

Coupled Clusters

THE COUPLED CLUSTER METHOD.§1 Second quantisation

This is a method which takes care of the antisymmetry of electronic wave functions in a more elegant way than the Slater method does. It is very useful for explaining the Coupled Cluster and related methods. In the second quantisation formalism the expansion of a determinant into $N!$ terms is bypassed. Instead, the antisymmetry is taken care of by defining creation operators \mathbf{a}_i^\dagger and annihilation operators \mathbf{a}_i (for MSO ψ_i) and defining an antisymmetric function by :

$$|\psi_1 \dots \psi_I \dots \psi_J \dots \psi_N| \equiv |1 \dots I \dots J \dots N\rangle \equiv \mathbf{a}_1^\dagger \dots \mathbf{a}_I^\dagger \dots \mathbf{a}_J^\dagger \dots \mathbf{a}_N^\dagger |vac\rangle \quad (2.1)$$

where $|vac\rangle$ is the "true vacuum" state, i.e. a state without any electrons. Each creation operator \mathbf{a}_i^\dagger puts (an electron into) MSO ψ_i in the first position of the state on which it acts. The annihilation operator \mathbf{a}_i removes (an electron in) MSO ψ_i if it is in the first position of the state on which it acts, otherwise the orbital has to be moved to the front first, which results in a factor $(-1)^{p(I)-1}$, where $p(I)$ indicates the position of ψ_i in the state (determinant).

The antisymmetry condition is satisfied by imposing the following anticommutation rule :

$$[\mathbf{a}_i^\dagger, \mathbf{a}_j^\dagger]_+ \equiv \mathbf{a}_i^\dagger \mathbf{a}_j^\dagger + \mathbf{a}_j^\dagger \mathbf{a}_i^\dagger = 0 \quad (2.2)$$

The anticommutation rule for the annihilation operators is obtained by taking the hermitian conjugate of Eq (2.2) :

$$[\mathbf{a}_i, \mathbf{a}_j]_+ \equiv \mathbf{a}_i \mathbf{a}_j + \mathbf{a}_j \mathbf{a}_i = 0 \quad (2.3)$$

We see that the sign changes if the operators are exchanged. This is the reason for using the designation anticommutation instead of commutation. From Eq (2.2) it follows that :

$$|\psi_i \psi_j\rangle \equiv \mathbf{a}_i^\dagger \mathbf{a}_j^\dagger |vac\rangle \equiv -\mathbf{a}_j^\dagger \mathbf{a}_i^\dagger |vac\rangle \equiv -|\psi_j \psi_i\rangle \quad (2.4)$$

and Eq (2.4) implies that:

$$\mathbf{a}_i^\dagger \mathbf{a}_i^\dagger = \mathbf{a}_i \mathbf{a}_i = 0 \quad (2.5)$$

In the following it will be assumed throughout that the MSOs are orthonormal. Using Eqs (2.2-4) it may then be shown that :

$$\mathbf{a}_i \mathbf{a}_j^\dagger + \mathbf{a}_j^\dagger \mathbf{a}_i = \delta_{ij} \quad (2.6)$$

Exercise 2.1

Derive Eq (2.6) from Eqs (2.2-4) by checking Eq (2.6) for all possibilities for the occupations of ψ_i and ψ_j , that is, for the 4 cases ψ_i occupied or empty and ψ_j occupied or empty.

Why is the orthonormality of the orbitals needed ?

Coupled Clusters

The second quantisation representation of an operator may be constructed by imposing the condition that the results will be the same as in the original formalism. The result is as follows.

