

## Chapter 4

# Coupled Cluster Approach for Molecular Electronic Structure: an Elementary Introduction

Jozef Noga

Institute of Inorganic Chemistry,

Slovak Academy of Sciences, SK - 842 36 Bratislava,

Slovak Republic <sup>1</sup>

### 4.1 Introduction

What is commonly understood under the *ab initio* calculations is solving of the time-independent Schrödinger equation (SE)

$$\hat{H}\Psi = E\Psi, \quad (4.1)$$

without introducing empirical parameters, most often also using the Born-Oppenheimer approximation. Usually, to solve the Eq. 4.1, we use the so called algebraic approximation, when the wave function is expanded in a basis of known functions. Then, rather complicated differential Eq. 4.1 transforms to algebraic equations in matrix form. Most frequently these functions are some types of atomic orbitals – e.

---

<sup>1</sup>L.A. Montero, L.A. Díaz and R. Bader (eds.), Introduction to Advanced Topics of Computational Chemistry, 71 - 103, 2003, ©2003 Editorial de la Universidad de La Habana, Havana.

g. when one solves the problem in independent particle approximation such as the Hartree-Fock (HF). The basis for subsequent correlated calculation can be a set of Slater determinants obtained from the ground state reference HF determinant ( $\Phi_0$ ) through a substitution of occupied orbitals by virtual ones, thus creating excited (or sometimes also called substituted) determinants ( $\Phi_I$ ). The final "exact" wave function can be then expressed via the well known configuration interaction (CI) expansion

$$\Psi = c_0\Phi_0 + \sum_I c_I\Phi_I, \quad (4.2)$$

where all possible excited determinants should be included.

Of course, the above expression "exact" wave function is far from being true, because of approximations introduced. The word "exact" is rather related to a mathematical model using finite basis set expansions. Practical limitation to finite basis expansions causes at least two serious difficulties:

- a) One is faced with the problem of the *basis set superposition error* [1]
- b) Within a finite one-electron basis one is unable to describe the *correlation cusp* correctly [2], i. e. the cusp relation [3]

$$\lim_{r_{12} \rightarrow 0} \left( \frac{\partial \Psi}{\partial r_{12}} \right)_{av} = \frac{1}{2} \Psi(r_{12} = 0), \quad (4.3)$$

with  $r_{12}$  being the inter-electronic distance, is not satisfied.

While the problem listed under a) concerns molecules, and, in particular, it must be emphasized for weak intermolecular interactions, the improper description of the correlation cusp gives rise to a rather slow convergence of the CI expansion Eq. 4.2 towards the exact limit with respect to the extension of the basis. The latter problem concerns both atoms and molecules, and the error caused by this deficiency can be fairly large.

The expansion Eq. 4.2 taken exactly, i. e. including all possible determinants (full CI) for the given finite one electron spin-orbital basis, leads very soon to intractable number of configurations and further truncation is inevitable in practice. Nevertheless, developments either towards approaching the (orbital) basis set limit by using alternative (e. g. two-particle) functions, or towards approaching the full CI limit with possible avoiding most of the terms in the wave function expansion itself, or at least their explicit calculation, are still topical. Some of the trends in both of these directions are subjects of this paper.

There were hardly a need to look for alternative expansions for the wave function, if one could always take the CI expansion Eq. 4.2 in full. Unfortunately, if the

expansion is truncated, the simple CI theory suffers from a serious drawback – the energy does not scale properly with the number of electrons involved – it is not *size extensive*. This can be demonstrated for two noninteracting systems  $A$  and  $B$  associated with the Hamiltonians  $H^A$  and  $H^B$ . For the supersystem we have

$$H^{AB} = H^A + H^B \quad (4.4)$$

Let us restrict the CI to  $n$ -tuple excitations and for simplicity assume that with this the solution for the subsystems is exact. For an orthonormal basis we have ( $M = A, B$ ):

$$\Psi^M = \sum_{i=0}^n c_i^M \Phi_i^M \quad (4.5)$$

$$E^M = \frac{\langle \Psi^M | H^M | \Psi^M \rangle}{\langle \Psi^M | \Psi^M \rangle} = \frac{\sum_{i,j=0}^n (c_i^* c_j)^M \langle \Phi_i^M | H^M | \Phi_j^M \rangle}{\sum_{i=0}^n (c_i^* c_i)^M} \quad (4.6)$$

For simplicity, subscripts  $i, j$  (and later  $k$ ) represent all possible  $i$ -tuple,  $j$ -tuple, and  $k$ -tuple excitations. For noninteracting case  $H^M$  only acts on  $\Psi^M$ . If the CI is unrestricted, the final wave-function is a product  $\Psi^A \Psi^B$ . For this demonstrative example we skip the complication due to antisymmetrization. Hence, we have

$$\Psi^{AB} = \sum_i c_i^A \Phi_i^A \sum_k c_k^B \Phi_k^B \quad (4.7)$$

$$\begin{aligned} E^{AB} &= \frac{\sum_{i,j=0}^n (c_i^* c_j)^A \langle \Phi_i^A | H^A | \Phi_j^A \rangle \sum_k (c_k^* c_k)^B}{\sum_{i=0}^n (c_i^* c_i)^A \sum_k (c_k^* c_k)^B} \\ &+ \frac{\sum_{k,j=0}^n (c_k^* c_j)^B \langle \Phi_k^B | H^B | \Phi_j^B \rangle \sum_i (c_i^* c_i)^A}{\sum_{i=0}^n (c_i^* c_i)^A \sum_k (c_k^* c_k)^B} = E^A + E^B \end{aligned} \quad (4.8)$$

For a case when the CI expansion is restricted to  $n$ -tuple excitations we obtain

$$\Psi_n^{AB} = \sum_{i+k=n} c_i^A c_k^B \Phi_i^A \Phi_k^B \quad (4.9)$$

$$E_n^{AB} = \frac{\sum_{i,j=0}^n (c_i^* c_j)^A \langle \Phi_i^A | H^A | \Phi_j^A \rangle \sum_k^{n-\max(i,j)} (c_k^* c_k)^B}{\sum_{i=0}^n (c_i^* c_i)^A \sum_k^{n-i} (c_k^* c_k)^B} \quad (4.10)$$

$$\begin{aligned} &+ \frac{\sum_{k,j=0}^n (c_k^* c_j)^B \langle \Phi_k^B | H^B | \Phi_j^B \rangle \sum_i^{n-\max(j,k)} (c_i^* c_i)^A}{\sum_{i=0}^n (c_i^* c_i)^A \sum_k^{n-i} (c_k^* c_k)^B} \\ &\neq E^A + E^B \end{aligned} \quad (4.11)$$

Hence for a truncated CI the total energy of noninteracting systems is not simply a sum of the energies for the individual subsystems. This particular failure disappears in the coupled cluster theory, though, not completely free of charge, as will be seen from the forthcoming.

Currently, among the high precision *ab initio* methods the coupled cluster approaches play undoubtedly one of the leading role. In particular, the nondegenerate one reference variants became routine and they are implemented in several of high performance *ab initio* program packages. Yet, the theoretical background remains often unclear for applied quantum chemists, whose primary aim is to get reliable numbers to solve their chemical problems. The purpose of this contribution is to provide a simple introduction to the non-degenerate coupled cluster theory, as well as to elucidate what is hidden behind the so called *non-iterative* corrections. A link between CC, CI and the Many Body Perturbation Theory (MBPT) is outlined. Finally, a recent CC theory which explicitly takes care of the *correlation cusp* is briefly presented.

Development in the CC theory during the last decade has been enormous, including gradient methods, methods to treat open shell single reference, multi-reference theories, methods to calculate excitation energies, ionization potentials, electron affinities e.t.c.; but it would be far beyond the purpose of this paper to write about those methods. This article should possibly give a simple explanation of the CC approach, without going deep into details. More detailed excellent reviews and analyses have been written recently [4, 5, 6, 7] and also some time ago [8, 9] and are recommended to the interested reader. Latest advances in this field were summarized in a special monograph [10].

