Lucas Visscher: Post Hartree-Foch Correlation Methods

Outline

1) Post Hartree-Foch correlation methods   (Tuesday)
   - Hartree-Foch and Brillouin's Theorem
   - Configuration Interaction
   - Perturbation Theory
   - Multireference Methods

2) The coupled cluster approach   (Tuesday)
   - Size consistency, extensivity and intensive properties
   - Model wave function

3) Second quantization   (Tuesday)
   - Foch space
   - Creation and Annihilation operators
   - Representation of operators
   - Contraction and Wich's Theorem

4) Exercises (Wednesday first hour)

5) Coupled Cluster equations   (Wednesday)
   - Similarity transformed Hamiltonian
   - Energy and amplitude equations
   - Optimization methods
   - Molecular properties and Excited States

6) A few applications   (Wednesday)
   - Energy convergence
   - Excited states

1) Post Hartree-Fock correlation methods

We assume that molecular orbitals are expanded in a basis of M spinorbitals (or spinors) and that we want to describe a system containing N electrons. We do not consider spin explicitly, so making the equations equally valid for spin-restricted, spin-unrestricted and relativistic (spin-orbit coupled) Hartree-Fock procedures. Taking a single determinant as model wave function and solving the Hartree-Fock equations self consistently leads to a set of N occupied spinorbitals $\{\psi_i\}$ and a set of $M-N$ unoccupied or virtual spinorbitals $\{\psi_a\}$. We will use the convention that indices i, j, k and l denote spinors occupied in the Hartree-Fock determinant, indices a, b, c and d unoccupied spinors and p, q, r and s general spinors (either occupied or virtual).

The Hamiltonian is written in the generic form:

$$\hat{H} = \sum_{i=1}^{N} \hat{h}(i) + \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \hat{g}(i,j) \qquad (1)$$

where i denotes now the electron coordinate. Matrix elements (integrals) of the 1- and 2-particle parts of the Hamiltonian are denoted as

$$h_{pq} \text{ and } g_{pqrs} = \iint \psi_p^*(1) \psi_q(1) \hat{g}(1,2) \psi_r^*(2) \psi_s(2) d\tau_1 d\tau_2$$

The Hartree-Fock equations are given by the equation

$$\hat{f}(i) \psi_p(i) = \varepsilon_p \psi_p(i) \qquad (2)$$

with $$\hat{f}(i) = \hat{h}(i) + \sum_{j=1}^{N} \hat{J}_j(i) - \hat{K}_j(i) \qquad (3)$$

in which $\hat{J}_j(i)$ and $\hat{K}_j(i)$ are the Coulomb and exchange operators, respectively

$J_j(i)$ and $K_j(i)$ are local and non-local operators that depend on the occupied orbitals:

$$J_j(i)\psi_P(i) = \int \psi_j^*(j)\psi_j(j)g(i,j)d\tau_j \, \psi_P(i) \quad (4)$$

$$K_j(i)\psi_P(i) = \int \psi_j^*(j)\psi_P(j)g(i,j)d\tau_j \, \psi_j(i) \quad (5)$$

Note that the integration over τ_j (the spatial and spin coordinates of electron j) can be carried out without knowledge of ψ_P for the J-operator, whereas it requires knowledge of ψ_P for the K-operator.

At convergence of the SCF the orbitals are eigenfunctions and we have

$$f_{Pq} = \int \psi_P^*(i)\hat{f}(i)\psi_q(i)d\tau_i = \int \psi_P^*(i)\varepsilon_q\psi_q(i)d\tau_i$$

$$= \varepsilon_q \int \psi_P^*(i)\psi_q(i)d\tau_i = \varepsilon_q \delta_{Pq} = \varepsilon_P \quad (6)$$

Looking at the energy expression

$$E^{HF} = \sum_{i=1}^{N} h_{ii} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\left(g_{iijj} - g_{ijji}\right) \quad (7)$$

We see that we may equivalently write it as a sum over occupied orbitals plus a correction term which corrects for double counting of electron interactions

$$E^{HF} = \sum_{i=1}^{N}\varepsilon_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\left(g_{iijj} - g_{ijji}\right) \quad (8)$$

While the Hartree-Fock wave function and energy account for the fact that electrons with the same spin will never be at the same place, the motion of electrons with different spin is uncorrelated, leading to significant errors.