1-electron operators :

$$\mathbf{O} = \sum_J \sum_I \mathbf{a}_I^\dagger \mathbf{O}_{IJ} \mathbf{a}_J$$

with

$$O_{IJ} = \langle \psi_I | \mathbf{O} | \psi_J \rangle$$

2-electron operators :

$$\mathbf{O} = \sum_{KL} \sum_{IJ} \mathbf{a}_K^\dagger \mathbf{a}_L^\dagger \mathbf{O}_{KL,IJ} \mathbf{a}_J \mathbf{a}_I$$

with

$$O_{KL,IJ} = \langle | \dots \psi_K \dots \psi_L \dots | | \mathbf{O} | | \dots \psi_I \dots \psi_J \dots | \rangle$$

For the unit operator in the MSO space this yields (m is the number of MSOs) :

$$\mathbf{1} = \sum_I^m \mathbf{a}_I^\dagger \mathbf{a}_I \quad (2.9)$$

This is the "resolution of the identity" in second quantisation. Eq (2.8) is equivalent to :

$$\mathbf{1} = \sum_I^m | \psi_I \rangle \langle \psi_I | \quad (2.10)$$

in the usual formalism.

In the following the HF determinant is designated by $| 0 \rangle$, the so-called (Fermi) vacuum state. The second quantisation operators may be used to generate excited determinants from this state by defining the excitation operators :

$$\tau_I^A \equiv \mathbf{a}_A^\dagger \mathbf{a}_I \quad (2.11)$$

for a single excitation. This operator replaces ψ_I by ψ_A in the same position.

Exercise 2.2

Show that applying the single excitation operator does not involve a change of sign.

Higher excitations are obtained by repeatedly applying these operators :

$$\tau_{IJ}^{AB} \equiv \tau_J^B \tau_I^A \quad (2.12)$$

Because of Eqs (2.2) and (2.3) these are antisymmetric in the indices :

$$\tau_{IJ}^{AB} = -\tau_{JI}^{BA} = -\tau_{IJ}^{BA} = \tau_{JI}^{AB} \quad (2.13)$$

Therefore we also have :

$$\tau_{IJ}^{AA} = \tau_{II}^{AB} = 0 \quad (2.14)$$

Coupled Clusters

Finally commutation relations for the excitation operators will be needed. By using Eqs (2.2), (2.3) and (2.6) the following result is found :

$$[\tau_1^A, \tau_j^B] = \delta_{jB} \tau_j^A - \delta_{jA} \tau_1^B \quad (2.15)$$

Exercise 2.3

Derive Eq (2.15) from Eqs (2.2), (2.3) and (2.6).

§2 The cluster expansion of the wave function.

We start with a HF single determinant function as the reference function. The Coupled Cluster function is then defined by :

$$|\Psi_{CC}\rangle = \exp(\mathbf{T}) |0\rangle$$

with (2.16)

$$\mathbf{T} = \sum_m \mathbf{T}_m$$

In Eq (2.16) \mathbf{T}_m contains all m-fold excitation operators (products of m single excitation operators), such as :

$$\mathbf{T}_1 = \sum_I^{\text{occ}} \sum_A^{\text{virt}} s_I^A \tau_I^A \quad (2.17)$$

$$\mathbf{T}_2 = \sum_{IJ}^{\text{occ}} \sum_{AB}^{\text{virt}} d_{IJ}^{AB} \tau_{IJ}^{AB} \quad (2.18)$$

Here s_I^A and d_{IJ}^{AB} are the coefficients to be optimized. All operators in Eq (2.16) commute, since no lower index coincides with any upper index.

The exponential operator in Eq (2.16) is defined by the power expansion :

$$\exp(\mathbf{T}) = 1 + \mathbf{T} + \frac{\mathbf{T}^2}{2} + \frac{\mathbf{T}^3}{6} + \frac{\mathbf{T}^4}{24} + \dots \quad (2.19)$$

The first term in the expansion shows that the CC function satisfies the intermediate normalisation condition.

