

# 13

## COUPLED-CLUSTER THEORY

The coupled-cluster method represents the most successful approach to accurate many-electron molecular wave functions. It can be applied to relatively large systems and is capable of recovering a large part of the correlation energy. It is size-extensive and presents few if any problems with respect to optimization. It does, however, require the existence of a reasonably accurate single-determinant wave function and cannot – at least in its more common formulation – be applied to systems with degenerate or nearly degenerate electronic configurations. In practice, therefore, the application of the coupled-cluster method is restricted to systems that are dominated by a single electronic configuration and the coupled-cluster wave function is best regarded as providing an accurate correction to the Hartree–Fock description.

The present chapter contains a general discussion of the coupled-cluster method and also a detailed exposition of the coupled-cluster singles-and-doubles (CCSD) model. We begin our presentation of coupled-cluster theory in Section 13.1, where we introduce the underlying physical model of coupled-cluster theory and the concept of clusters. In Section 13.2, we introduce the important exponential ansatz of coupled-cluster theory and employ this ansatz to study in more detail the structure of the coupled-cluster wave function and its optimization. Following these introductory sections, we go on to discuss various aspects of coupled-cluster theory such as size-extensivity in Section 13.3 and optimization techniques in Section 13.4. The variational Lagrangian formulation of coupled-cluster theory and the Hellmann–Feynman theorem are then discussed in Section 13.5, followed by a treatment of the calculation of excited states and excitation energies in Section 13.6. In Section 13.7, we turn our attention to an important special case of coupled-cluster theory – the CCSD model, for which all expressions needed for the calculation of the energy and the optimization of the wave function are explicitly derived. In Section 13.8, we consider some important treatments of the correlation problem that are modifications of the standard coupled-cluster theory (i.e. the Brueckner and quadratic CI models). We conclude this chapter with a discussion of open-shell systems in Section 13.9. In this chapter, multireference systems are not discussed since at present no consensus regarding the treatment of such systems in the context of coupled-cluster theory has been established.

### 13.1 The coupled-cluster model

The purpose of the present section is to introduce the coupled-cluster model. First, in Section 13.1.1, we consider the description of virtual excitation processes and correlated electronic states by means of pair clusters. Next, in Section 13.1.2, we introduce the coupled-cluster model as a generalization of the concept of pair clusters. After a discussion of connected and disconnected clusters in Section 13.1.3, we consider the conditions for the optimized coupled-cluster state in Section 13.1.4.

## 13.1.1 PAIR CLUSTERS

In the independent-particle model, the wave function, which corresponds to a product of creation operators working on the vacuum state, describes an uncorrelated motion of the electrons. The variationally optimized electronic system is represented by a set of occupied spin orbitals from which no virtual excitations ever occur since the electrons do not interact. As a refinement to this model, we note that, within the orbital picture, the correlated motion of interacting electrons manifests itself in virtual excitations of electrons from occupied to unoccupied spin orbitals. With each such excitation, we may associate an amplitude, representing the probability that this particular excitation will occur as a result of interactions among the electrons.

To a first approximation, we may restrict ourselves to a pairwise-correlated treatment of the electrons. Consider two electrons that, in the independent-particle model, occupy the spin orbitals  $I$  and  $J$ . As a result of the instantaneous interaction between the two electrons, their motion is disturbed and the electrons are excited to a different set of spin orbitals  $A$  and  $B$ , initially unoccupied. With each excitation process, we associate an amplitude  $t_{IJ}^{AB}$ . Our description of the motion of the two electrons is thus improved in the following manner

$$a_I^\dagger a_J^\dagger \rightarrow a_I^\dagger a_J^\dagger + \sum_{A>B} t_{IJ}^{AB} a_A^\dagger a_B^\dagger \quad (13.1.1)$$

and this expansion is known as a *two-electron cluster* or a *pair cluster*. We now introduce an operator that describes this 'correlation process' of two electrons initially found in the spin orbitals  $I$  and  $J$  of the reference state. Introducing the notation

$$\hat{\tau}_{IJ}^{AB} = a_A^\dagger a_I a_J^\dagger a_B \quad (13.1.2)$$

and using the relation

$$\hat{\tau}_{IJ}^{AB} \hat{\tau}_{IJ}^{CD} = 0 \quad (13.1.3)$$

we may write the pair cluster  $IJ$  in the following way

$$\left[ \prod_{A>B} (1 + t_{IJ}^{AB} \hat{\tau}_{IJ}^{AB}) \right] a_I^\dagger a_J^\dagger | \rangle = a_I^\dagger a_J^\dagger | \rangle + \sum_{A>B} t_{IJ}^{AB} a_A^\dagger a_B^\dagger | \rangle \quad (13.1.4)$$

assuming that  $I$  and  $J$  are unoccupied in  $| \rangle$ . Each operator of the form  $1 + t_{IJ}^{AB} \hat{\tau}_{IJ}^{AB}$  provides an improved correlated description of the electronic system, generating a superposition of the original state with a new state that represents the outcome of the excitation process.

Assuming a Hartree–Fock reference state and introducing cluster expansions of type (13.1.4) for all pairs of occupied spin orbitals, we arrive at the following expression for our pair-correlated electronic state [1]:

$$|\text{CCD}\rangle = \left[ \prod_{A>B, I>J} (1 + t_{IJ}^{AB} \hat{\tau}_{IJ}^{AB}) \right] |\text{HF}\rangle \quad (13.1.5)$$

Since the excitation operators (13.1.2) commute among one another, there are no problems with the order of the operators in this expression. The resulting wave function (13.1.5) corresponds to a particularly simple realization of the *coupled-cluster model*, in which only double excitations are allowed: the *coupled-cluster doubles (CCD) wave function*. For a complete specification of this model, we must also describe the method by which the amplitudes are determined. We shall

return to this point in Section 13.1.4, following the introduction and discussion of the general coupled-cluster model in Sections 13.1.2 and 13.1.3.