The oldest post Hartree-Fock correlation method is Configuration Interaction in which we write the wave function as a sum of determinants

$$\Psi^{CI} = c_0 \Phi_0 + \sum_I c_I \Phi_I \tag{9}$$

in which Φ_0 is the Hartree-Fock determinant

$$\Phi_0 = \| \psi_1(1)\, \psi_2(2) \,\text{--}\, \psi_N(N) \| \tag{10}$$

and Φ_I are determinants in which one or more occupied spinorbitals are replaced by virtual orbitals. This larger flexibility of the wave functions allows for correction of the Hartree-Fock error and is particularly important when electrons are strongly correlated, such as in a dissociating H_2 molecule in which each electron should end up at a different nucleus.

The CI coefficients are determined using the variational principle by diagonalizing the CI matrix:

$$\underline{\underline{H}}\,\underline{C} = E\,\underline{C} \tag{11}$$

matrix elements $H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle$ are found using the Slater-Condon rules. Here the sparsity of the matrix is apparent because the Hamiltonian can connect only determinants that differ by at most 2 replacements. Denoting determinants differing by Single, Double, Triple and Quadruple excitation symbolically as S, D, T, Q the following structure emerges:

	O	S	D	T	Q
O	E^{HF}	0	▨	0	0
S	0	▨	▨	▨	0
D	▨	▨	▨	▨	▨
T	0	▨	▨	▨	▨
Q	0	0	▨	▨	▨

The blocks OT, OQ, SQ are zero because determinants will differ by at least 3 replacements. The block OO consists of a single value which is identical to the Hartree Fock energy because $E^{HF} = \langle \underline{\Phi}_0 | \hat{H} | \underline{\Phi}_0 \rangle$

and the block OS is zero because of Brillouins theorem

$$\langle \underline{\Phi}_0 | \hat{H} | \underline{\Phi}_i^a \rangle = f_{ia} = 0 \qquad (12)$$

Here the notation Φ^a was introduced to denote a determinant in which occupied orbital i has replaced by virtual orbital a relative to the Hartree-Fock determinant.

From the structure of the CI matrix we can easily deduce that double excitations will give the largest contribution to the ground state wave function as these are the only ones that couple directly to the Hartree-Fock determinant. The CID method is therefore the simplest method that leads to an improved ground state energy. CIS is also used, but for a different purpose: the singly excited states provide a first approximation to excited states and provides thereby access to spectroscopic transitions.

The contributions of different excitation levels become also apparent if we consider perturbation theory. The most popular variant is Møller-Plesset perturbation theory in which the sum of Fock operators functions as zeroth order Hamiltonian:

$$\hat{H}^{(0)} = \sum_{i=1}^{N} \hat{f}(i) \qquad (13)$$

this then implies a perturbation operator $\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)}$

$$\hat{H}^{(1)} = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} \hat{g}(i,j) - \sum_{i=1}^{N}\sum_{j=1}^{N} \hat{J}_j(i) - \hat{K}_j(i) \qquad (14)$$

this operator is often called the "fluctuation" operator as it describes the deviation of the true Hamiltonian from that of a mean-field Hamiltonian.

The zeroth order energy is calculated as

$$\langle \bar{\Phi}_0 | H^{(0)} | \bar{\Phi}_0 \rangle = \langle \bar{\Phi}_0 | \sum_i^N \hat{f}(i) | \bar{\Phi}_0 \rangle = \langle \bar{\Phi}_0 | \sum_{i=1}^N \varepsilon_i | \bar{\Phi}_0 \rangle$$

$$= \sum_i \varepsilon_i \qquad (15)$$

where we used the fact that the Hartree-Fock orbitals are eigenfunctions of the Fock operator.

The first order energy correction is

$$\langle \bar{\Phi}_0 | H^{(1)} | \bar{\Phi}_0 \rangle = \frac{1}{2} \langle \bar{\Phi}_0 | \sum_{i=1}^N \sum_{j\neq i}^N \hat{g}(i,j) | \bar{\Phi}_0 \rangle - \langle \bar{\Phi}_0 | \sum_{ij} \hat{J}_j^{(i)} \hat{K}_j(i) | \bar{\Phi}_0 \rangle$$

$$= -\frac{1}{2} \sum_{i=1}^N \sum_{j\neq i}^N \left(g_{iijj} - g_{ijji} \right) \qquad (16)$$

$$(17)$$

so that $E^{(0)} + E^{(1)} = E^{MF}$

Improvement relative to Hartree-Fock comes at second order

$$E^{(2)} = \sum_{I \neq 0} \frac{\langle \bar{\Phi}_0 | \hat{H}^{(1)} | \bar{\Phi}_I \rangle \langle \bar{\Phi}_I | \hat{H}^{(1)} | \bar{\Phi}_0 \rangle}{E_0^{(0)} - E_I^{(0)}} \qquad (18)$$