If only single and double excitations are included, we have :

$$\begin{aligned} \exp(\mathbf{T}) = \exp(\mathbf{T}_1 + \mathbf{T}_2) &= 1 + (\mathbf{T}_1 + \mathbf{T}_2) + \frac{1}{2}(\mathbf{T}_1^2 + 2\mathbf{T}_1\mathbf{T}_2 + \mathbf{T}_2^2) \\ &+ \frac{1}{6}(\mathbf{T}_1^3 + 3\mathbf{T}_1^2\mathbf{T}_2) + \frac{1}{24}\mathbf{T}_1^4 + \dots \text{higher excitations...} \end{aligned} \quad (2.20)$$

Coupled Clusters

The first term in the expansion yields the HF function. The second term yields the single and double excitations (with coefficients) and the third term yields double, triple and quadruple excitations (with products of single and/or double excitation coefficients).

The coefficients in \mathcal{T} are called the connected cluster amplitudes. Eq (2.20) shows that the relation of an expansion coefficient of a determinant in Ψ_{CC} to the amplitudes in \mathcal{T} may be rather complicated, because the coefficients of the determinants in the expansion generally also contain products of contributions from lower excitations. These products are the disconnected (factorisable) contributions.

§3 Size consistency

Functions of the form of Eq (2.16) satisfy the product rule of Eq (1.3). This may be shown as follows. If it is assumed that the system contains two subsystems A and B without interaction, we have

$$\mathcal{T}_{AB} = \mathcal{T}_A + \mathcal{T}_B \quad (2.21)$$

and, since all excitation operators commute,

$$\exp(\mathcal{T}_{AB}) = \exp(\mathcal{T}_A)\exp(\mathcal{T}_B) \quad (2.22)$$

where it is assumed that the orbitals are localised on one of the subsystems. All excitations which transfer one or more electrons from one subsystem to the other, do not contribute since they do not interact with any excitation generated by Eq (2.22).

For the HF reference function we may write (cf the discussion in § 1.2) :

$$\Phi_0^{AB} = \Phi_0^A \Phi_0^B \quad (2.23)$$

and we have :

$$\Psi_{CC}^{AB} = \exp(\mathcal{T}_A)\exp(\mathcal{T}_B)\Phi_0^A\Phi_0^B = \exp(\mathcal{T}_A)\Phi_0^A\exp(\mathcal{T}_B)\Phi_0^B = \Psi_{CC}^A\Psi_{CC}^B \quad (2.24)$$

Exercise 2.4

Verify that

$$\exp(\mathcal{A}+\mathcal{B}) = \exp(\mathcal{A})\exp(\mathcal{B})\exp(-c/2) = \exp(\mathcal{B})\exp(\mathcal{A})\exp(c/2)$$

$$\text{if } [\mathcal{A},\mathcal{B}] = \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = c \text{ (c scalar).}$$

From this it follows that $\exp(\mathcal{A}+\mathcal{B}) = \exp(\mathcal{A})\exp(\mathcal{B})$ only holds if $c = 0$.

Coupled Clusters

§4 Calculation of the CCD wave function and energy

In the following we will restrict ourselves to the double excitations by using $\mathbf{T}=\mathbf{T}_2$. This is called the CCD method. The function containing all double excitations is given by

$$|\mathbf{T}|0\rangle = \sum_{I<J} \sum_{A<B} |D_{IJ}^{AB}\rangle = d_{IJ}^{AB} \quad (2.25)$$

$$\langle D_{IJ}^{AB} | \Psi_{CC} \rangle = \langle D_{IJ}^{AB} | \exp(\mathbf{T}) | 0 \rangle = \langle D_{IJ}^{AB} | \mathbf{T} | 0 \rangle = \langle D_{IJ}^{AB} | d_{IJ}^{AB} \tau_{IJ}^{AB} | 0 \rangle = d_{IJ}^{AB}$$

so in this case the connected amplitudes are the same as the expansion coefficients. It will be assumed that the amplitudes d in Eq (2.25) are reasonably small, such that the power expansion Eq (2.19) converges quickly. The factors $1/n!$ ensure that each combination of double excitations appears only once. It is also seen that the coefficients of the quadruply excited determinants correspond to products of double excitation coefficients. There are no connected quadruple excitation amplitudes. For the model system of 2-electron pairs without interaction this is identical to the FCI result.