## 4.2 Basic considerations

In the following we will use a tensor notation for antisymmetrized integrals over arbitrary n-particle operator

$$\begin{aligned}\bar{O}_{\lambda_1\lambda_2\dots\lambda_n}^{\kappa_1\kappa_2\dots\kappa_n} &= \sum_{P_{i_1\dots i_n}} (-1)^{P_{i_1\dots i_n}} O_{\lambda_1\lambda_2\dots\lambda_n}^{\kappa_{i_1}\kappa_{i_2}\dots\kappa_{i_n}} \\ O_{\lambda_1\lambda_2\dots\lambda_n}^{\kappa_{i_1}\kappa_{i_2}\dots\kappa_{i_n}} &= \langle \lambda_1\lambda_2\dots\lambda_n | O | \kappa_{i_1}\kappa_{i_2}\dots\kappa_{i_n} \rangle\end{aligned}\quad (4.12)$$

where the sum runs over all permutations among  $\kappa$ 's, with the corresponding parity  $P_{i_1\dots i_n}$ .

We restrict our consideration to a single determinant reference wave function  $|\Phi\rangle$ . For our purpose it is convenient to work in the second quantized formalism<sup>1</sup> in which the Hamiltonian is expressed as:

$$H = \sum_{\kappa,\lambda} h_{\kappa}^{\lambda} a_{\lambda}^{\kappa} + \frac{1}{4} \sum_{\kappa,\lambda,\mu,\nu} \bar{g}_{\kappa\lambda}^{\mu\nu} a_{\mu\nu}^{\kappa\lambda} \quad (4.13)$$

---

<sup>1</sup>An elementary introduction to second quantization, which is especially designed for quantum chemists is given e. g. in the book by Surján [11] or in the chapter by Paldus [5].

$$h_{\kappa}^{\lambda} = \langle \kappa | h | \lambda \rangle \quad (4.14)$$

$$\bar{g}_{\kappa\lambda}^{\mu\nu} = g_{\kappa\lambda}^{\mu\nu} - g_{\kappa\lambda}^{\nu\mu} \quad (4.15)$$

$$g_{\kappa\lambda}^{\mu\nu} = \langle \kappa\lambda | \frac{1}{r_{12}} | \mu\nu \rangle \quad (4.16)$$

represent one- and two-electron integrals in an orthonormal spin-orbital (MO) basis.

$$a_{\kappa_1 \kappa_2 \dots \kappa_n}^{\lambda_1 \lambda_2 \dots \lambda_n} = a^{\lambda_1} a^{\lambda_2} \dots a^{\lambda_n} a_{\kappa_n} \dots a_{\kappa_2} a_{\kappa_1} \quad (4.17)$$

is an n-body replacement operator, with  $a^{\lambda} = a_{\lambda}^{\dagger}$  and  $a_{\kappa}$  the creation and annihilation operators, respectively. The latter operators satisfy the anti-commutation relations:

$$[a^{\lambda}, a_{\kappa}]_{+} = \delta_{\kappa}^{\lambda}, \quad [a_{\lambda}, a_{\kappa}]_{+} = [a^{\lambda} a^{\kappa}]_{+} = 0 \quad (4.18)$$

Acting of the replacement operator (Eq. 4.17) on a Slater determinant means that spin-orbitals  $|\kappa_i\rangle$  are replaced by  $|\lambda_i\rangle$  if such a replacement is at all possible. Otherwise the result is zero.

We use Greek letters  $\kappa$ ,  $\lambda$ ,  $\mu$ ,  $\nu$  to denote spin-orbitals within a *complete* basis for which the Hamiltonian Eq. 4.13 is exact (assuming the Born-Oppenheimer and non-relativistic approximation). We shall need such a notation in Sec. 4.8. In the following  $i$ ,  $j$ ,  $k$ ,  $\dots$  denote occupied spin-orbitals that build the reference determinant,  $a$ ,  $b$ ,  $c$ ,  $\dots$  denote virtual spin-orbitals resulting from the calculation within the independent particle approximation, while  $p$ ,  $q$ ,  $\dots$  are used for arbitrary spin-orbitals resulting from the calculation within the independent particle approximation using finite one-electron basis set. Virtual orbitals referred to the *complete* basis in Sec. 4.8 are denoted as  $\alpha$ ,  $\beta$ .

When all  $\lambda$ 's in Eq. 4.17 are occupied spin-orbitals and at the same time  $\kappa$ 's are virtual orbitals, the operator is called excitation operator. Acting on a reference determinant, such operators create the orthonormal set of excited determinants used in the CI expansion.

All the manipulations simplify considerably if the operators are expressed in the so called normal product form with respect to the given reference (Fermi vacuum). Within the hole-particle formalism [12] the normal product, which we denote in braces  $\{\dots\}$ , is defined by putting all the creation operators for occupied ( $a^i$ ) and all the annihilation operators for virtual orbitals ( $a_a$ ) to the right, changing the sign at each interchange between neighbors, i. e.

$$\{X_1 Y_1 X_2 \dots X_n \dots Y_m\} = (-1)^P X_1 \dots X_n Y_1 \dots Y_m \quad (4.19)$$

where  $P$  is the parity of the permutations and  $Y$ 's are operators whose action on the reference function (Fermi vacuum) produces zero ( $a^i |\Phi\rangle = a_a |\Phi\rangle = 0$ ), i. e.  $Y \equiv a^i$

or  $a_a$ , unlike  $X \equiv a_i$  or  $a^a$ . In any case the expectation value of the normal product is zero for the reference wave function

$$\langle \Phi | \{ \dots \} | \Phi \rangle = 0. \quad (4.20)$$

The difference between the simple and the normal product of two (either creation or annihilation) operators ( $X, Y$ ) is called the contraction. This corresponds to the reference function expectation value for the given product ( $\langle XY \rangle = \langle \Phi | XY | \Phi \rangle$ )

$$\overline{XY} = XY - \{XY\} = \langle XY \rangle \quad (4.21)$$

Hence, the following is valid:

$$\overline{a^p a_q} = n_p \delta_q^p \quad \overline{a_p a^q} = (1 - n_p) \delta_p^q \quad \overline{a^p a^q} = \overline{a_p a_q} = 0 \quad (4.22)$$

where  $n_p$  is the occupation number for the spin-orbital  $|p\rangle$  in the reference determinant.

Any string of creation and/or annihilation operators can be expressed through normal products, just by subsequent application of anticommutation relations (4.18). This is behind the Wick's theorem [13] which states that an arbitrary string of creation/annihilation operators can be expressed as the normal product, plus normal products with all possible single contractions, plus normal products with all possible double contractions, plus etc. up to normal products with maximum possible number of contractions:

$$\begin{aligned} X_1 X_2 \dots X_n &= \{X_1 X_2 \dots X_n\} + \sum_{j>i} \{ \dots \overline{X_i \dots X_j} \dots \} \\ &+ \sum_{\substack{l>k, j>i \\ k \neq j}} \{ \dots \overline{X_i \dots X_k \dots X_j \dots X_l} \dots \} + \sum_{\substack{\text{triple} \\ \text{contracted}}} \{ \dots \} + \dots + \sum_{\substack{\text{fully} \\ \text{contracted}}} \{ \dots \} \end{aligned} \quad (4.23)$$

Normal products with contractions are determined by evaluating particular contractions after the corresponding operators are put together. Each interchange of the operators changes the sign. Uncontracted operators remain in the normal product.