for the ~~denominator~~ numerator we use the original definition of $\hat{H}^{(1)}$ and the fact that all determinants are eigenfunctions of $H^{(0)}$:

$$\langle \bar{\Phi}_I | \hat{H}^{(1)} | \bar{\Phi}_0 \rangle = \langle \bar{\Phi}_I | \hat{H} | \bar{\Phi}_0 \rangle - \langle \bar{\Phi}_I | \hat{H}^{(0)} | \bar{\Phi}_0 \rangle$$

$$= \langle \bar{\Phi}_I | \hat{H} | \bar{\Phi}_0 \rangle - E_0^{(0)} \langle \bar{\Phi}_I | \bar{\Phi}_0 \rangle$$

$$= \langle \bar{\Phi}_I | \hat{H} | \bar{\Phi}_0 \rangle \qquad (19)$$

From the Slater-Condon rules and the Brillouin Theorem we see that only double excitations will contribute:

$$\langle \bar{\Phi}_{ij}^{ab} | \hat{H} | \bar{\Phi}_0 \rangle = g_{aibj} - g_{ajbi} \qquad (20)$$

The zeroth order energy of $\bar{\Phi}_{ij}^{ab}$ differs from $E_0^{(0)}$ by two orbital energies so that

$$E_0^{(0)} - E_I^{(0)} = \varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b \qquad (21)$$

Inserting equations (20) and (21) in (18) gives then

$$E^{(2)} = -\sum_{\substack{ij \\ ab}} \frac{|g_{aibj} - g_{ajbi}|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \qquad (22)$$

in which we defined the denominator as positive, and used the hermiticity of the Hamiltonian: $\langle \bar{\Phi}_{ij}^{ab} | \hat{H} | \bar{\Phi}_0 \rangle = \langle \bar{\Phi}_0 | \hat{H} | \bar{\Phi}_{-ij}^{ab} \rangle^*$ to get a sum in which each element is positive and the correlation energy is strictly negative.

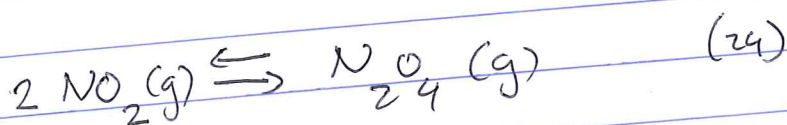The first-order wave function is given by

$$\bar{\Psi}^{(1)} = -\sum_{ij} \sum_{ab} \frac{(g_{aibj} - g_{ajbi})}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} | \bar{\Phi}_{ij}^{ab} \rangle \qquad (23)$$

and we see that the method will only work reliably if the HOMO-LUMO gap $\min\limits_{\forall a, \forall i} (\varepsilon_a - \varepsilon_i)$ is sufficiently large.

If the gap is small, one may expect large individual contribution to the wave function and energy correction and slow convergence of the perturbation series. In such cases one should resort to (quasi) degenerate perturbation theory and multideterminant reference functions.

2) Coupled Cluster

While CI theory is still used for small systems, it has become apparent that the method is not suited to describe larger molecules. The reason is that only the so-called full CI method is size-consistent and size-extensive and while all truncated methods fail to describe systems of different sizes in a consistent manner. Let's consider the reaction

$$2 NO_2 (g) \rightleftharpoons N_2O_4 (g) \qquad (24)$$

as an example. If we want to calculate the thermodynamical equilibrium constant we need to determine the electronic energies of the NO_2 and N_2O_4 molecules. The straightforward way to do this is to calculate the energy of NO_2 in one calculation, multiply this by 2 and compare it to the energy of N_2O_4, the molecule on the right-hand side. A less efficient way is to calculate the left hand side energy in a calculation of two NO_2 molecules simultaneously, placing them so far apart that they will not interact. It would be desirable that both ways of calculating the left hand energy yield the same result, i.e. that the energy is <u>size-extensive</u>

$$E(n A^{non-interacting}) = n E (A) \qquad (25)$$

in which A can be any molecule. Carrying out the CISD ~~method~~ calculation for both cases does, however, give a difference of more than 50 kcal/mol for reaction (24) with the energy calculated as $2 E (NO_2)$ being much lower. The reason is that this calculation does effectively include quadruple excitations because it considers double excitations on one NO_2 molecule as well as double excitations on the other molecule. These excitations are absent in the calculation in which the NO_2 molecules are placed at large distance but calculated simultaneously. Should we then simply use the $2* E (NO_2)$ energy?

