \phi_\mu(\mathbf{r}_1) \]

The Hartree-Fock-Roothaan equations can be written in matrix form as:

\[ \sum_{\mu} F_{\nu \mu} C_{\mu j} = \varepsilon_j \sum_{\mu} S_{\nu \mu} C_{\mu j} \]

Or in a more simple way as the Roothaan equations:

\[ FC = SC\varepsilon \]

One method to solve the HF problem in this formulation is:

1. We start with some choice of MOs as LCAO out of a convenient set of AOs.
2. We calculate the matrix \( F \) from the choice of the previous point.
3. We solve for the matrix \( C \) in the Roothaan equation
4. If the new matrix \( C \) is the same as the one in the previous step, the one used to calculate \( F \) (or close enough) we stop. Otherwise, we loop to point 2

To solve for \( C \) we can just calculate the eigenvalues of \( S^{-1}F \). In practice, this and other aspects or SCF are optimized. For instance, the Fock equation only guarantees that we are at a stationary point, not at a minimum. This
is also a problem that one needs to work around in practice. In essence, the Fock equation comes from a linear expansion of the energy, and one needs to check the second order expansion (the Hessian). There is also the question of convergence. The method presented can sometimes converge slowly and oscillate between solutions. Indeed, in theory HF is NP-complete (this is a very recent result [arXiv:0712.0483]), but this is more than anything an statement about worst case contrive examples. In any case, the convergence is improved by using the history of past solutions and proposing a new solution that averages over that history in a smart way. This method is called \textit{direct inversion in the iterative subspace} (DIIS). This does not affect the cost, and improves the convergence. In particular, it avoids oscillations.

Let's now look into the matrix $F$ to get some understanding of the cost of running HF. The main cost comes from the Coulomb and exchange terms, which give expressions like

$$
F_{\mu\nu} = \cdots \sum_{\lambda\sigma} \left(N/2 \sum_a C_{\lambda a} C_{\sigma a}^* \right) 2 \int dr_1 dr_2 \phi_{\mu}^*(r_1) \phi_{\sigma}^*(r_2) \frac{1}{r_{12}} \phi_{\nu}(r_1) \phi_{\lambda}(r_2) + \cdots \\
\equiv \cdots + \sum_{\lambda\sigma} P_{\lambda\sigma} \left( (\mu\nu|\sigma\lambda) - \frac{1}{2} (\mu\lambda|\sigma\nu) \right) + \cdots.
$$

We are again using the chemist notation for the integrals. The main observation is that we have to calculate the terms of the tensor of integrals $(\mu\nu|\sigma\lambda)$ where each index runs over each of the $K$ possible AO that we have previously chosen, in principle. That is, we have $K^4$ integrals. For a very modest $K = 100$, this gives, in principle, $10^8$ integrals. There are some methods to improve this scaling which are used for big systems, but we are not going to cover them here. Another point is that these integrals are easier if the AOs are Gaussian instead of Slater orbitals, and this is the reason why normally basis sets are made out of Gaussians. In theory, this integrals need to be calculated only at the initialization step, but they still dominate the cost. In fact, they are often too many to be stored in memory and so they are recalculated when needed: this is called direct HF theory.