Exercise 2.5

Show that the product $\tau_{IJ}^{AB} \tau_{KL}^{CD}$ shows up only once in the quadratic term of Eq (2.19).

And what about $(\tau_{IJ}^{AB})^2$?

We have seen that the CC wave function behaves correctly for systems without interaction. However, the CC method also has some disadvantages. The best way to optimize the wave function would be to minimise the energy by a variational procedure. However, the wave function contains all higher excitations, which hopelessly complicates the calculation of the energy as the expectation value of the Hamiltonian. Therefore another method is used, which enables us to get rid of the explicit appearance of all higher excitations. The disadvantage of this method is, that the energy is no longer an upper bound to the FCI energy. However, in practice this disadvantage is much less important than the size consistency error in DCI calculations.

The procedure for calculating the coefficients d in the CCD function

$$|\Psi_{CCD}\rangle = \exp(\mathbf{T}) | 0 \rangle \quad \text{where} \quad \mathbf{T} = \mathbf{T}_2 \quad (2.26)$$

is as follows. The Schrödinger Equation to be satisfied is given by :

$$(\mathbf{H} - E_{CCD}) |\Psi_{CCD}\rangle = 0 \quad (2.27)$$

First this equation is projected to the HF state and Eq (2.19) is used :

$$\langle 0 | \mathbf{H} - E_{CCD} | \Psi_{CCD} \rangle = \langle 0 | (\mathbf{H} - E_{CCD})(1 + \mathbf{T}) | 0 \rangle = 0 \quad (2.28)$$

Coupled Clusters

or

$$E_{\text{CCD}} = E_0 + \langle 0 | \mathbf{H} \mathbf{T} | 0 \rangle \quad (2.29)$$

The higher order terms do not contribute, since they involve at least 4-fold excitations.

Then we project to the space spanned by the double excitations :

$$\langle D_{IJ}^{\text{AB}} | \mathbf{H} - E_{\text{CCD}} | \Psi_{\text{CCD}} \rangle = \langle D_{IJ}^{\text{AB}} | (\mathbf{H} - E_{\text{CCD}}) (1 + \mathbf{T} + \frac{1}{2} \mathbf{T}^2) | 0 \rangle = 0 \quad (2.30)$$

or :

$$\langle D_{IJ}^{\text{AB}} | \mathbf{H} (1 + \mathbf{T} + \frac{1}{2} \mathbf{T}^2) | 0 \rangle = E_{\text{CCD}} \langle D_{IJ}^{\text{AB}} | \mathbf{T} | 0 \rangle = E_{\text{CCD}} d_{IJ}^{\text{AB}} \quad (2.31)$$

Here the terms with \mathbf{T}^3 etc are not needed. Since we restrict ourselves to the double excitation connected amplitudes, only the quadruple excitations are needed explicitly. All other possible projections are ignored. The CCD energy is calculated by solving Eqs (2.29) and (2.30).

Eqs (2.29) and (2.30) are the equations for the so-called Coupled Cluster Approximation (CCA, Hurley). It is also known as the Coupled Pair Many Electron Theory (CPMET, Cizek and Paldus) or the Coupled Cluster with Doubles (CCD) method. In the CCSD method the single excitations are also included.

Note that the projection method leads to a non-variational algorithm. Therefore Eq (2.24) does not prove that the CC method is size consistent. Also the fact that Eq (2.29) is additive is not a sufficient condition : we also have to prove that the dimer CC equations reduce to the monomer equations if the subsystems are far apart. This may be done by showing that in the equation for an excitation localised on A all terms depending on B cancel, which is indeed the case.