No, this is not so simple because also in the N_2O_4 case, in which the molecules are interacting, such excitations are missing. In fact, the energy calculated in this way is strongly repulsive which is contradiction to the experimental observation of an exothermic reaction.

What is needed in chemistry are size-consistent methods in which a whole potential energy surface can be described in a consistent manner and in which it is possible to perform separate calculations if molecules are so far apart that they can be considered non-interacting.

A necessary, but not sufficient, condition for size-consistency is size-extensivity, a mathematical proof that for a certain method equation (25) will hold. This may be complemented by a size-intensivity criterium for molecular properties that should not depend on system size. An example is the first electronic excitation energy which should be the same, regardless of whether you have one or a multitude of non-interacting molecules.

Before looking at the coupled cluster approach which guarantees size-extensivity, we first note that for non-interacting systems the wave function can be written as an (anti-symmetrized) product of the individual wave functions of the ~~separate~~ separate systems.

$$\Psi_{nA}^{non-interacting} = \Psi_{A_1} \otimes \Psi_{A_2} \otimes \cdots \otimes \Psi_{A_N} \qquad (26)$$

It is easily verified that the Hartree-Fock model fulfills (26), whereas the CISD model does not. We therefore need a better ansatz for the correlated wave function.

A 4-electron wave function

As a final example of a wave function in first quantization we will follow Crawford and Schaeffer (Reviews in Computational Chemistry) and write out a correct wave function for 4 electrons. This example will serve to introduce the second quantization technique that makes electron correlation theory much simpler.

Starting point is a Hartree-Fock determinant consisting of 4 spinors that are labeled as i, j, k and ℓ (NB these are not summation indices here)

$$\overline{\Phi}_0 = \| \psi_i(1) \psi_j(2) \psi_k(3) \psi_\ell(4) \| \qquad (27)$$

and introduce a cluster function

$$f_{ij}(m, n) = \sum_{a > b} t_{ij}^{ab} \, \psi_a(m) \psi_b(n) \qquad (28)$$

which will correlate an arbitrary pair of electrons m and n that occupy the orbitals i and j. The sum over a and b runs over all virtual orbitals and we restrict the two summations indices to avoid double counting the same excitation $\left(\psi_{ij}^{ab} = -\psi_{ij}^{ba} \right)$. The cluster amplitudes t_{ij}^{ab} give the weight of this particular excitation and form the free variables in the wave function.

Correlating all possible orbital pairs yields the following wave function

$$\overline{\Phi} = | \psi_i \psi_j \psi_k \psi_\ell \rangle + | f_{ij} \psi_k \psi_\ell \rangle - | f_{ik} \psi_j \psi_\ell \rangle + | f_{i\ell} \psi_j \psi_k \rangle$$

$$+ | \psi_i f_{jk} \psi_\ell \rangle - | \psi_i f_{j\ell} \psi_k \rangle + | \psi_i \psi_j f_{k\ell} \rangle + | f_{ij} f_{k\ell} \rangle$$

$$- | f_{ik} f_{j\ell} \rangle + | f_{i\ell} f_{jk} \rangle \qquad (29)$$

in which we recognize the Hartree-Fock determinant as the first term

and see both doubly and quadruply excited determinants. These will be the most important, but it is possible to further improve the wave function by defining also three-orbital cluster functions f_{ijk} and extend the wave function with the terms

$$|f_{ijk}\psi_\ell\rangle - |f_{ije}\psi_u\rangle + |f_{ike}\psi_j\rangle + |\psi_i f_{jke}\rangle \tag{30}$$

(note here and in the previous the signs arising from permuting the orbitals to make the spinors to be correlated adjacent).
More important is, however, to also include single orbital excitations

$$f_i(m) = \sum_a t_i^a \,\psi_a(m)$$

written out in full the coupled cluster singles and doubles wave function then becomes

$$\Psi^{CCSD} = |\psi_i\psi_j\psi_u\psi_\ell\rangle + |f_i\psi_j\psi_u\psi_\ell\rangle + |\psi_i f_j\psi_u\psi_\ell\rangle + |\psi_i\psi_j f_u\psi_\ell\rangle$$

$$+|\psi_i\psi_j\psi_u f_\ell\rangle + |f_i f_j\psi_u\psi_\ell\rangle + |f_i\psi_j f_u\psi_\ell\rangle + |\psi_i f_j f_u\psi_\ell\rangle$$

$$+|\psi_i f_j\psi_u f_\ell\rangle + |\psi_i\psi_j f_u f_\ell\rangle + |f_i f_j f_u\psi_\ell\rangle + |f_i f_j\psi_u f_\ell\rangle +$$