In the following we will use:

$$\tilde{a}_{\kappa_1 \kappa_2 \dots \kappa_n}^{\lambda_1 \lambda_2 \dots \lambda_n} = \left\{ a_{\kappa_1 \kappa_2 \dots \kappa_n}^{\lambda_1 \lambda_2 \dots \lambda_n} \right\} \quad (4.24)$$

Employing the Wick's theorem to replacement operators in Eq. 4.13 we have:

$$a_\lambda^\kappa = \tilde{a}_\lambda^\kappa + \langle a^\kappa a_\lambda \rangle \quad (4.25)$$

$$\begin{aligned} a_{\mu\nu}^{\kappa\lambda} &= \tilde{a}_{\mu\nu}^{\kappa\lambda} + \langle a^\kappa a_\mu \rangle \tilde{a}_\nu^\lambda + \langle a^\lambda a_\nu \rangle \tilde{a}_\mu^\kappa - \langle a^\kappa a_\nu \rangle \tilde{a}_\mu^\lambda - \langle a^\lambda a_\mu \rangle \tilde{a}_\nu^\kappa \\ &+ \langle a^\lambda a_\nu \rangle \langle a^\kappa a_\mu \rangle - \langle a^\kappa a_\nu \rangle \langle a^\lambda a_\mu \rangle \end{aligned} \quad (4.26)$$

Using this, and Eqs. 4.22, the Hamiltonian Eq. 4.13 can be rewritten:

$$\begin{aligned}
H &= \sum_{\kappa,\lambda} \left( h_{\kappa}^{\lambda} + \sum_i \bar{g}_{i\kappa}^{i\lambda} \right) \tilde{a}_{\lambda}^{\kappa} + \frac{1}{4} \sum_{\kappa,\lambda,\mu,\nu} \bar{g}_{\kappa\lambda}^{\mu\nu} \tilde{a}_{\mu\nu}^{\kappa\lambda} \\
&+ \sum_i h_i^i + \frac{1}{2} \sum_{i,j} \bar{g}_{ij}^{ij} = H_N + \langle \Phi | H | \Phi \rangle
\end{aligned} \tag{4.27}$$

Subtracting the reference energy  $\langle \Phi | H | \Phi \rangle$ , Eq. 4.1 reads

$$H_N \Psi = \Delta E \Psi \tag{4.28}$$

with  $\Delta E$  being directly the correlation energy.  $H_N$  is called the normal-ordered form of the Hamiltonian. One easily recognizes that the matrix elements in the one electron part correspond to the usual Fock matrix elements.

$$\begin{aligned}
H_N &= H - \langle \Phi | H | \Phi \rangle = F_N + W_N \\
F_N &= \sum_{\kappa,\lambda} f_{\kappa}^{\lambda} \tilde{a}_{\lambda}^{\kappa}; \quad W_N = \frac{1}{4} \sum_{\kappa,\lambda,\mu,\nu} \bar{g}_{\kappa\lambda}^{\mu\nu} \tilde{a}_{\mu\nu}^{\kappa\lambda}
\end{aligned} \tag{4.29}$$

A product of  $n$  normal ordered products can be expressed again in a normal ordered form by employing the generalized Wick's theorem which resembles the ordinary Wick's theorem [12] when we replace  $X_i$  by a string of operators

$$\mathcal{X}_i \equiv X_{i_1} X_{i_2} \dots X_{i_m} \tag{4.30}$$

$$\begin{aligned}
\{\mathcal{X}_1\}\{\mathcal{X}_2\}\dots\{\mathcal{X}_n\} &= \{\mathcal{X}_1 \mathcal{X}_2 \dots \mathcal{X}_n\} + \sum_{j>i} \{\dots \overline{\mathcal{X}_i \dots \mathcal{X}_j} \dots\} \\
&+ \sum_{l>k,j>i} \{\dots \overline{\mathcal{X}_i \dots \mathcal{X}_k \dots \overline{\mathcal{X}_j \dots \mathcal{X}_l}} \dots\} + \dots + \sum_{\substack{\text{fully} \\ \text{contracted}}} \{\dots\}
\end{aligned} \tag{4.31}$$

where only contractions among operators from different original normal products are allowed. Since  $\mathcal{X}_i$  are normal ordered products, contractions within them give zero. The right hand side of Eq. 4.31 can be separated to so called *connected* and *disconnected* terms. *Connected* are terms which contain at least  $n - 1$  contractions from  $n - 1$  different pairs of strings. In such case all the original normal products are joined somehow together. Other terms are *disconnected*.

Let us have  $n$ -particle excitation operators associated with amplitudes ( $t$ ) [12]:

$$T_n = \frac{1}{n!} \sum_{\substack{i_1, \dots, i_n \\ a_1, \dots, a_n}} t_{a_1, \dots, a_n}^{i_1, \dots, i_n} \tilde{a}_{i_1, \dots, i_n}^{a_1, \dots, a_n} \tag{4.32}$$

Of course, excitation operators commute among themselves. Any product of these operators can only be fully *disconnected* since there is no possible nonzero contraction

among them, i. e. only the first term on the r.h.s in Eq. 4.31 remains, i. e. the product of such excitation operators is automatically a normal product. A product of an arbitrary normal ordered m-particle operator

$$O_m = \frac{1}{m!} \sum_{\substack{p_1 \dots p_m \\ q_1 \dots q_m}} \bar{O}_{q_1 \dots q_m}^{p_1 \dots p_m} \tilde{a}_{p_1 \dots p_m}^{q_1 \dots q_m} \quad (4.33)$$

from the left, with n-particle excitation operator results in *disconnected* (subscript "D") and *connected* ("C") terms, i. e.

$$\begin{aligned} O_m T_n &= (O_m T_n)_D + (O_m T_n)_C \\ &= \frac{1}{n!m!} \sum_{\substack{p_1 \dots p_m \\ q_1 \dots q_m}} \sum_{\substack{i_1, \dots, i_n \\ a_1, \dots, a_n}} \bar{O}_{q_1 \dots q_m}^{p_1 \dots p_m} t_{a_1 \dots a_n}^{i_1, \dots, i_n} \tilde{a}_{p_1 \dots p_m}^{q_1 \dots q_m a_1 \dots a_n} + (O_m T_n)_C \end{aligned} \quad (4.34)$$

where the *disconnected* term is an (m+n)-particle operator while *connected* terms include (m+n-1)-, (m+n-2)-, ... (|m-n|)-particle operators. Since  $O$  and  $t$  are scalar quantities, and the normal ordered replacement operator is invariant with respect to simultaneous permutations of pairs of lower and upper indices, we can write

$$\begin{aligned} (O_m T_n)_D &= \frac{1}{n!m!} \sum_{\substack{p_1 \dots p_m \\ q_1 \dots q_m}} \sum_{\substack{i_1, \dots, i_n \\ a_1, \dots, a_n}} \bar{O}_{q_1 \dots q_m}^{p_1 \dots p_m} t_{a_1 \dots a_n}^{i_1, \dots, i_n} \tilde{a}_{p_1 \dots p_m}^{q_1 \dots q_m a_1 \dots a_n} \\ &= \frac{1}{m!n!} \sum_{\substack{i_1, \dots, i_n \\ a_1, \dots, a_n}} \sum_{\substack{p_1 \dots p_m \\ q_1 \dots q_m}} t_{a_1 \dots a_n}^{i_1, \dots, i_n} O_{q_1 \dots q_m}^{p_1 \dots p_m} \tilde{a}_{i_1 \dots i_n}^{a_1 \dots a_n p_1 \dots p_m} = T_n O_m \end{aligned} \quad (4.35)$$

This follows from the fact that in accord with Eq. 4.22 a product of two replacement operators  $\tilde{a}_{p_1 p_2 \dots}^{q_1 q_2 \dots} \tilde{a}_{r_1 r_2 \dots}^{s_1 s_2 \dots}$  contains connected terms only if at least one index  $q$  denoting occupied spin-orbital coincides with  $r$  or at least one  $p$  denoting virtual spin-orbitals coincides with  $s$ , which is never true for a product of an excitation operator with any other operator. Hence, if  $T$  is from the left, the whole product only contains *disconnected* terms.

Using the generalized Wick's theorem for a product of  $O_m$  with two excitation operators we obtain:

$$\begin{aligned} O_m T_n T_{n'} &= [(O_m T_n)_D T_{n'}]_D + [(O_m T_n)_C T_{n'}]_D + [(O_m T_n)_C T_{n'}]_C \\ &\quad + [(O_m T_n)_C T_{n'}]_C \end{aligned} \quad (4.36)$$

As *connected*  $(O_m T_n)_C$  is a new normal ordered operator (e. g.  $O'_m$ ), in analogy with Eq. 4.35 we can write:

$$[(O_m T_n)_C T_{n'}]_D = (O'_m T_{n'})_D = T_{n'} O'_m = T_{n'} (O_m T_n)_C \quad (4.37)$$

These relations are also valid for a global excitation operator:

$$T = \sum_n T_n \quad (4.38)$$



where  $n$  of course must not exceed the number of electrons. In practice, however, it is much more limited.