### 13.1.2 THE COUPLED-CLUSTER WAVE FUNCTION

Excitations within the pair clusters of Section 13.1.1 provide the dominant contributions to the description of the complicated correlated motion of interacting electrons. The pair clusters dominate since the correlated motion is especially important for electrons that are close to each other and since at most two electrons (with opposite spins) may coincide in space. However, for an accurate treatment of the correlated motion of the electrons (within the orbital model), we must consider clusters of all sizes. For the three-electron clusters, we thus introduce amplitudes  $t_{IJK}^{ABC}$  that represent the simultaneous interaction of three electrons, resulting in the excitation of three electrons from three occupied spin orbitals to three unoccupied ones. Furthermore, we must allow excitations to occur also within 'clusters' containing a single electron. Such one-electron processes represent a relaxation of the spin orbitals and occur since the Hartree–Fock mean field experienced by each electron before the excitations were introduced is modified by the many-electron excitation processes occurring within the remaining clusters.

For a general discussion of the coupled-cluster method, we introduce the generic notation  $\hat{\tau}_\mu$  for an excitation operator of unspecified excitation level (single, double, etc.) and  $t_\mu$  for the associated amplitude. In our model, the effect of a given excitation  $\mu$  on any state  $| \rangle$  is to change this state as follows

$$| \rangle \rightarrow (1 + t_\mu \hat{\tau}_\mu) | \rangle \quad (13.1.6)$$

and the result of such *correlating operators* working on the Hartree–Fock state is the *coupled-cluster wave function*

$$|\text{CC}\rangle = \left[ \prod_{\mu} (1 + t_\mu \hat{\tau}_\mu) \right] |\text{HF}\rangle \quad (13.1.7)$$

Several points should be noted about the form of this wave function. First, the coupled-cluster state is manifestly in a product form, leading to a size-extensive treatment of the electronic system as discussed in Sections 4.3 and 13.3. In this respect, the coupled-cluster model differs fundamentally from the linear CI model, for which the corresponding wave function is not in a product form:

$$|\text{CI}\rangle = \left( 1 + \sum_{\mu} C_{\mu} \hat{\tau}_{\mu} \right) |\text{HF}\rangle \quad (13.1.8)$$

Second, the order of the operators in (13.1.7) is unimportant since the excitation operators  $\hat{\tau}_\mu$  commute:

$$[\hat{\tau}_\mu, \hat{\tau}_\nu] = 0 \quad (13.1.9)$$

Commutation occurs since the excitation operators always excite from the set of occupied Hartree–Fock spin orbitals to the virtual ones – see (13.1.2) for the double-excitation operators. The creation and annihilation operators of the excitation operators therefore anticommute.

### 13.1.3 CONNECTED AND DISCONNECTED CLUSTERS

Let us now analyse the coupled-cluster state in terms of determinants

$$|\mu\rangle = \hat{\tau}_\mu |\text{HF}\rangle \quad (13.1.10)$$

Upon expansion of the product operator in (13.1.7), we obtain the expression [1]

$$\begin{aligned} |\text{CC}\rangle &= \left( 1 + \sum_{\mu} t_{\mu} \hat{\tau}_{\mu} + \sum_{\mu>\nu} t_{\mu} t_{\nu} \hat{\tau}_{\mu} \hat{\tau}_{\nu} + \dots \right) |\text{HF}\rangle \\ &= |\text{HF}\rangle + \sum_{\mu} t_{\mu} |\mu\rangle + \sum_{\mu>\nu} t_{\mu} t_{\nu} |\mu\nu\rangle + \dots \end{aligned} \quad (13.1.11)$$

where, in the summation, the restriction  $\mu > \nu$  applies since each excitation operator appears only once in the product (13.1.7). Clearly, we cannot associate each excitation operator and its amplitude with a single unique determinant – the operator  $\hat{\tau}_{\mu}$  not only generates the determinant  $|\mu\rangle$  but also gives rise to a large number of determinants in collaboration with the other excitation operators, for example

$$|\mu\nu\rangle = \hat{\tau}_{\mu} \hat{\tau}_{\nu} |\text{HF}\rangle = \hat{\tau}_{\nu} \hat{\tau}_{\mu} |\text{HF}\rangle \quad (13.1.12)$$

As seen from the expansion (13.1.11), the amplitude of such a composite excitation is obtained by taking the product of the amplitudes of the contributing excitations.

Owing to the presence of product excitations in the coupled-cluster state (13.1.7), each determinant may be reached in several distinct ways. For example, the introduction of a composite excitation operator equivalent to the product of two separate excitations

$$\hat{\tau}_{\mu\nu} = \hat{\tau}_{\mu} \hat{\tau}_{\nu} \quad (13.1.13)$$

means that the determinant  $|\mu\nu\rangle$  may be reached by at least two distinct processes, with an overall amplitude equal to the sum of the individual amplitudes:

$$t_{\mu\nu}^{\text{total}} = t_{\mu\nu} + t_{\mu} t_{\nu} + \dots \quad (13.1.14)$$

With respect to the determinant  $|\mu\nu\rangle$ , the amplitude  $t_{\mu\nu}$  is referred to as a *connected cluster amplitude* and  $t_{\mu} t_{\nu}$  as a *disconnected cluster amplitude*. In general, high-order excitations can be reached by a large number of processes or mechanisms, each contributing to the total amplitude with a weight equal to the product of the amplitudes of the individual excitations.

From (13.1.11) it is apparent that a coupled-cluster wave function – generated, for example, by all possible single- and double-excitation operators  $\hat{\tau}_{\mu}$  – contains contributions from *all* determinants entering the FCI wave function (although the number of free parameters is usually much smaller). In practice, therefore, we cannot work with the coupled-cluster state in the expanded form (13.1.11) but we must instead retain the wave function in the more compact form (13.1.7), avoiding references to the individual determinants.