The main problem when solving the CCD equations is caused by the \mathbf{T}^2 term. This term generates the quadruple excitations, each combined with a product of double excitation coefficients. However, each quadruple excitation may be generated in a number of different ways :

$$\tau_{IJ}^{\text{AB}} \tau_{KL}^{\text{CD}} | 0 \rangle = \tau_{KL}^{\text{AB}} \tau_{IJ}^{\text{CD}} | 0 \rangle = \tau_{IJ}^{\text{AD}} \tau_{KL}^{\text{BC}} | 0 \rangle = \dots \quad (2.32)$$

In fact, the coefficient of a quadruple excitation appears to contain 18 independent contributions. Moreover, because of the antisymmetry of the indices in the excitation operators (see Eq (2.13)) we have to keep track of the sign of each contribution. As a result Eq (2.31) has a complicated algebraic structure, which leads to rather time-consuming calculations (compared to DCI).

Eqs (2.29) and (2.31) have the following explicit form :

$$E_{\text{CCD}} = E_0 + \sum_{I < J} \sum_{A < B} \langle 0 | \mathbf{H} | D_{IJ}^{\text{AB}} \rangle d_{IJ}^{\text{AB}} \quad (2.33)$$

Coupled Clusters

$$\begin{aligned}
& \langle D_{IJ}^{AB} | \mathbf{H} | 0 \rangle + \langle D_{IJ}^{AB} | \mathbf{H} | \sum_{K < L < C < D} D_{KL}^{CD} \rangle d_{KL}^{CD} \\
& + \langle D_{IJ}^{AB} | \mathbf{H} | \sum_{K < L < C < D} Q_{IJKL}^{ABCD} \rangle q_{IJKL}^{ABCD} = E_{CCD} d_{IJ}^{AB}
\end{aligned} \tag{2.34}$$

Eq (2.34) may be simplified as follows. The $\langle D | \mathbf{H} | Q \rangle$ matrix elements are related to the $\langle 0 | \mathbf{H} | D \rangle$ matrix elements, since they correspond to the same Slater-Condon H-matrix element rule. The interaction between a quadruple and a double excitation is given by :

$$\langle D_{IJ}^{AB} | \mathbf{H} | Q_{IJKL}^{ABCD} \rangle = (KC/LD) - (KD/LC) \tag{2.35}$$

The same result is obtained for the interaction between the double excitation $|D_{KL}^{CD}\rangle$ with the HF function $|0\rangle$:

$$\langle 0 | \mathbf{H} | D_{KL}^{CD} \rangle = (KC/LD) - (KD/LC) \tag{2.36}$$

Therefore we have :

$$\langle D_{IJ}^{AB} | \mathbf{H} | Q_{IJKL}^{ABCD} \rangle = \langle 0 | \mathbf{H} | D_{KL}^{CD} \rangle \tag{2.37}$$

The result is :

$$\begin{aligned}
& \langle D_{IJ}^{AB} | \mathbf{H} | 0 \rangle + \langle D_{IJ}^{AB} | \mathbf{H} | \sum_{K < L < C < D} D_{KL}^{CD} \rangle d_{KL}^{CD} \\
& + \langle 0 | \mathbf{H} | \sum_{K < L < C < D} D_{KL}^{CD} \rangle q_{IJKL}^{ABCD} = E_{CCD} d_{IJ}^{AB}
\end{aligned} \tag{2.38}$$

where each q-coefficients is a linear combination of products of 2 d-coefficients. Eq (2.38) shows that the CCD method uses exactly the same information in terms of H-matrix elements as the DCI method.