Note that in accord with Eq. 4.20

$$\langle \Phi | T O_m | \Phi \rangle = \langle \Phi | (O_m T)_D | \Phi \rangle = 0, \quad (4.39)$$

because such a product cannot be fully contracted, i. e. it always contains a normal ordered replacement operator whose expectation value is zero.

### 4.3 The Coupled Cluster Ansatz

Under the cluster expansion of a quantity  $A$  connected with an  $n$ -particle system the expression

$$A = A_o + \sum_i A_i + \sum_{i>j} A_{ij} + \sum_{i>j>k} A_{ijk} + \dots, \quad (4.40)$$

is understood, where  $A_o$  is the value for independent particles,  $A_i, A_{ij}, A_{ijk} \dots$ , are one-, two-, three- ... body effects. If  $A$  is additively separable, one can write a similar expansion for the multiplicatively separable quantity

$$B = \exp(A). \quad (4.41)$$

The idea of the cluster expansion characterized by Eq. 4.41 was introduced in the classical statistical mechanics [14]. Exploiting of a similar exponential structure for the wave operator<sup>2</sup> in the framework of the Many Body Perturbation Theory (MBPT) was pointed out by Gell-Mann and Low [15], Hubbard [16] and Hugenholtz [17]. These considerations which became a basis for the coupled cluster theory. A common feature of all CC methods is the exponential ansatz for the wave function

$$|\Psi\rangle = e^S |\Phi\rangle, \quad \langle \Psi | = \langle \Phi | e^{S^\dagger} \quad (4.42)$$

where in *conventional* approaches  $S = T$  is usually a global excitation operator (Eq. 4.38), and, as mentioned in the foregoing, we restrict ourselves to a single reference CC with  $|\Phi\rangle$  being the reference function. The global excitation operator  $T$  consists of single- ( $T_1$ ), double- ( $T_2$ ), triple- ( $T_3$ ), e.t.c.  $n$ -excitation operators.

Let us now return to noninteracting systems described in the **Introduction** by Eqs. 4.4-4.11. Assuming a truncation on  $T$  to  $n$ -excitations, and using the fact that excitation operators commute, we obtain

$$\Psi^{AB} = \Psi_n^{AB} = e^{T_A} e^{T_B} \Phi^A \Phi^B = e^{T_A + T_B} \Phi^A \Phi^B \quad (4.43)$$

---

<sup>2</sup>An operator ( $\Omega$ ) through which the final wave function is created from the reference ( $\Psi = \Omega\Phi$ ).

$$\begin{aligned}
E^{AB} &= E_n^{AB} = \frac{\langle e^{T_A} \Phi^A | H_A | e^{T_A} \Phi^A \rangle \langle e^{T_B} \Phi^B | e^{T_B} \Phi^B \rangle}{\langle e^{T_A} \Phi^A | e^{T_A} \Phi^A \rangle \langle e^{T_B} \Phi^B | e^{T_B} \Phi^B \rangle} \\
&+ \frac{\langle e^{T_B} \Phi^B | H_B | e^{T_B} \Phi^B \rangle \langle e^{T_A} \Phi^A | e^{T_A} \Phi^A \rangle}{\langle e^{T_A} \Phi^A | e^{T_A} \Phi^A \rangle \langle e^{T_B} \Phi^B | e^{T_B} \Phi^B \rangle} = E^A + E^B, \quad (4.44)
\end{aligned}$$

which is a sign for *size extensive* method. Generally, for interacting systems one can say that the method is *size extensive* if merely such terms survive in the pertinent equations that linearly depend on the number of particles.

For the many fermion system the cluster expansion ansatz was first suggested by Coester and Kümmel [18]. These general ideas were formulated into the systematic procedure by Čížek [12], yielding explicit equations to determine the amplitudes within the single reference approach.

Let us assume intermediate normalization, which in our case directly follows from the orthogonality of the basis,

$$\langle \Phi | \Psi \rangle = 1. \quad (4.45)$$

Expanding  $\exp(T)$  in a Taylor series the wave function defined by Eq. 4.42 reads:

$$|\Psi\rangle = |\Phi\rangle + T|\Phi\rangle + \frac{1}{2}T^2|\Phi\rangle + \frac{1}{3!}T^3|\Phi\rangle + \dots \quad (4.46)$$

If we define

$$C_0 = 1, \quad (4.47)$$

$$C_1 = T_1, \quad (4.48)$$

$$C_2 = T_2 + \frac{1}{2}T_1^2, \quad (4.49)$$

$$C_3 = T_3 + T_1T_2 + \frac{1}{3!}T_1^3, \quad (4.50)$$

$$C_4 = T_4 + T_1T_3 + \frac{1}{2}T_2^2 + \frac{1}{2}T_2T_1^2 + \frac{1}{4!}T_1^4, \quad (4.51)$$

⋮

then the CI expansion Eq. 4.2 with single reference can be rewritten as:

$$|\Psi\rangle = (1 + C_1 + C_2 + C_3 + C_4 + \dots)|\Phi\rangle = C|\Phi\rangle. \quad (4.52)$$

Of course, expansion Eq. 4.46 or 4.52 terminates after all possible excitations are exhausted, since an occupied spin orbital cannot be twice annihilated without creating it in-between, and the same is valid for creating the virtual orbital twice without annihilating it in-between (cf. Ref. [11]). CI and CC expansions are hence fully equivalent as soon as no approximations are introduced in them. In practical applications one truncates the CI wave function after a certain rank of excitations, while in the coupled cluster approach the truncation is limited to  $T$ , or, in some

cases Eq. 4.46 is restricted, as well. Through the exponential expansion, however, a substantial portion of higher excitations is included, without a need to treat them explicitly.

Projecting the Schrödinger equation (Eq. 4.28) from the left onto the reference state, and onto the subspaces of various excited determinants, one gets a system of either nonlinear equations (for CC)

$$\langle \Phi | H_N e^T | \Phi \rangle = \Delta E = E - \langle \Phi | H | \Phi \rangle, \quad (4.53)$$

$$\langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} (H_N - \Delta E) e^T | \Phi \rangle = 0, \quad (4.54)$$

or linear equations (for CI)

$$\langle \Phi | H_N C | \Phi \rangle = \Delta E, \quad (4.55)$$

$$\langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} (H_N - \Delta E) C | \Phi \rangle = 0. \quad (4.56)$$

These equations determine the energy and the amplitudes. In practice, Eqs. 4.54 (or 4.56) are solved for those excitation determinants (from the left) which can be created by the actual (i. e. mostly truncated) operator  $T$  (or  $C$ ). This means that for  $T \approx T_2$ , Eq. 4.54 is only projected to double-excitations, for  $T \approx T_1 + T_2$ , the projection is made to single- and double-excitations e.t.c.