#### 13.1.4 THE COUPLED-CLUSTER SCHRÖDINGER EQUATION

Given the product ansatz for the coupled-cluster wave function (13.1.7), let us consider its optimization. We recall that, in CI theory, the wave function (13.1.8) is determined by minimizing the expectation value of the Hamiltonian with respect to the linear expansion coefficients:

$$E_{\text{CI}} = \min_{c_{\mu}} \frac{\langle \text{CI} | \hat{H} | \text{CI} \rangle}{\langle \text{CI} | \text{CI} \rangle} \quad (13.1.15)$$

By analogy with CI theory, therefore, we might attempt to determine the coupled-cluster state by minimizing the expectation value of the Hamiltonian with respect to the amplitudes:

$$E_{\min} = \min_{t_{\mu}} \frac{\langle \text{CC} | \hat{H} | \text{CC} \rangle}{\langle \text{CC} | \text{CC} \rangle} \quad (13.1.16)$$

We recall that the derivatives of the CI wave function with respect to the variational parameters may be written in a particularly simple form

$$\frac{\partial}{\partial C_{\mu}} |\text{CI}\rangle = |\mu\rangle \quad (13.1.17)$$

giving rise to a standard eigenvalue problem for the CI coefficients (11.1.3)

$$\langle \mu | \hat{H} | \text{CI} \rangle = E_{\text{CI}} \langle \mu | \text{CI} \rangle \quad (13.1.18)$$

By contrast, the nonlinear parametrization of the coupled-cluster model (13.1.7) means that the derivatives of the coupled-cluster state become complicated functions of the amplitudes

$$\frac{\partial}{\partial t_{\mu}} |\text{CC}\rangle = \left[ \prod_{\nu} (1 + t_{\nu} \hat{\tau}_{\nu}) \right] |\mu\rangle \quad (13.1.19)$$

The variational conditions on the amplitudes (13.1.16) therefore give rise to an intractable set of nonlinear equations for the amplitudes

$$\langle \mu | \left[ \prod_{\nu} (1 + t_{\nu} \hat{\tau}_{\nu}^{\dagger}) \right] \hat{H} | \text{CC} \rangle = E_{\min} \langle \mu | \left[ \prod_{\nu} (1 + t_{\nu} \hat{\tau}_{\nu}^{\dagger}) \right] | \text{CC} \rangle \quad (13.1.20)$$

which involves the full set of FCI determinants and high-order products of the amplitudes. The variational minimization of the coupled-cluster energy is thus a complicated undertaking, which can be carried out only for small molecular systems. We shall therefore abandon the variation principle as a basis for the optimization of the coupled-cluster wave function.

To establish a different principle for the optimization of the coupled-cluster wave function, we note that, for the linear CI wave function, the variational minimization of the energy is entirely equivalent to the solution of the *projected Schrödinger equation* (13.1.18). By contrast, for nonlinearly parametrized wave functions, the solution of the projected Schrödinger equation is, in general, not equivalent to the minimization of the energy. For such models, therefore, we may regard the solution of the projected Schrödinger equation as an alternative to the minimization of the energy. In particular, applied to the coupled-cluster model, projection of the Schrödinger equation against those determinants that enter the coupled-cluster state (13.1.11) with connected amplitudes

$$\langle \mu | = \langle \text{HF} | \hat{\tau}_{\mu}^{\dagger} \quad (13.1.21)$$

gives us the projected coupled-cluster equations

$$\langle \mu | \hat{H} | \text{CC} \rangle = E \langle \mu | \text{CC} \rangle \quad (13.1.22)$$

where the coupled-cluster energy is obtained by projection against the Hartree–Fock state

$$E = \langle \text{HF} | \hat{H} | \text{CC} \rangle \quad (13.1.23)$$

since

$$\langle \text{HF} | \text{CC} \rangle = 1 \quad (13.1.24)$$

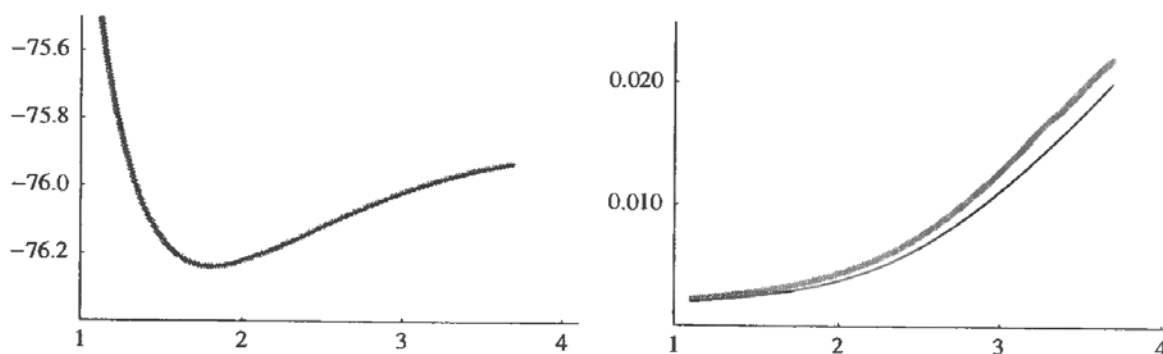
Like the variational coupled-cluster conditions (13.1.20), the projected equations (13.1.22) are nonlinear in the amplitudes. However, unlike the variational conditions, the expansion of the wave function in (13.1.22) and (13.1.23) terminates after a few terms since the Hamiltonian operator couples determinants that differ by no higher than double excitations, making the solution of the projected equations and the calculation of the energy tractable. Of course, the calculated coupled-cluster energy no longer represents an upper bound to the FCI energy. In practice, the deviation from the variational energy turns out to be small and of little practical consequence.

The nonvariational character of the coupled-cluster energy may be observed in different situations. For example, in the course of the optimization of the wave function, we may sometimes find that the partially optimized energy is below the final, fully optimized coupled-cluster energy – see Table 13.2. Similarly, the coupled-cluster energy calculated in a given basis and at a given truncation level may sometimes be below the FCI energy – see Table 5.11, where the coupled-cluster energies of the stretched water molecule are lower than the FCI energy at some levels of truncation. In general, this somewhat unsystematic behaviour of the coupled-cluster energy (compared with a variationally determined energy) is irrelevant since the calculated energies are nevertheless rather accurate.

The small difference between the energy calculated as an expectation value and by projection is illustrated in Figure 13.1, where we have plotted both the standard CCSD energy (13.1.23) and the energy obtained from the expression

$$E_{\text{ave}} = \frac{\langle \text{CC} | \hat{H} | \text{CC} \rangle}{\langle \text{CC} | \text{CC} \rangle} \quad (13.1.25)$$

for the water molecule in the cc-pVDZ basis. The left-hand plot shows that the two energies are essentially the same, the difference being scarcely detectable. The right-hand plot, which contains the errors relative to FCI, reveals that, for this system, the expectation value is always lower than the standard CCSD energy, even though they have been obtained from the same amplitudes. Obviously, a variational minimization of (13.1.25) would lead to a further (presumably small) lowering of the energy. All things considered, we conclude that the more complicated variational expression (13.1.25) does not improve the description significantly.