§5 Comparison to DCI

The DCI function in intermediate normalisation contains only the $(1+\mathbf{T})$ part of Eq (2.26) defining the CCD function. The number of variables is the same for both methods. The DCI equations for a function with intermediate normalisation are :

$$\langle 0 | \mathbf{H} - E | \Psi_{DCI} \rangle = \langle 0 | (\mathbf{H} - E)(1 + \mathbf{T}) | 0 \rangle = 0 \tag{2.39}$$

$$\langle D_{IJ}^{AB} | \mathbf{H} - E | \Psi_{DCI} \rangle = \langle D_{IJ}^{AB} | (\mathbf{H} - E)(1 + \mathbf{T}) | 0 \rangle = 0 \tag{2.40}$$

Eq (2.39) has the same form as the first CC equation (2.28). However, Eq (2.40) is different from Eq (2.30) since the \mathbf{T}^2 term is missing in the DCI equations. This is the term that takes care of the size consistency in the CCD equations. Fig 2.1 illustrates the difference between DCI and CCD, taking the FCI H-matrix as the reference.

Coupled Clusters

Excitation level

DCI

CCD

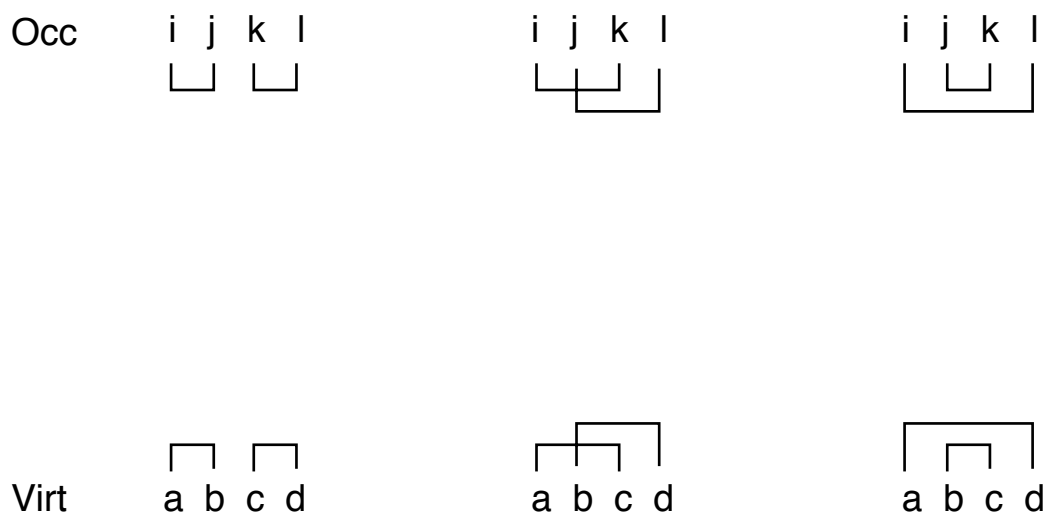
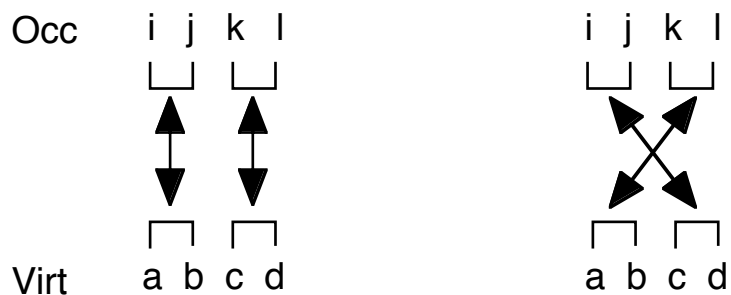
	0	2	4	6
0	X	X	0	0
2	X	X	X	0
4	0	X	X	X
6	0	0	X	X
	0	0	0	X

Fig 2.1 Structure of the FCI and DCI matrices and the projection to the CCD equations.

Eq (2.38) may be used to construct size consistent methods which are less time-consuming than the CCD method. This is done by relating the q -coefficients to the d -coefficients by neglecting certain terms in Eq (2.32). This results in the Coupled Electron Pairs Approach (CEPA).

Coupled Clusters

Quadruple excitation coefficients in CCD



$2 \times 3 \times 3 = 18$ combinations
if all indices are different