Before we will demonstrate where is the origin of the *size extensivity* failure of the truncated CI, let us mention the price which one pays in CC for the size extensivity. Actually, Eqs. 4.53 and 4.54 correspond to first projecting the SE from the left onto  $\langle \Phi | (1 + T^\dagger)$ , making the resulting (non-symmetric) energy functional

$$\frac{\langle \Phi | (1 + T^\dagger) H_N e^T | \Phi \rangle}{\langle \Phi | (1 + T^\dagger) e^T | \Phi \rangle} = \Delta E, \quad (4.57)$$

stationary with respect to variations of  $T^\dagger$ , i. e. with respect to any change in amplitudes of  $T^\dagger$ . But the resulting equations determine the amplitudes of  $T$ , not the amplitudes of  $T^\dagger$ . If the cluster operator  $T$  were complete, then due to the validity of Schrödinger equation,

$$\langle \Phi | (1 + T^\dagger) (H_N - \Delta E) e^T | \Phi \rangle = \langle \Phi | e^{T^\dagger} (H_N - \Delta E) e^T | \Phi \rangle. \quad (4.58)$$

Inserting the identity resolution (for  $N_{occ}$  being number of occupied spin-orbitals)

$$\sum_{n=0}^{N_{occ}} \sum_{\substack{i_1 \dots i_n \\ \alpha_1 \dots \alpha_n}} \tilde{a}_{i_1 \dots i_n}^{\alpha_1 \dots \alpha_n} | \Phi \rangle \langle \Phi | \tilde{a}_{\alpha_1 \dots \alpha_n}^{i_1 \dots i_n} = 1 \quad (4.59)$$

we obtain

$$\begin{aligned} & \sum_{n=0}^{N_{occ}} \sum_{\substack{i_1 \dots i_n \\ \alpha_1 \dots \alpha_n}} \langle \Phi | (1 + T^\dagger) \tilde{a}_{i_1 \dots i_n}^{\alpha_1 \dots \alpha_n} | \Phi \rangle \langle \Phi | \tilde{a}_{\alpha_1 \dots \alpha_n}^{i_1 \dots i_n} (H_N - \Delta E) e^T | \Phi \rangle = \\ & = \sum_{n=0}^{N_{occ}} \sum_{\substack{i_1 \dots i_n \\ \alpha_1 \dots \alpha_n}} \langle \Phi | e^{T^\dagger} \tilde{a}_{i_1 \dots i_n}^{\alpha_1 \dots \alpha_n} | \Phi \rangle \underbrace{\langle \Phi | \tilde{a}_{\alpha_1 \dots \alpha_n}^{i_1 \dots i_n} (H_N - \Delta E) e^T | \Phi \rangle}_{=0} = 0. \end{aligned} \quad (4.60)$$

For exact  $T$ , Eqs. 4.53, 4.54 are satisfied for all possible excitations within the actual basis set and the last expression is always equal to zero. In this case, independently determined amplitudes of  $T^\dagger$  are hermitian conjugates to the calculated amplitudes of  $T$ , i. e. both sets of amplitudes would minimize the energy functional Eq. 4.57.

For  $T$  truncated to certain excitations (e. g.  $T = T_1 + T_2$ ) Eq. 4.54 is not fulfilled (and not solved) for the missing excitations. Hence, Eq. 4.58 does not hold, because there are nonzero terms in Eq. 4.60. Consequently, solving independently for  $T^\dagger$ , one would paradoxically obtain  $(T^\dagger)^\dagger \neq T$ . The method is not anymore variational since calculated  $T$  does not minimize the energy functional, which is the mentioned price to pay for the *size extensivity* (see also Sec. 4.7).

Unlike for CC, the CI Eqs. 4.55, 4.56 result from the differentiation of a symmetric functional,

$$\frac{\langle \Phi | (C^\dagger) H_N C | \Phi \rangle}{\langle \Phi | (C^\dagger) C | \Phi \rangle} = \Delta E, \quad (4.61)$$

even for a truncated  $C$ , which implies that the method is always variational.

To demonstrate the *size extensivity* problem let us restrict to double excitations in operators, for simplicity. It means that  $T \approx T_2$  (CCD) and  $C \approx 1 + C_2$  (CID). Eqs. 4.55 and 4.56 then yield:

$$\Delta E = \langle \Phi | H_N (1 + C_2) | \Phi \rangle = \langle \Phi | H_N C_2 | \Phi \rangle, \quad (4.62)$$

$$0 = \langle \Phi | \tilde{a}_{ab}^{ij} H_N (1 + C_2) | \Phi \rangle - \underbrace{\Delta E \langle \Phi | \tilde{a}_{ab}^{ij} C_2 | \Phi \rangle}_{\sim N^2 \text{ depend. in } \Delta E}, \quad (4.63)$$

where the second (renormalization) term in Eq. 4.63 causes an improper scaling of the  $C$ -amplitudes, and, consequently, terms that scale quadratically with the number of particles ( $N$ ) are introduced to the energy. Since the renormalization term remains uncanceled, it causes an improper scaling of the energy with the number of particles involved.

From Eqs. 4.53 and 4.54 we get:

$$\Delta E = \langle \Phi | H_N (1 + T_2) | \Phi \rangle = \langle \Phi | H_N T_2 | \Phi \rangle, \quad (4.64)$$

$$0 = \langle \Phi | \tilde{a}_{ab}^{ij} H_N (1 + T_2) | \Phi \rangle + \frac{1}{2} \langle \Phi | \tilde{a}_{ab}^{ij} H_N T_2 T_2 | \Phi \rangle - \Delta E \langle \Phi | \tilde{a}_{ab}^{ij} T_2 | \Phi \rangle. \quad (4.65)$$

The expansion of  $e^{T_2}$  is terminated after the quadratic term, because the Hamiltonian is a two-particle operator (at most in our case), hence nonzero matrix elements arise just for cases where the resulting determinants on the right and the left side around the operator differ at most by two excitations.

In accord with Sec. 4.2, in Eqs. 4.64 and 4.65 we can distinguish *connected* and *disconnected* terms. Disconnected terms in Eq. 4.64 vanish trivially in accord with Eq. 4.39, i.e.

$$\Delta E = \langle \Phi | (H_N T_2)_C | \Phi \rangle + \langle \Phi | (H_N T_2)_D | \Phi \rangle = \langle \Phi | (H_N T_2)_C | \Phi \rangle. \quad (4.66)$$

Employing the resolution of identity one can easily show that disconnected term in the linear part of Eq. 4.65 vanishes as well.

$$\begin{aligned} \langle \Phi | \tilde{a}_{ab}^{ij} H_N T_2 | \Phi \rangle &= \langle \Phi | \tilde{a}_{ab}^{ij} (H_N T_2)_C | \Phi \rangle + \langle \Phi | \tilde{a}_{ab}^{ij} (H_N T_2)_D | \Phi \rangle \\ \langle \Phi | \tilde{a}_{ab}^{ij} (H_N T_2)_D | \Phi \rangle &= \langle \Phi | \tilde{a}_{ab}^{ij} T_2 H_N | \Phi \rangle \\ &= \sum_{n=0}^{N_{occ}} \sum_{\substack{i_1 \dots i_n \\ a_1 \dots a_n}} \langle \Phi | \tilde{a}_{ab}^{ij} T_2 \tilde{a}_{i_1 \dots i_n}^{a_1 \dots a_n} | \Phi \rangle \langle \Phi | \tilde{a}_{a_1 \dots a_n}^{i_1 \dots i_n} H_N | \Phi \rangle = 0, \end{aligned} \quad (4.67)$$

because the first term in the product vanishes for  $n \neq 0$  and the second term vanishes for  $n = 0$  ( $\langle \Phi | H_N | \Phi \rangle = 0$ ). The quadratic term can be written as

$$\frac{1}{2} H_N T_2 T_2 = \frac{1}{2} (H_N T_2 T_2)_C + \frac{1}{2} (H_N T_2 T_2)_D. \quad (4.68)$$

and in accord with Eq. 4.36 the second term on r.h.s. can be further decomposed as

$$\frac{1}{2} (H_N T_2 T_2)_D = \frac{1}{2} [(H_N T_2)_D T_2]_D + [(H_N T_2)_C T_2]_D, \quad (4.69)$$

where we have introduced a factor of 2 in the second term, because both  $T_2$  operators are identical. Following Eq. 4.37 we can write

$$\langle \Phi | \tilde{a}_{ab}^{ij} [(H_N T_2)_C T_2]_D | \Phi \rangle = \langle \Phi | \tilde{a}_{ab}^{ij} T_2 (H_N T_2)_C | \Phi \rangle. \quad (4.70)$$

Inserting a full projector (Eq. 4.59) yields:

$$\begin{aligned} &\sum_{n=0}^{N_{occ}} \sum_{\substack{i_1 \dots i_n \\ a_1 \dots a_n}} \langle \Phi | \tilde{a}_{ab}^{ij} T_2 \tilde{a}_{i_1 \dots i_n}^{a_1 \dots a_n} | \Phi \rangle \langle \Phi | \tilde{a}_{a_1 \dots a_n}^{i_1 \dots i_n} | (H_N T_2)_C | \Phi \rangle = \\ &\langle \Phi | \tilde{a}_{ab}^{ij} T_2 | \Phi \rangle \langle \Phi | (H_N T_2)_C | \Phi \rangle = \langle \Phi | \tilde{a}_{ab}^{ij} T_2 | \Phi \rangle \Delta E, \end{aligned} \quad (4.71)$$

where we have used the identity Eq. 4.66. Only the term with the reference state survives, since  $\langle \Phi | \tilde{a}_{ab}^{ij} T_2 \tilde{a}_{i_1 \dots i_n}^{a_1 \dots a_n} | \Phi \rangle$  is evidently zero for  $n \neq 0$  due to the orthogonality of the excited determinants. Similarly, one can show that the first term on the r.h.s. of Eq. 4.69 vanishes. As a result, the last (renormalization) term in Eq. 4.65 is exactly canceled by the disconnected part of the quadratic term.