**Fig. 13.1.** The CCSD energy (in  $E_h$ ) calculated in the standard manner (grey line) and as an expectation value (black line) for the water molecule in the cc-pVDZ basis, at a fixed bond angle of  $110.565^\circ$ . On the left, we have plotted the potential-energy curve as a function of the bond distance (in  $a_0$ ); on the right, we have plotted the difference relative to the FCI energy.

## 13.2 The coupled-cluster exponential ansatz

In Section 13.1, the coupled-cluster wave function was expressed as a product of correlating operators working on the Hartree–Fock state (13.1.7). This expression for the coupled-cluster wave function is useful for displaying the relation to CI theory and for exhibiting the fundamental role played by the excitation operators in coupled-cluster theory. In general, however, the coupled-cluster wave function is expressed as the exponential of an operator applied to the Hartree–Fock determinant. In the present section, we introduce and explore this exponential ansatz for the coupled-cluster wave function.

### 13.2.1 THE EXPONENTIAL ANSATZ

In the product form of the coupled-cluster wave function (13.1.7), the spin-orbital excitation operators satisfy the relation

$$\hat{\tau}_\mu^2 = 0 \quad (13.2.1)$$

and the correlating operators may therefore be written as exponentials of the excitation operators:

$$1 + t_\mu \hat{\tau}_\mu = \exp(t_\mu \hat{\tau}_\mu) \quad (13.2.2)$$

Using the commutators (13.1.9), we may then write the wave function in the form

$$|\text{CC}\rangle = \exp(\hat{T})|\text{HF}\rangle \quad (13.2.3)$$

known as the *exponential ansatz* for the coupled-cluster wave function [2–4], where the *cluster operator*

$$\hat{T} = \sum_{\mu} t_{\mu} \hat{\tau}_{\mu} \quad (13.2.4)$$

is a linear combination of excitation operators multiplied by the associated (connected) cluster amplitudes. In the spin-orbital basis, expansion of the exponential operator in (13.2.3) leads to the same expression as does the expansion of the product function (13.1.7), producing contributions from all possible configurations that may be obtained by excitations from the Hartree–Fock state – as appropriate for an electronic-structure model that emphasizes the physical process of excitations rather than the resulting configurations.

### 13.2.2 THE COUPLED-CLUSTER HIERARCHY OF EXCITATION LEVELS

In coupled-cluster theory, the wave function is written as an exponential of a cluster operator (13.2.4) acting on a single-determinant reference state (13.2.3). In constructing the wave function, the excitations included in the cluster operator are not selected individually. Rather, a *hierarchy of approximations* is established by partitioning the cluster operator into classes comprising all *single (one-electron) excitations*, all *double (two-electron) excitations*, all *triple (three-electron) excitations*, and so on. We may then write the cluster operator in the form

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_N \quad (13.2.5)$$

where, for example, the one- and two-electron parts are given by

$$\hat{T}_1 = \sum_{AI} t_I^A a_A^\dagger a_I = \sum_{AI} t_I^A \hat{\tau}_I^A \quad (13.2.6)$$

$$\hat{T}_2 = \sum_{\substack{A>B \\ I>J}} t_{IJ}^{AB} a_A^\dagger a_I a_B^\dagger a_J = \frac{1}{4} \sum_{AIBJ} t_{IJ}^{AB} a_A^\dagger a_I a_B^\dagger a_J = \frac{1}{4} \sum_{AIBJ} t_{IJ}^{AB} \hat{\tau}_{IJ}^{AB} \quad (13.2.7)$$

As before, the indices  $I$  and  $J$  are used for the occupied Hartree–Fock spin orbitals and the indices  $A$  and  $B$  for the unoccupied (virtual) spin orbitals. The cluster amplitudes  $t_{IJ}^{AB}$  are antisymmetric with respect to permutations of  $A$  and  $B$  and permutations of  $I$  and  $J$ .

Each excitation operator in (13.2.5) excites at least one electron from an occupied to a virtual spin orbital. Since there are  $N$  electrons in the system, the expansion (13.2.5) terminates after  $\hat{T}_N$ . We further note that the excitation operators given by (13.2.6) and (13.2.7) satisfy the commutation relation (13.1.9) and that the following relationship holds

$$\hat{\tau}_\mu^\dagger |\text{HF}\rangle = 0 \quad (13.2.8)$$

since it is impossible to excite an electron from an unoccupied orbital. We also note that

$$[\hat{T}_i, \hat{T}_j] = 0 \quad (13.2.9)$$

as a simple consequence of the commutation relation for the individual excitation operators.

To compare the excitation-based coupled-cluster model with the configuration-based CI model, we expand the exponential operator in (13.2.3) and collect terms to the same order in the excitation level:

$$\exp(\hat{T})|\text{HF}\rangle = \sum_{i=0}^N \hat{C}_i |\text{HF}\rangle \quad (13.2.10)$$

The lowest-order ‘configuration’ operators  $\hat{C}_i$  are given by

$$\hat{C}_0 = 1 \quad (13.2.11)$$

$$\hat{C}_1 = \hat{T}_1 \quad (13.2.12)$$

$$\hat{C}_2 = \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 \quad (13.2.13)$$

$$\hat{C}_3 = \hat{T}_3 + \hat{T}_1 \hat{T}_2 + \frac{1}{6} \hat{T}_1^3 \quad (13.2.14)$$

$$\hat{C}_4 = \hat{T}_4 + \hat{T}_1 \hat{T}_3 + \frac{1}{2} \hat{T}_2^2 + \frac{1}{2} \hat{T}_1^2 \hat{T}_2 + \frac{1}{24} \hat{T}_1^4 \quad (13.2.15)$$

These equations show which excitation processes contribute at each excitation level. Thus, the quadruply excited configurations are generated by five distinct mechanisms, where, for instance, the disconnected  $\hat{T}_2^2$  term represents the independent interactions within two distinct pairs of electrons and the connected  $\hat{T}_4$  term describes the simultaneous interaction of four electrons. The disconnected terms represent interactions of product clusters within disjoint sets of electrons and vanish whenever two or more spin-orbital indices are identical.