$$|f_i\psi_j f_u f_\ell\rangle + |\psi_i f_j f_u f_\ell\rangle + |f_{ij}\psi_u\psi_\ell\rangle - |f_{iu}\psi_j\psi_\ell\rangle$$

$$+|f_{i\ell}\psi_j\psi_u\rangle + |\psi_i f_{ju}\psi_\ell\rangle - |\psi_i f_{j\ell}\psi_u\rangle + |\psi_i\psi_j f_{u\ell}\rangle$$

$$+|f_{ij}f_{u\ell}\rangle \mp |f_{iu}f_{j\ell}\rangle + |f_{i\ell}f_{ju}\rangle + |f_i f_j f_u f_\ell\rangle$$

$$+|f_{ij}f_{uk}\psi_\ell\rangle + |f_{ij}\psi_u f_\ell\rangle + |f_{ij}f_u f_\ell\rangle - |f_{iu}f_j\psi_\ell\rangle$$

$$-|f_{iu}f_j f_\ell\rangle + |f_{i\ell}f_j\psi_u\rangle + |f_{i\ell}f_j f_u\rangle + |f_i f_{ju}\psi_\ell\rangle + |\psi_i f_{ju}f_\ell\rangle$$

$$+|f_i f_{ju}f_\ell\rangle - |f_i f_{j\ell}\psi_u\rangle - |\psi_i f_{j\ell}f_u\rangle - |f_i f_{j\ell}f_u\rangle + |\psi_i f_j f_{u\ell}\rangle \tag{31}$$

$$+|f_i f_j f_{u\ell}\rangle$$

It is clear that even this simple case already yields very long expressions so that we need a better notation. Let's consider one of these terms:

$$| \Psi_{ij}^{ab} \Psi_k \Psi_\ell \rangle = \sum_{a>b} \sum_c t_{ij}^{ab} t_\ell^c | \Psi_a \Psi_b \Psi_k \Psi_c \rangle \qquad (32)$$

and note that this corresponds to determinants that are triply excited relative to the Hartree-Fock reference. This means that while the ~~coupled~~ CCSD wave function contains just as many independent variables as the CISD wave function, it also includes triply and quadruply excited determinants.

To arrive at a more compact notation we define a creation operator

$$a_p^+ | \Psi_q \cdots \Psi_s \rangle = | \Psi_p \Psi_q \cdots \Psi_s \rangle \qquad (33)$$

and annihilation operator

$$a_p | \Psi_p \Psi_q \cdots \Psi_s \rangle = | \Psi_q \cdots \Psi_s \rangle \qquad (34)$$

to extend and reduce the number of electrons in a determinant. Excitations are now represented by the operator $a_a^+ a_i$ which removes an electron in spinor i and puts it back in spinor a.

With these operators the lengthy expression (31) may be written as

$$e^{\hat{T}_1 + \hat{T}_2} \;\rule{3cm}{0pt}\; | \Phi_o \rangle$$

by defining $\hat{T}_1 = \sum_{i,a} t_i^a a_a^+ a_i$ and $\hat{T}_2 = \sum_{i<j} \sum_{a<b} t_{ij}^{ab} a_a^+ a_b^+ a_j a_i \quad (35)$

3) Second Quantization

Before continuing with coupled cluster theory we will now first consider how wave functions and operators can be expressed in second quantization.

We start with wave functions and define a **Fock space** consisting of all determinants that can be formed by occupying $0, 1, 2, \dots M$ of our n spinorbitals. The **Hilbert space** of N electron determinants in a subspace of this larger space.

~~waxx~~ Determinants are defined by their occupation number vector in which is specified whether spinor p is occupied: $k_p = 1$ or unoccupied: $k_p = 0$

$$|\vec{k}\rangle = |k_1, k_2, \dots, k_M\rangle \qquad (36)$$

because determinants that differ in occupation are orthogonal we ~~kxxx~~ define

$$\langle \vec{k} | \vec{\ell} \rangle = \delta_{\vec{k}, \vec{\ell}} = \prod_{p=1}^{M} \delta_{k_p \ell_p} \qquad (37)$$