What remains from Eqs. 4.64 and 4.65 are connected terms

$$\Delta E = \langle \Phi | (H_N T_2)_C | \Phi \rangle, \quad (4.72)$$

$$0 = \langle \Phi | \tilde{a}_{ab}^{ij} (H_N (1 + T_2))_C | \Phi \rangle + \frac{1}{2} \langle \Phi | \tilde{a}_{ab}^{ij} (H_N T_2 T_2)_C | \Phi \rangle, \quad (4.73)$$

which scale properly with the size of the system (number of particles). It is also evident that the renormalization term is formally canceled out with terms which originate in contributions from quadruple excitations (note that  $T_2T_2$  is hidden in  $C_4$ ), i. e. it cannot be canceled within the CID.

## 4.4 Standard Coupled Cluster Hierarchy

The result that CC equations contain just connected terms is valid generally, not only for the approximation  $T \approx T_2$ . It can be obtained more elegantly than by Eqs. 4.64-4.71, which, of course, was not a general proof. We just wanted to demonstrate where is the crucial difference between "corresponding" approximated CI and CC approaches. Actually, the Schrödinger equation Eq. 4.28 can be first multiplied from the left by  $e^{-T}$ :

$$e^{-T} H_N e^T |\Phi\rangle = \Delta E |\Phi\rangle, \quad (4.74)$$

and subsequently projected to various excitation subspaces. Using the operator identity [19]:

$$e^{-T} H_N e^T = \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{[\dots [H_N, T], T], \dots, T]}_n = (H_N e^T)_C, \quad (4.75)$$

one finally arrives at the connected series simply by subsequent using of Eqs. 4.35 and 4.37

$$[H_N, T] = H_N T - T H_N = (H_N T)_C + (H_N T)_D - T H_N = (H_N T)_C, \quad (4.76)$$

$$\begin{aligned} [[H_N, T], T] &= (H_N T)_C T - T (H_N T)_C = ((H_N T)_C T)_C \\ &+ ((H_N T)_C T)_D - T (H_N T)_C = (H_N T^2)_C, \end{aligned} \quad (4.77)$$

$$\begin{aligned} &\vdots \\ \underbrace{[\dots [H_N, T], \dots T]}_n &= (H_N T^{n-1})_C T - T (H_N T^{n-1})_C = ((H_N T^{n-1})_C T)_C \\ &+ ((H_N T^{n-1})_C T)_D - T (H_N T^{n-1})_C = (H_N T^n)_C, \end{aligned} \quad (4.78)$$

Moreover, Eq. 4.75 terminates after the fourth power in  $T$ , because there are maximum four (two creation and two annihilation) operators from  $H_N$  available for contractions with  $T$ 's. Now we can rewrite Eqs. 4.53 and 4.54 to:

$$\langle \Phi | (H_N e^T)_C | \Phi \rangle = \Delta E, \quad (4.79)$$

$$\langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} (H_N e^T)_C | \Phi \rangle = 0, \quad (4.80)$$

**Table 1.**  
Standard single reference coupled cluster variants.

CC variant	$\approx \exp(T)$ in projection of $(W_N \exp(T))_C  \Phi\rangle$ onto:		
	$\langle \Phi   \tilde{a}_a^i$	$\langle \Phi   \tilde{a}_{ab}^{ij}$	$\langle \Phi   \tilde{a}_{abc}^{ijk}$
LCCD <sup>a,b,c</sup>	0	$1 + T_2$	0
CCD <sup>a,b,d</sup>	0	$\exp(T_2)$	0
CCSD <sup>e,f</sup>	$\exp(T_2 + T_1)$	$\exp(T_2 + T_1)$	0
QCISD <sup>b,g</sup>	$T_2 + T_1 + T_1 T_2$	$\exp(T_2) + T_1$	0
CCSDT-1a <sup>e,h</sup>	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1) + T_3$	$T_2$
CCSDT-1b <sup>e,i</sup>	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1 + T_3)$	$T_2$
CCSDT-2 <sup>e,j</sup>	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2)$
CCSDT-3 <sup>e,j</sup>	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1)$
CCSDT-4 <sup>e,k</sup>	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1) + T_3$
CCSDT <sup>e,k</sup>	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1 + T_3)$	$\exp(T_2 + T_1 + T_3)$

<sup>a</sup>  $\langle \Phi | (F_N e^T)_C | \Phi \rangle \approx \langle \Phi | (F_N T_2)_C | \Phi \rangle$     <sup>b</sup>  $\Delta E = \langle \Phi | (W_N T_2)_C | \Phi \rangle$

<sup>c</sup> Linearized CCD, also L-CPMET [12], CEPA-0 [25, 26], D-MBPT( $\infty$ ) [27]

<sup>d</sup> known also as CPMET [12]; first large scale applications [21, 22, 23].

<sup>e</sup>  $\Delta E = \langle \Phi | [W_N (T_2 + \frac{1}{2} T_1 T_1)]_C | \Phi \rangle$ ;  $\langle \Phi | (F_N e^T)_C | \Phi \rangle$  taken in full  
for  $T=T_1+T_2$  (CCSD) and  $T=T_1+T_2+T_3$  (CCSDT)

<sup>f</sup> Ref [24]

<sup>g</sup> Quadratic CI [28] is actually a CC variant cf. [29]

HF based method;  $\langle \Phi | (F_N e^T)_C | \Phi \rangle \approx \langle \Phi | (F_N T)_C | \Phi \rangle$  for  $T=T_1+T_2$

<sup>h</sup> Ref. [30];    <sup>i</sup> Ref. [31]    <sup>j</sup> Ref. [32];    <sup>k</sup> Refs. [33, 34]    <sup>l</sup> Refs. [20, 35]

If HF reference is used – due to the Brillouin theorem – the doubly excited determinants in the wave function expansion play the most important role besides the reference function. CC equations with  $T \approx T_2$  were given already in the fundamental papers by Čížek [12], together with some model calculations on semi-empirical level. The method was called Coupled Pair Many Electron Theory (CPMET). Although an extended CPMET (ECPMET) method with approximate inclusion of single and triple excitations has been applied at *ab initio* level for very small systems quite soon [20], the large scale *ab initio* programs and applications of CPMET were reported much later in various groups [21, 22, 23]. The method had been used under the name Coupled Cluster Doubles (CCD) and it became a starting point for subsequent developments creating a hierarchy of CC methods which were gradually approaching the full CI limit. These CC variants differ in the approximation of  $e^T$  in the  $(W_N e^T)_C$  expansion for projections of the Eq. 4.28 onto different subspaces of excited determinants.

A list of the most frequent methods is given in Table 1. Besides variants included in this table, approaches which include the cluster operator of quadruple excitations,  $T_4$ , either approximatively [36], or in full [37] were reported. The practical use of the last two methods was limited to very model systems, and routine applications are hardly expectable.

It is beyond the purpose of this paper to give a detailed explicit form of the CC equations. One could of course evaluate the matrix elements using the standard Slater rules [38]. This would be rather tedious if we realize that even for the simple CCD case we deal with quadruple substituted determinants. Generally, the use of replacement operator algebra (see e. g. [5]) simplifies the task considerably, and, in particular for the single reference case the use of diagrammatic techniques [12, 35, 39, 40] proved to be extremely helpful.