Without truncation, the FCI and full coupled-cluster functions contain the same number of parameters since there is then one connected cluster amplitude for each determinant. In this special case, the CI and coupled-cluster models provide linear and nonlinear parametrizations of the same state and there is then no obvious advantage in employing the more complicated exponential parametrization. The advantages of the cluster parametrization become apparent only upon truncation and are related to the fact that, even at the truncated level, the coupled-cluster state contains contributions from all determinants in the FCI wave function, with weights obtained from the different excitation processes leading to the determinants.



The most common approximation in coupled-cluster theory is to truncate the cluster operator at the doubles level, yielding the *coupled-cluster singles-and-doubles (CCSD) model* [5]. In this model, the  $\hat{T}_2$  operator describes the important electron-pair interactions and  $\hat{T}_1$  carries out the orbital relaxations induced by the field set up by the pair interactions. The CCSD wave function contains contributions from all determinants of the FCI wave function, although the highly excited determinants, generated by disconnected clusters, are in general less accurately described than those that also contain connected contributions. However, the disconnected contributions may also in many situations be dominant. For example, even though the  $\hat{T}_4$  and  $\hat{T}_3\hat{T}_1$  contributions to the quadruples (13.2.15) are neglected at the CCSD level, these determinants are still quite accurately described since the  $\hat{T}_2^2$  contributions, which represent the simultaneous but independent interactions within disjoint pairs of electrons, usually constitute the dominant contributions to the  $\hat{C}_4$  amplitudes.

To investigate the importance of the disconnected clusters, we have in Table 13.1 listed the weights of the various excitation levels in the (normalized) CCSD and FCI wave functions of the water molecule in the cc-pVDZ basis. At the reference geometry  $R_{\text{ref}}$ , the Hartree–Fock determinant is a good zero-order approximation and the singles and doubles weights are similar at the CCSD and FCI levels, differing by only about 5%. Turning our attention to the higher excitations (which in CCSD theory are represented by disconnected clusters), we note that the CCSD wave function recovers as much as 84% of the FCI quadruples weight. The quadruples are therefore well described in CCSD theory, mainly by means of  $\hat{T}_2^2$ . On the other hand, the CCSD wave function is incapable of describing the triple excitations, recovering as little as 4% of their total weight in the FCI wave function – apparently, the triples are not well represented by disconnected clusters.

At the stretched geometry  $2R_{\text{ref}}$ , the single-determinant reference state is no longer a good zero-order approximation. The CCSD model is consequently a poor one, underestimating the doubles by as much as 14%. Again the triples are poorly described but we note that the disconnected quadruples represent as much as 81% of the total FCI quadruples weight. In conclusion, the size-extensive CCSD model appears to recover most of the effects of the quadruples – at least for systems without large contributions from static correlation. The CCSD model is discussed in detail in Section 13.7.

For high accuracy, we must take into account also the connected triple excitations. Truncating the cluster expansion (13.2.5) at the  $\hat{T}_3$  level, we arrive at the *coupled-cluster singles-doubles-and-triples (CCSDT) model* [6,7]. Although highly accurate for the description of dynamical correlation, the CCSDT model is computationally very demanding and can be applied only to small systems.

**Table 13.1** The weights of the excitation levels in the normalized CCSD and FCI wave functions for the water molecule in the cc-pVDZ basis

Excitation level	$R_{\text{ref}}$		$2R_{\text{ref}}$	
	CCSD	FCI	CCSD	FCI
0	0.94410	0.94100	0.65114	0.58966
1	0.00056	0.00053	0.02494	0.02680
2	0.05413	0.05650	0.28762	0.33300
3	0.00002	0.00055	0.00371	0.01040
4	0.00115	0.00137	0.03225	0.03970
5	0.00000	0.00002	0.00006	0.00080
6	0.00001	0.00002	0.00028	0.00044

Unlike the CCSD model, the CCSDT and higher models – such as the *coupled-cluster singles-doubles-triples-and-quadruples (CCSDTQ) model* [8] – are therefore not treated in detail in this chapter. Fortunately, accurate coupled-cluster models have been developed that include the effects of the connected triples in an approximate fashion. We defer the discussion of such approximate CCSDT treatments to Chapter 14.

### 13.2.3 THE PROJECTED COUPLED-CLUSTER EQUATIONS

In a given orbital basis, the full coupled-cluster wave function satisfies the Schrödinger equation

$$\hat{H} \exp(\hat{T})|\text{HF}\rangle = E \exp(\hat{T})|\text{HF}\rangle \quad (13.2.16)$$

Truncated coupled-cluster wave functions cannot satisfy this equation exactly and, as discussed in Section 13.1.4, we use projection to determine the wave function. The optimized coupled-cluster wave function then satisfies the Schrödinger equation (13.2.16) projected onto the Hartree–Fock state and onto the excited projection manifold

$$\langle \mu | = \langle \text{HF} | \hat{\tau}_\mu^\dagger \quad (13.2.17)$$

The resulting projected coupled-cluster equations may be written as

$$\langle \text{HF} | \hat{H} \exp(\hat{T}) | \text{HF} \rangle = E \quad (13.2.18)$$

$$\langle \mu | \hat{H} \exp(\hat{T}) | \text{HF} \rangle = E \langle \mu | \exp(\hat{T}) | \text{HF} \rangle \quad (13.2.19)$$

In the CCSD model, for example, the excited projection manifold comprises the full set of all singly and doubly excited determinants, giving rise to one equation (13.2.19) for each connected amplitude. For the full coupled-cluster wave function, the number of equations is equal to the number of determinants and the solution of the projected equations recovers the FCI wave function. The nonlinear equations (13.2.19) must be solved iteratively, substituting in each iteration the coupled-cluster energy as calculated from (13.2.18).