Note that this also encompasses overlaps between determinants with a different number of electrons (these are zero). The state with no spinors occupied is called the vacuum state

$$|vac\rangle = |0_1, 0_2 \dots, 0_n\rangle \qquad (38)$$

Operating with a creation operator yields a one electron state

$$a_p^{+} |vac\rangle = |0_1, \dots, 1_p, \dots 0_n\rangle \qquad (39)$$

while operating with an annihilation operator gives zero

$$a_p |vac\rangle = 0 \qquad (40)$$

if a_p^+ operates on a state that has already occupied spinors its action depends on the occupation numbers

$$a_p^+ |k_1, k_2, \dots, 0_p, \dots, k_m\rangle = \Gamma_p^{\vec{k}} |k_1, k_2, \dots, 1_p, \dots, k_m\rangle$$

$$a_p^+ |k_1, k_2, \dots, 1_p, \dots, k_m\rangle = 0$$

$$\text{with } \Gamma_p^{\vec{k}} = \prod_{q=1}^{p-1} (-1)^{k_q} \qquad (41)$$

The sign factor $\Gamma_p^{\vec{k}}$ preserves the antisymmetry of the wave functions, the fact that a_p^+ gives zero when spinor p is already occupied is consistent with the Pauli principle.

The effect of an annihilation operator is defined as

$$a_p |\vec{k}\rangle = 0 \qquad \text{if } k_p = 0$$

$$a_p |k_1, k_2, \dots, 1_p, \dots, k_m\rangle = \Gamma_p^{\vec{k}} |k_1, k_2, \dots, 0_p, \dots, k_m\rangle \qquad (4$$

From these properties follow the important anti commutation rules

$$[a_p^+, a_q^+]_+ = a_p^+ a_q^+ + a_q^+ a_p^+ = 0$$

$$[a_p, a_q]_+ = a_p a_q + a_q a_p = 0 \qquad (43)$$

$$[a_p^+, a_q]_+ = \delta_{pq}$$

It is easy to verify that the number operator

$$\hat{N} = \sum_{P=1}^{M} a_P^+ a_P \qquad (44)$$

can be used to count the number of occupied spinors in an occupation vector.

As already mentioned before we can furthermore define excitation operators $a_a^+ a_i$ to replace spinors.

With these ingredients we are ready to write compact representations of complicated wave functions. Before doing so we, however, also need to consider the representation of the Hamiltonian.

This takes the generic form

$$\hat{H} = \sum_{P,q}^{M} h_{Pq} a_P^+ a_q + \frac{1}{2} \sum_{P,q,r,s}^{M} g_{pqrs} a_P^+ a_r^+ a_s a_q \qquad (45)$$

in which we recognize the matrix elements of the first quantized operators $\hat{h}$ and $\hat{g}$. Note that all reference to the actual number of electrons has disappeared from the Hamiltonian, this information is carried by the occupation vectors that represent the wave function. Note also the peculiar order of the indices p, q, r and s in the second term, this is a consequence of writing the integrals in "Mulliken" ordering in which functions are grouped on electron coordinate rather than on bra/ket character.

As an illustration of the above, we shall now calculate the Hartree–Fock energy of the 4-electron system considered in equation (27) and verify that this is consistent with the energy expression found earlier. For simplicity we take $i=1$, $j=2$, $h=3$, and $l=4$ so that the determinant becomes

$$\overline{\Phi}_0 = || \psi_1 \psi_2 \psi_3 \psi_4 || \qquad (44)$$

This becomes in second quantization:

$$|HF\rangle = a_1^+ a_2^+ a_3^+ a_4^+ |vac\rangle \qquad (45)$$

and $\langle HF| = (|HF\rangle)^+ = \langle vac| a_4 a_3 a_2 a_1 \qquad (46)$

This wave function is obviously normalized because

$$\langle \underset{1234\,567\quad n}{\underbrace{1111\,000\cdots 0}} | \underset{1234\,967\quad n}{\underbrace{1111\,000\cdots 0}} \rangle = 1 \qquad (47)$$

but this may also be demonstrated using the anticommutation relations to bring the operator string to normal order: the ordering in which all annihilation operators stand to the right of the creation operators

$$\langle vac| a_4 a_3 a_2 a_1 a_1^+ a_2^+ a_3^+ a_4^+ \rangle = -\langle vac| a_4 a_3 a_2 a_1 a_1^+ a_2^+ a_3^+ a_4^+ |vac\rangle$$

$$+ \langle vac| a_4 a_3 a_2 a_2^+ a_3^+ a_4^+ |vac\rangle \qquad (48)$$

$$= -\langle vac| a_1 a_4 a_3 a_2 a_2^+ a_3^+ a_4^+ a_1 |vac\rangle - \langle vac| a_4 a_3 a_2 a_2^+ a_3^+ a_4^+ |vac\rangle$$

$$+ \langle vac| a_4 a_3 a_3^+ a_4^+ |vac\rangle \qquad (49)$$

in which we obtained (48) from the application of the anticommutation rule (43). The first term of (49) is zero because a_1 acts on the vacuum state. The same holds for the second term which after we have brought a_2 completely to the right. This leaves only