Let us at least briefly mention how the CC equations are solved in practice for the ground state. We start from the Hamiltonian  $H_N$  in which we recognize the Møller-Plesset partition [41], with  $W_N$  being the perturbation. Then the logical suggestion for practical calculations is an iterative procedure based on moving the diagonal part of  $F_N$  to one side and the rest to other side of the equations. We use  $F_N^d$  for diagonal and  $F_N^{nd}$  for non-diagonal part of  $F_N$ . In  $(k+1)$ st iteration we solve:

$$-\langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} (F_N^d T^{(k+1)})_C | \Phi \rangle = \langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} [(F_N^{nd} + W_N) e^{T^{(k)}}]_C | \Phi \rangle, \quad (4.81)$$

$$\Delta E^{(k+1)} = \langle \Phi | (H_N e^{T^{(k+1)}})_C | \Phi \rangle. \quad (4.82)$$

Moreover, the nonlinear terms are expected to be much less important than the linear ones, which suggests quasi-linearization of the problem. If we define an intermediate quantity:



$$G(T^{(k)}) = \sum_{n=0}^3 \frac{1}{(n+1)!} \left[ (F_N^{nd} + W_N)(T^{(k)})^n \right]_C, \quad (4.83)$$

in each iteration we have a system of quasi-linear equations:

$$-\langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} (F_N^d T^{(k+1)})_C | \Phi \rangle = \langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} [F_N^{nd} + W_N + G(T^{(k)}) T^{(k)}]_C | \Phi \rangle. \quad (4.84)$$

The Eq. 4.84 is computationally equivalent to what is done in direct CI [42]. The only extra computational demand with respect to CI is evaluation of the intermediate quantities  $G(T)$ . This is, however, less computationally demanding than Eq. 4.84 itself.

Using of Eq. 4.83 actually means that one can handle higher excitations without explicitly having matrix elements due to those configurations. This is a great advantage of CC methods with respect to CI. Actually, in each individual step the computation can be restricted to a product of two matrices. This follows from the fact that the  $T$  operators cannot be contracted among themselves, i. e. the nonlinear terms in Eq. 4.83 can be further factorized. Hence, e. g.  $(W_N(T^{(k)}))^3$  is obtained in a chain:

$$\underbrace{(W_N T^{(k)})_C}_{G_1} \rightarrow \underbrace{(G_1 T^{(k)})_C}_{G_2} \rightarrow \underbrace{(G_2 T^{(k)})_C}_{G_3} = (W_N(T^{(k)}))^3 \quad (4.85)$$

Actually, in all standard methods a contraction with  $T_2$  only appears once in this chain, while additional contractions with  $T_1$  can be very easily evaluated, so to say free of charge in comparison with the computationally most demanding step.

Through the above factorization e. g. the computational complexity for CCSD is the same as for the CISD, just the timing is slightly higher, because of the additional contraction in Eqs. 4.85. At the same time in CCSD one covers high portion of the contributions from triple and especially quadruple excitations, despite those are not treated explicitly. Similarly, the complexity of CCSDT is not worse than for CISDT; but CCSDT covers even quintuple-excitations to a large extent.

Rough estimates of the computational demands for various methods are given in Table 2.

## 4.5 Relation to the Perturbation Theory

As mentioned in the foregoing, the idea of the coupled cluster ansatz is closely related to an observation that the wave operator in MBPT can be expressed in exponential form [16]. In this section we will try to give a simple insight into the relation of both methods and we will show how the individual orders of MBPT can be obtained within the CC iterative procedure.

**Table 2.**  
Formal accuracy of the energy in terms of MBPT and  
computational demands for various CC methods.

Method	accuracy <sup>a</sup>	approximate dependence <sup>b</sup>
CCD	3	$N_o^2 N_v^4 / \text{it}$ , $N_o^3 N_v^3 / \text{it}$
CCSD	3	$N_o^2 N_v^4 / \text{it}$ , $N_o^3 N_v^3 / \text{it}$
CCSDT-1	4	$N_o^3 N_v^4 / \text{it}$
CCSD(T) <sup>c</sup>	4	$N_o^2 N_v^4 / \text{it} + N_o^3 N_v^4$ (for (T))
CCSDT	4	$N_o^4 N_v^4 / \text{it}$

<sup>a</sup> Energy covers all terms which appear in PT up to the given order (4.5) which does not mean *precision* of the method

<sup>b</sup>  $N_o \equiv$  number of occupied;  $N_v \equiv$  number of virtual orbitals  
Formally, the most demanding step within the iteration.  
The permutation symmetry factor is not included.

<sup>c</sup> See Sec. 4.6

We start from  $H_N$  in which  $W_N$  is treated as a perturbation with the strength  $\lambda$  (Møller-Plesset partition in which we assume the validity of Brillouin theorem, i. e.  $f_a^i = 0$ ):

$$H_N = F_N + \lambda W_N \quad (4.86)$$

Within the perturbation theory (PT) the correlation energy and the wave function are expanded in a power series of  $\lambda$ :

$$\Delta E = E - E^{(0)} = \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \lambda^4 E^{(4)} + \dots \quad (4.87)$$

$$|\Psi\rangle = |\Phi^{(0)}\rangle + \lambda |\Phi^{(1)}\rangle + \lambda^2 |\Phi^{(2)}\rangle + \lambda^3 |\Phi^{(3)}\rangle + \lambda^4 |\Phi^{(4)}\rangle + \dots \quad (4.88)$$

Equivalently to Eq. 4.88 we can expand the cluster excitation operator  $T$  (note that unlike in Sec. 4.4 now the superscript " $k$ " denotes order in PT, not the iteration):

$$T = T^{(0)} + \lambda T^{(1)} + \lambda^2 T^{(2)} + \lambda^3 T^{(3)} + \lambda^4 T^{(4)} + \dots \quad (4.89)$$

which means that this expansion is valid for each particular excitation operator from the global T. Using Eq. 4.46 we obtain:

$$|\Phi^{(0)}\rangle = |\Phi\rangle = (1 + T^{(0)})|\Phi\rangle \Rightarrow T^{(0)} = 0, \quad (4.90)$$

$$|\Phi^{(1)}\rangle = T^{(1)}|\Phi\rangle, \quad (4.91)$$

$$|\Phi^{(2)}\rangle = \left(T^{(2)} + \frac{1}{2}(T^{(1)})^2\right)|\Phi\rangle, \quad (4.92)$$

$$\begin{aligned}
& \vdots & \vdots \\
|\Phi^{(n)}\rangle &= \left( T^{(n)} + \frac{1}{2} \sum_{i+j=n} T^{(i)} T^{(j)} + \frac{1}{3!} \sum_{i+j+k=n} T^{(i)} T^{(j)} T^{(k)} \right. \\
& \quad \left. + \dots + \frac{1}{n!} (T^{(1)})^n \right) |\Phi\rangle.
\end{aligned} \tag{4.93}$$

The  $n$ -th order correlation energy can be obtained by inserting the expansion Eq. 4.89 into Eq. 4.79, using the Hamiltonian 4.86, and separating terms which belong to the same power of  $\lambda$ :

$$E^{(n)} = \langle \Phi | \left( W_N T_2^{(n-1)} + \frac{1}{2} W_N \sum_{i=1}^{n-2} T_1^{(i)} T_1^{(n-i-1)} \right)_C | \Phi \rangle, \tag{4.94}$$