For many purposes, it is convenient to express the projected coupled-cluster equations in a slightly different form. First, we multiply the Schrödinger equation (13.2.16) from the left by the operator  $\exp(-\hat{T})$  to obtain

$$\exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle = E | \text{HF} \rangle \quad (13.2.20)$$

which may be regarded as a Schrödinger equation with an effective, non-Hermitian *similarity-transformed Hamiltonian*:

$$\hat{H}^T = \exp(-\hat{T}) \hat{H} \exp(\hat{T}) \quad (13.2.21)$$

Projecting the similarity-transformed Schrödinger equation (13.2.20) against the same determinants as in (13.2.18) and (13.2.19), we arrive at the following set of equations for the coupled-cluster amplitudes and energy:

$$\langle \text{HF} | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle = E \quad (13.2.22)$$

$$\langle \mu | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle = 0 \quad (13.2.23)$$

As will be shown shortly, these similarity-transformed equations are equivalent to (13.2.18) and (13.2.19), yielding the same amplitudes and the same energy upon solution. In the following, we shall refer to (13.2.23) as the *linked coupled-cluster equations* and to (13.2.19) as the *unlinked coupled-cluster equations*. The reason for the use of the terms ‘linked’ and ‘unlinked’ is that, in diagrammatic coupled-cluster theory, the energy-independent (similarity-transformed) equations (13.2.23) give rise to only linked diagrams, whereas the energy-dependent equations (13.2.19) give rise to unlinked as well as to linked diagrams [9]. As we shall see in Section 13.3, even though the linked and unlinked coupled-cluster equations are equivalent and both provide a size-extensive treatment of the electronic system, the treatment of size-extensivity is rather different in the two formulations.

For the full coupled-cluster wave function, the equivalence of the equations (13.2.16) and (13.2.20) is trivial; for truncated cluster expansions, on the other hand, the equivalence of the linked and unlinked forms of the amplitude equations is less obvious and requires special attention. First, the equivalence of the energy expressions (13.2.18) and (13.2.22) is easily established since for any choice of amplitudes

$$\langle \text{HF} | \exp(-\hat{T}) = \langle \text{HF} | \quad (13.2.24)$$

Next, to demonstrate the equivalence of the amplitude equations (13.2.19) and (13.2.23), we introduce the unsymmetric matrices  $\mathbf{T}^+$  and  $\mathbf{T}^-$  with elements given by

$$T_{\mu\nu}^{\pm} = \langle \mu | \exp(\pm \hat{T}) | \nu \rangle = \langle \mu | \nu \rangle \pm \langle \mu | \hat{T} | \nu \rangle + \frac{1}{2} \langle \mu | \hat{T}^2 | \nu \rangle \pm \dots \quad (13.2.25)$$

The lower triangular structure of these two matrices, illustrated in Figure 13.2, follows from the fact that  $\hat{T}$  contains only excitation operators and it is also retained if the definition of  $\mathbf{T}^{\pm}$  is extended to include the Hartree–Fock state. Since the upper-triangular elements are 0 and since the diagonal elements are equal to 1, the matrices  $\mathbf{T}^{\pm}$  are nonsingular for any projection manifold, irrespective of whether or not we omit, for example, excited determinants higher than doubles. We note that the diagonal blocks with  $\mu$  and  $\nu$  belonging to the same excitation level – both singles or both doubles, for instance – are equal to the identity matrices.

	HF	S	D	...	N
HF	1	0	0	...	0
S		1	0	...	0
D			1	...	0
⋮				⋮	0
N					1

**Fig. 13.2.** The lower triangular structure of the matrices  $\mathbf{T}^{\pm}$  in (13.2.25). The diagonal blocks contain identity matrices, the lower triangular blocks are nonzero and the upper triangular blocks zero.

Let us now assume that the similarity-transformed amplitude equations (13.2.23) are satisfied. Since the matrix  $\mathbf{T}^+$  is nonsingular, the conditions (13.2.23) are equivalent to the conditions

$$\begin{aligned} A_\mu &= \sum_\nu T_{\mu\nu}^+ \langle \nu | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle \\ &= \sum_\nu \langle \mu | \exp(\hat{T}) | \nu \rangle \langle \nu | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle = 0 \end{aligned} \quad (13.2.26)$$

where  $\mu$  and  $\nu$  belong to the excitation manifold of (13.2.23). From the structure of  $\mathbf{T}^+$ , we note that  $\langle \mu | \exp(\hat{T})$  contains determinants of excitation levels lower than or equal to that of  $\langle \mu |$ . Assuming that the projection manifold is *closed under de-excitation* (i.e. if  $\langle \mu |$  is a member of the projection manifold, then so is  $\langle \mu | \hat{T}$ ), we may invoke the resolution of the identity

$$\langle \mu | \exp(\hat{T}) = \sum_\nu \langle \mu | \exp(\hat{T}) | \nu \rangle \langle \nu | + \langle \mu | \exp(\hat{T}) | \text{HF} \rangle \langle \text{HF} | \quad (13.2.27)$$

where the summation is again restricted to the excitation manifold. Equation (13.2.26) can now be written as

$$\begin{aligned} A_\mu &= \langle \mu | \exp(\hat{T}) \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle - \langle \mu | \exp(\hat{T}) | \text{HF} \rangle \langle \text{HF} | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle \\ &= \langle \mu | \hat{H} \exp(\hat{T}) | \text{HF} \rangle - E \langle \mu | \exp(\hat{T}) | \text{HF} \rangle \end{aligned} \quad (13.2.28)$$

where we have used (13.2.27) and (13.2.22). We conclude that the conditions (13.2.26) are equivalent to the unlinked equations (13.2.19) if the projection manifold is closed under de-excitation. We have thus established the equivalence of the linked and unlinked coupled-cluster conditions (13.2.23) and (13.2.19) for the standard models CCSD, CCSDT, and so on. Equivalence also holds for all models containing only even-order excitations such as CCD. Finally, the projection manifold is closed and equivalence is maintained if, at the highest excitation level, only selected excitations are retained – for example, if selected triple excitations are included in addition to all singles and doubles.