$$\langle vac | a_4 a_3 a_3^\dagger a_4^\dagger | vac \rangle = -\langle vac | a_4 a_3^\dagger a_3 a_4^\dagger \rangle$$

$$+ \langle vac | a_4 a_4^\dagger | vac \rangle = 0 + 1 = 1 \qquad (50)$$

For the Hamiltonian we can apply a similar technique, except that now we have now free indices. Let's consider the one-body term first

$$\langle HF | \sum_{p,q} h_{pq} a_p^\dagger a_q | HF \rangle = \sum_{p,q} \langle vac | a_4 a_3 a_2 a_1 a_p^\dagger a_q a_1^\dagger a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle$$
$$\ast h_{pq}$$

$$= \sum_{p,q} -h_{pq} \langle vac | a_4 a_3 a_2 a_p^\dagger a_1 a_q a_1^\dagger a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle$$

$$+ \delta_{p1} h_{pq} \langle vac | a_4 a_3 a_2 a_q a_1^\dagger a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle$$

$$= \sum_{p,q} +h_{pq} \langle vac | a_4 a_3 a_2 a_p^\dagger a_1 a_1^\dagger a_q a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle$$

$$- h_{pq} \delta_{q1} \langle vac | a_4 a_3 a_2 a_p^\dagger a_1 a_1^\dagger a_q a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle$$

$$- h_{pq} \delta_{p1} \langle vac | a_4 a_3 a_2 a_1^\dagger a_q a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle$$

$$+ h_{pq} \delta_{p1} \delta_{q1} \langle vac | a_4 a_3 a_2 a_2^\dagger a_3^\dagger a_4^\dagger | vac \rangle \qquad (51)$$

The last term gives h_{11}, one of the four terms in the sum over

matrix elements of $\hat{h}$. The other three terms emerge from repeated anticommutation of the first term. The second and third term are zero because we can ~~take~~ either commute a_q directly to the right (first term, $q=1$, so not equal to 2, 3 and 4)

or commute $a_1^\dagger$ to the left and use $\langle vac|a_1^\dagger = (a_1|vac\rangle)^\dagger = 0$.

These manipulations are an example of Wichs theorem that states that for evaluation of the expectation value of an operator string over the vacuum state one only needs to consider fully contracted terms.

For the matrix elements of $\hat{g}$ the procedure is similar, but more tedious, and I will just give the result

$$\langle HF| \hat{H} |HF\rangle = \sum_{i=1}^{4} h_{ii} + \frac{1}{2} \sum_{i=1}^{4} \sum_{j=1}^{4} \left(g_{iijj} - g_{ijji} \right) \qquad (52)$$

which is indeed identical to equation (7) if we realize that the additional term with $i=j$: $g_{iijj} - g_{ijji} = g_{iiii} - g_{iiii} = 0$.

Using these simple manipulations, which can also be carried out by a computer, it is possible to derive explicit expressions for much more complicated wave function than the one above. This is used to derive explicit expression for coupled cluster energies and amplitudes as we shall see next.

5) Coupled Cluster Equations

The coupled cluster wave function is defined by the action of the cluster operators on a reference determinant. In principle this can be any determinant but it is convenient to take the Hartree-Fock determinant written in terms of canonical orbitals. We already saw an explicit form of the wave function in equation (31) and noted that this could be shortened to equation (35) in chapter 2. We will now take a more general approach and define a state $|\mu\rangle$ that is generated from the Hartree-Fock determinant by the action of an excitation operator

$$|\mu\rangle = \hat{\tau}_{\mu} |HF\rangle \qquad (53)$$

operating with two different excitation operators will in general give a higher excited state

$$|\mu\nu\rangle = \hat{\tau}_{\mu}\hat{\tau}_{\nu} |HF\rangle \qquad (54)$$

unless one more annihilation, or one more creation, operators that are contained in $\hat{\tau}_{\mu}$ and $\hat{\tau}_{\nu}$ are identical. This is always the case if $\mu = \nu$ so

$$\hat{\tau}_{\mu}^{2} |HF\rangle = 0 \qquad (55)$$

The coupled cluster wave function takes the general form

$$|cc\rangle = e^{\hat{T}} |HF\rangle \qquad (56)$$

with $\quad \hat{T} = \sum_{\mu} t_{\mu} \hat{\tau}_{\mu} \qquad (57)$

Note that (35) is a special case of (57).

In order to determine the energy and amplitudes, we will carry out the similarity transformation

$$\hat{H}^T = e^{-\hat{T}} \hat{H} e^{\hat{T}} \qquad (58)$$

and use the Schrödinger equation

$$\hat{H} |cc\rangle = E^{cc} |cc\rangle$$

$$H e^{\hat{T}} |HF\rangle = E^{cc} e^{\hat{T}} |HF\rangle \qquad (59)$$

$$e^{-\hat{T}} H e^{\hat{T}} |HF\rangle = E^{cc} |HF\rangle$$

to obtain equations for the energy and amplitudes. The energy is obtained by left projecting with $\langle HF|$

$$E^{cc} = \langle HF| e^{-\hat{T}} H e^{\hat{T}} |HF\rangle = \langle HF| \hat{H}^T |HF\rangle \qquad (60)$$

while equations for the amplitudes are obtained by left projecting with the states μ:

$$\langle \mu| e^{-\hat{T}} H e^{\hat{T}} |HF\rangle = \langle \mu| E^{cc} |HF\rangle = 0 \qquad (61)$$

Let us now consider the explicit expression for the energy. First we note that

$$\langle HF| e^{-\hat{T}} = \left(e^{-\hat{T}^\dagger} |HF\rangle \right)^\dagger = 0 \qquad (62)$$

because $\hat{T}^\dagger$ is a de-excitation operator that produces zero when

applied to the reference wave function. We then expand $e^{\hat{T}}$ in a Taylor series as

$$E^{cc} = \langle HF | \hat{H} e^{\hat{T}} | HF \rangle \qquad (63)$$

$$= \langle HF | \hat{H}(1 + \hat{T} + \frac{1}{2}\hat{T}^2 + \cdots) | HF \rangle \qquad (64)$$

$$= \langle HF | \hat{H}(1 + \hat{T} + \frac{1}{2}\hat{T}^2) | HF \rangle \qquad (65)$$

where we have used that powers higher than 2 do not contribute because these yield at least triply excited determinants. which ~~For these are~~ can not be connected via the Hamiltonian to the Hartree-Fock state (see the structure of the CI matrix on page 4). For the same reason there is no direct contribution of $\hat{\tau}_\mu$ excitation operators that produce triple or higher excitations. These may of course indirectly influence the energy by modifying the values of the ~~sing~~ ~~excit~~ double excitation amplitudes. The single excitation amplitudes contribute to the energy via the $\hat{T}^2$ term, their contribution via $\langle HF | H\hat{T} | HF \rangle$ is zero due to Brillouins theorem.

Working out the anticommutators gives

$$E^{cc} = \sum_{i<j} \sum_{a<b} \left(g_{iajb} - g_{ibja} \right) \left(t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a \right)$$

$$(66)$$

an expression which is remarkably similar to the MP2 energy expression (23). The two expressions give the same value if the coupled cluster amplitudes are initialized with the first order perturbation expressions $t_i^a = 0$ $\forall i, a$ and

$$t_{ij}^{ab} = \frac{(g_{aibj} - g_{ajbi})}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}.$$

We see that for the energy we only need to know the two-electron integrals and the amplitudes. How do we determine the latter?

For this we use the Baker-Campbell-Hausdorff expansion that was proven in one of the exercises.

We have $\hat{H}^T = e^{-\hat{T}} \hat{H} e^{\hat{T}}$

$$= \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{6}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}]$$

$$+ \frac{1}{24}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}]$$

The expansion terminates after the 5th term (the one with a fourfold nested commutator) because each commutator reduces the number of free indices in the Hamiltonian by one and the Hamiltonian has at most 4 indices (in the two-electron operator). This is an important result which does not depend on the excitation level of the $\hat{T}_\mu$ operators but solely on the nature of the Hamiltonian. Also for CCSDTQQH methods it suffices to evaluate just 4 commutators.

Of course the resulting equations may still be lengthy and I will not write them out in full here. Explicit formula's for the CCSD model are easily found in the literature, for higher models developers have often used symbolic algebra programs to directly generate computer code without having to write out all equations manually first and subsequently program them.

Reference

Most of the material here can be found in the excellent
textbook

Molecular Electronic Structure Theory
by Helgaker, Jørgensen and Olsen
Wiley.

of which I have also taken some exercises.

Another excellent introduction that also discusses diagremmatic
techniques is provided by Crawford and Schaeffer

TD Crawford, HF ~~Schaeffer~~ Schaefer

An Introduction to coupled cluster theory for Computatinal Chemists
Reviews in Computational Chemistry, volume 14

the cluster functions f_{ij} are discussed in more detail there.