Let us now compare the coupled-cluster equations in the linked and unlinked forms. We begin by reiterating that these two forms of the coupled-cluster equations are equivalent for the standard models in the sense that they have the same solutions. Moreover, applied at the important CCSD level of theory, neither form is superior to the other, requiring about the same number of floating-point operations. The energy-dependent unlinked form (13.2.19) exhibits more closely the relationship with CI theory, where the projected equations may be written in a similar form (13.1.18). On the other hand, the linked form (13.2.23) has some important advantages over the unlinked one (13.2.19), making it the preferred form in most situations.

First, when the similarity-transformed Hamiltonian is expressed as a BCH expansion (see Section 13.2.5), the coupled-cluster equations may be shown to be no higher than quartic in the cluster amplitudes – for any truncation of the cluster operator. As a bonus, the nested commutators of the BCH expansion reduce the rank of the operators, further simplifying the algebra. Second, although the linked and unlinked equations yield the same, size-extensive wave function, the linked equations have the useful additional property of being size-extensive term by term (see Section 13.3), allowing for a simple control of size-extensivity upon modification of the coupled-cluster equations and making this particular form a useful starting point for the development of perturbation theory as discussed in Section 14.3. Third, the similarity-transformed linked equations also constitute the starting point for the development of a coupled-cluster approach for

the calculation of excited states in Section 13.6. Fourth, by carrying out an explicit similarity transformation of the Hamiltonian using the singles amplitudes, the algebra of the coupled-cluster equations is significantly simplified, as discussed for the CCSD wave function in Section 13.7. For these reasons, we shall, in the remainder of this chapter, mostly employ the coupled-cluster equations in the linked form.

### 13.2.4 THE COUPLED-CLUSTER ENERGY

In coupled-cluster theory, the electronic energy is obtained from (13.2.22):

$$E = \langle \text{HF} | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle = \langle \text{HF} | \hat{H} \exp(\hat{T}) | \text{HF} \rangle \quad (13.2.29)$$

Expanding the cluster amplitudes, we obtain

$$E = \langle \text{HF} | \hat{H} \left( 1 + \hat{T} + \frac{1}{2} \hat{T}^2 + \dots \right) | \text{HF} \rangle = \langle \text{HF} | \hat{H} \left( 1 + \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 \right) | \text{HF} \rangle \quad (13.2.30)$$

Cluster operators higher than doubles do not contribute to the energy since  $\hat{H}$  is a two-particle operator. Because of the Brillouin theorem, the one-particle operators contribute only to second order:

$$\langle \text{HF} | \hat{H} \hat{T}_1 | \text{HF} \rangle = 0 \quad (13.2.31)$$

As a result, only singles and doubles amplitudes contribute directly to the coupled-cluster energy – irrespective of the truncation level in the cluster operator. Of course, the higher-order excitations contribute indirectly since all amplitudes are coupled by the projected equations (13.2.23).

### 13.2.5 THE COUPLED-CLUSTER AMPLITUDE EQUATIONS

Having examined the coupled-cluster energy and seen that it is no higher than quadratic in the cluster amplitudes, let us now turn our attention to the structure of the linked projected coupled-cluster equations (13.2.23):

$$\langle \mu | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \text{HF} \rangle = 0 \quad (13.2.32)$$

Since  $\hat{T}$  is not anti-Hermitian, it gives rise to a nonunitary transformation and the similarity-transformed Hamiltonian operator is therefore non-Hermitian. Naively, we would expect the BCH expansion (3.1.7) of the similarity-transformed Hamiltonian to yield an infinite sequence of nested commutators. Nevertheless, we shall see that the expansion terminates after five terms.

Let the *up rank*  $s_A^+$  and the *down rank*  $s_A^-$  of a string of elementary operators  $\hat{A}$  be given by

$$s_A^+ = \frac{1}{2}(n_v^c + n_o^a) \quad (13.2.33)$$

$$s_A^- = \frac{1}{2}(n_o^c + n_v^a) \quad (13.2.34)$$

where  $n_v^c$  and  $n_o^c$  are the numbers of creation operators in  $\hat{A}$  for the virtual and occupied spin orbitals, respectively; likewise,  $n_o^a$  and  $n_v^a$  are the numbers of occupied and virtual annihilation operators. We also introduce the *excitation rank*  $s_A$  as the difference between the up and down ranks

$$s_A = s_A^+ - s_A^- \quad (13.2.35)$$

and note that the sum of the up and down ranks is equal to the particle rank of  $\hat{A}$ . In Section 13.2.8, we prove the following *cluster-commutation condition* for the vanishing of  $k$  nested commutators of  $\hat{A}$  with cluster operators:

$$k > 2s_A^- \Rightarrow [[\dots [[\hat{A}, \hat{T}_{n_1}], \hat{T}_{n_2}], \dots], \hat{T}_{n_k}] = 0 \quad (13.2.36)$$

Since the highest down rank of the Hamiltonian is 2, it follows that the BCH expansion of the similarity-transformed Hamiltonian is no higher than quartic in the amplitudes:

$$\exp(-\hat{T})\hat{H}\exp(\hat{T}) = \hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{6}[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}] + \frac{1}{24}[[[[\hat{H}, \hat{T}], \hat{T}], \hat{T}], \hat{T}] \quad (13.2.37)$$

The projected coupled-cluster Schrödinger equation (13.2.32) therefore yields at most *quartic equations* in the cluster amplitudes – even for the full cluster expansion. The BCH expansion terminates because of the special structure of the cluster operators, which are linear combinations of commuting excitation operators of the form (13.2.6) and (13.2.7).

Although the similarity-transformed Hamiltonian is quartic in the cluster amplitudes, the equations for the cluster amplitudes (13.2.32) need not contain all the amplitudes to this order. In Section 13.2.8, we use the cluster-commutation condition (13.2.36) to show that, for a general operator  $\hat{O}$  of particle rank  $m_O$ , the state

$$|n_1 n_2 \dots n_k\rangle = [[\dots [[\hat{O}, \hat{T}_{n_1}], \hat{T}_{n_2}], \dots], \hat{T}_{n_k}]|\text{HF}\rangle \quad (13.2.38)$$

is a linear combination of determinants with excitation ranks  $s$  in the range

$$\sum_{i=1}^k n_i - m_O \leq s \leq \sum_{i=1}^k n_i + m_O - k \quad (13.2.39)$$

where  $n_i$  is the excitation rank of  $\hat{T}_{n_i}$ . Using these conditions, we may set up the following expressions for the CCSD amplitude equations:

$$\begin{aligned} &\langle \mu_1 | \hat{H} | \text{HF} \rangle + \langle \mu_1 | [\hat{H}, \hat{T}_1] | \text{HF} \rangle + \langle \mu_1 | [\hat{H}, \hat{T}_2] | \text{HF} \rangle + \frac{1}{2} \langle \mu_1 | [[\hat{H}, \hat{T}_1], \hat{T}_1] | \text{HF} \rangle \\ &+ \langle \mu_1 | [[\hat{H}, \hat{T}_1], \hat{T}_2] | \text{HF} \rangle + \frac{1}{6} \langle \mu_1 | [[[\hat{H}, \hat{T}_1], \hat{T}_1], \hat{T}_1] | \text{HF} \rangle = 0 \end{aligned} \quad (13.2.40)$$

$$\begin{aligned} &\langle \mu_2 | \hat{H} | \text{HF} \rangle + \langle \mu_2 | [\hat{H}, \hat{T}_1] | \text{HF} \rangle + \langle \mu_2 | [\hat{H}, \hat{T}_2] | \text{HF} \rangle + \frac{1}{2} \langle \mu_2 | [[\hat{H}, \hat{T}_1], \hat{T}_1] | \text{HF} \rangle \\ &+ \langle \mu_2 | [[\hat{H}, \hat{T}_1], \hat{T}_2] | \text{HF} \rangle + \frac{1}{2} \langle \mu_2 | [[\hat{H}, \hat{T}_2], \hat{T}_2] | \text{HF} \rangle \\ &+ \frac{1}{6} \langle \mu_2 | [[[\hat{H}, \hat{T}_1], \hat{T}_1], \hat{T}_1] | \text{HF} \rangle + \frac{1}{2} \langle \mu_2 | [[[\hat{H}, \hat{T}_1], \hat{T}_1], \hat{T}_2] | \text{HF} \rangle \\ &+ \frac{1}{24} \langle \mu_2 | [[[[\hat{H}, \hat{T}_1], \hat{T}_1], \hat{T}_1], \hat{T}_1] | \text{HF} \rangle = 0 \end{aligned} \quad (13.2.41)$$

Whereas the singles occur to fourth order, the doubles appear only *quadratically* in these expressions. Moreover, from (13.2.39), it is easily verified that, except for the singles and the doubles, the amplitudes of the highest excitation level always occur only *linearly* in the coupled-cluster equations. These additional simplifications in the coupled-cluster equations – beyond what is dictated by the termination of the BCH expansion (13.2.37) – occur because of the restrictions on the excitation levels in the projection space  $\langle \mu_i |$ . If instead we had calculated the energy from the variation principle, the expansion of the exponentials would not terminate (except, of course, for the fact that we have only a finite number of electrons to excite from the occupied spin orbitals). The evaluation

of the energy and the amplitudes in such a model is therefore considerably more difficult than in standard coupled-cluster theory.

### 13.2.6 COUPLED-CLUSTER THEORY IN THE CANONICAL REPRESENTATION

In coupled-cluster theory, it is often convenient to work in the canonical representation of the spin orbitals. The Hamiltonian operator is then partitioned into the Fock operator  $\hat{f}$ , the fluctuation potential  $\hat{\Phi}$  and the nuclear–nuclear contribution  $h_{\text{nuc}}$  as discussed in Section 10.4.5:

$$\hat{H} = \hat{f} + \hat{\Phi} + h_{\text{nuc}} \quad (13.2.42)$$

In the canonical representation, the Fock operator may be written in terms of the orbital energies  $\varepsilon_p$  as

$$\hat{f} = \sum_p \varepsilon_p a_p^\dagger a_p \quad (13.2.43)$$

The commutator of the Fock operator with the cluster operator is found to be

$$[\hat{f}, \hat{T}] = \sum_\mu \varepsilon_\mu t_\mu \hat{\tau}_\mu \quad (13.2.44)$$

where  $\varepsilon_\mu$  is the sum of all unoccupied orbital energies minus the sum of all occupied orbital energies of the spin orbitals in  $\hat{\tau}_\mu$ , for example,

$$\varepsilon_{AI} = \varepsilon_A - \varepsilon_I \quad (13.2.45)$$

$$\varepsilon_{AI BJ} = \varepsilon_A - \varepsilon_I + \varepsilon_B - \varepsilon_J \quad (13.2.46)$$

Inserting the Hamiltonian (13.2.42) into the expressions for the energy (13.2.22) and the nonlinear equations (13.2.23), we obtain the following expressions for the coupled-cluster energy and the amplitude equations:

$$E = E_{\text{HF}} + \langle \text{HF} | \hat{\Phi} \left( \hat{T}_2 + \frac{1}{2} \hat{T}_1^2 \right) | \text{HF} \rangle \quad (13.2.47)$$

$$\varepsilon_\mu t_\mu = -\langle \mu | \exp(-\hat{T}) \hat{\Phi} \exp(\hat{T}) | \text{HF} \rangle \quad (13.2.48)$$

where we have used the commutator (13.2.44) and the fact that all higher commutators of the Fock operator vanish since the down rank of the Fock operator is  $\frac{1}{2}$ . As we shall see in Section 14.3, (13.2.47) and (13.2.48) form a convenient starting point for the development of a perturbation theory with the Fock operator as the zero-order Hamiltonian. Moreover, in Section 13.4, we shall use (13.2.48) to set up an efficient iterative scheme for the optimization of the coupled-cluster wave function.

### 13.2.7 COMPARISON OF THE CI AND COUPLED-CLUSTER HIERARCHIES

Owing to the presence of the disconnected clusters, coupled-cluster wave functions truncated at a given excitation level also contain contributions from determinants corresponding to higher-order excitations. The terms that are missing relative to FCI represent higher-order connected clusters and the associated disconnected clusters. By contrast, CI wave functions truncated at the same level contain contributions from determinants only up to this level. Since the disconnected contributions to the energy are significant (and dominant for extended systems), the accuracy