Similarly, by inserting the expansion Eq. 4.89 into Eq. 4.80 we get equations which determine  $T^{(n)}$

$$-\langle \Phi | \tilde{a}_{ab}^{ij} (F_N T_2^{(1)})_C | \Phi \rangle = \langle \Phi | \tilde{a}_{ab}^{ij} W_N | \Phi \rangle, \tag{4.95}$$

$$T_1^{(1)} = T_3^{(1)} = T_4^{(1)} = \dots = T_{k \neq 2}^{(1)} = 0, \tag{4.96}$$

$$-\langle \Phi | \tilde{a}_a^i (F_N T_1^{(2)})_C | \Phi \rangle = \langle \Phi | \tilde{a}_a^i (W_N T_2^{(1)})_C | \Phi \rangle, \tag{4.97}$$

$$-\langle \Phi | \tilde{a}_{ab}^{ij} (F_N T_2^{(2)})_C | \Phi \rangle = \langle \Phi | \tilde{a}_{ab}^{ij} (W_N T_2^{(1)})_C | \Phi \rangle, \tag{4.98}$$

$$-\langle \Phi | \tilde{a}_{abc}^{ijk} (F_N T_3^{(2)})_C | \Phi \rangle = \langle \Phi | \tilde{a}_{abc}^{ijk} (W_N T_2^{(1)})_C | \Phi \rangle, \tag{4.99}$$

$$T_4^{(2)} = T_5^{(2)} = T_6^{(2)} = \dots = T_{k > 3}^{(2)} = 0. \tag{4.100}$$

For higher orders we have:

$$\begin{aligned}
-\langle \Phi | \tilde{a}_i^a (F_N T_1^{(n)})_C | \Phi \rangle &= \langle \Phi | \tilde{a}_i^a \left[ W_N (T_1^{(n-1)} + T_2^{(n-1)} + T_3^{(n-1)}) \right]_C | \Phi \rangle \\
&+ \langle \Phi | \tilde{a}_i^a \left[ W_N \sum_{\substack{i,j=1 \\ i+j=n-1}}^{n-2} \left( \frac{1}{2} T_1^{(i)} T_1^{(j)} + T_1^{(i)} T_2^{(j)} \right) \right]_C | \Phi \rangle, \\
&+ \frac{1}{3!} \langle \Phi | \tilde{a}_i^a \left[ W_N \sum_{\substack{i,j,k=1 \\ i+j+k=n-1}}^{n-2} T_1^{(i)} T_1^{(j)} T_1^{(k)} \right]_C | \Phi \rangle,
\end{aligned} \tag{4.101}$$

$$\begin{aligned}
-\langle \Phi | \tilde{a}_{ab}^{ij} (F_N T_2^{(n)})_C | \Phi \rangle &= \\
&\langle \Phi | \tilde{a}_{ab}^{ij} \left[ W_N (T_1^{(n-1)} + T_2^{(n-1)} + T_3^{(n-1)} + T_4^{(n-1)}) \right]_C | \Phi \rangle
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \langle \Phi | \tilde{a}_{ab}^{ij} [W_N \sum_{\substack{i,j=1 \\ i+j=n-1}}^{n-2} (T_1^{(i)} T_1^{(j)} + T_2^{(i)} T_2^{(j)})]_C | \Phi \rangle \\
& + \langle \Phi | \tilde{a}_{ab}^{ij} [W_N \sum_{\substack{i,j=1 \\ i+j=n-1}}^{n-2} (T_1^{(i)} T_2^{(j)} + T_1^{(i)} T_3^{(j)})]_C | \Phi \rangle \\
& + \langle \Phi | \tilde{a}_{ab}^{ij} [W_N \sum_{\substack{i,j,k=1 \\ i+j+k=n-1}}^{n-2} (\frac{1}{3!} T_1^{(i)} T_1^{(j)} T_1^{(k)} + \frac{1}{2} T_1^{(i)} T_1^{(j)} T_2^{(k)})]_C | \Phi \rangle \\
& + \frac{1}{4!} \langle \Phi | \tilde{a}_{ab}^{ij} [W_N \sum_{\substack{i,j,k,l=1 \\ i+j+k+l=n-1}}^{n-2} T_1^{(i)} T_1^{(j)} T_1^{(k)} T_1^{(l)}]_C | \Phi \rangle, \tag{4.102}
\end{aligned}$$

$$\begin{aligned}
& - \langle \Phi | \tilde{a}_{abc}^{ijk} (F_N T_3^{(n)})_C | \Phi \rangle = \\
& \quad \langle \Phi | \tilde{a}_{abc}^{ijk} [W_N (T_2^{(n-1)} + T_3^{(n-1)} + T_4^{(n-1)} + T_5^{(n-1)})]_C | \Phi \rangle \\
& \quad + \langle \Phi | \tilde{a}_{abc}^{ijk} [W_N \sum_{\substack{i,j=1 \\ i+j=n-1}}^{n-2} (\frac{1}{2} T_2^{(i)} T_2^{(j)} + T_1^{(i)} T_2^{(j)})]_C | \Phi \rangle \\
& \quad + \langle \Phi | \tilde{a}_{abc}^{ijk} [W_N \sum_{\substack{i,j=1 \\ i+j=n-1}}^{n-2} (T_1^{(i)} T_3^{(j)} + T_1^{(i)} T_4^{(j)} + T_2^{(i)} T_3^{(j)})]_C | \Phi \rangle \\
& \quad + \frac{1}{2} \langle \Phi | \tilde{a}_{abc}^{ijk} [W_N \sum_{\substack{i,j,k=1 \\ i+j+k=n-1}}^{n-2} T_1^{(i)} (T_2^{(j)} T_2^{(k)} + T_1^{(j)} T_2^{(k)} + T_1^{(j)} T_3^{(k)})]_C | \Phi \rangle \\
& \quad + \frac{1}{3!} \langle \Phi | \tilde{a}_{abc}^{ijk} (W_N \sum_{\substack{i,j,k,l=1 \\ i+j+k+l=n-1}}^{n-2} T_1^{(i)} T_1^{(j)} T_1^{(k)} T_2^{(l)})_C | \Phi \rangle, \tag{4.103} \\
& \quad \vdots
\end{aligned}$$

or more generally:

$$- (F_N | \Phi^{(n)} \rangle)_C = (W_N | \Phi^{(n-1)} \rangle)_C. \tag{4.104}$$

It is clear that the iterative CC procedure can be easily arranged in order to obtain first and second order amplitudes in the first two iterations. Since here we have not made any approximations for  $T$ , the Eq. 4.58 is valid and the energy can be equivalently expressed in a symmetric form. We remind that in this case independently determined amplitudes of  $T$  and  $T^\dagger$  are hermitian conjugates. The energy can be equivalently written as an expectation value, which is again connected, though in principle infinite series.

$$\Delta E = \frac{\langle \Phi | e^{T^\dagger} H_N | e^T | \Phi \rangle}{\langle \Phi | e^{T^\dagger} e^T | \Phi \rangle} = \langle \Phi | (e^{T^\dagger} H_N e^T)_C | \Phi \rangle. \tag{4.105}$$

Somewhat complicated proof for this was given in [12]. Using of Eq. 4.86 and expanding the Eq. 4.105 in power series of  $\lambda$  we obtain equivalent alternatives to Eq. 4.94 in various orders, after separating terms belonging to the same power of  $\lambda$ :

$$\begin{aligned} E_D^{(2)} &= \langle \Phi | (W_N T_2^{(1)})_C | \Phi \rangle \\ &+ \underbrace{\langle \Phi | (T_2^{\dagger(1)} F_N T_2^{(1)})_C | \Phi \rangle + \langle \Phi | (T_2^{\dagger(1)} W_N)_C | \Phi \rangle}_{=0}, \end{aligned} \quad (4.106)$$

$$\begin{aligned} E_D^{(3)} &= \langle \Phi | (W_N T_2^{(2)})_C | \Phi \rangle \\ &+ \underbrace{\langle \Phi | (T_2^{\dagger(1)} F_N T_2^{(2)})_C | \Phi \rangle + \langle \Phi | (T_2^{\dagger(1)} W_N T_2^{(1)})_C | \Phi \rangle}_{=0} \\ &+ \underbrace{\langle \Phi | (T_2^{\dagger(2)} F_N T_2^{(1)})_C | \Phi \rangle + \langle \Phi | (T_2^{\dagger(2)} W_N)_C | \Phi \rangle}_{=0}, \end{aligned} \quad (4.107)$$

Here we used subscript "D" to stress that the contributions are due to double excitations. Under-braced terms at the r.h.s. in both equations give zero contribution which follows from the definition Eqs. 4.95 and 4.98, and from the fact that in our case  $F_N$  must be contracted with both  $T$  and  $T^\dagger$ , otherwise the contribution is zero. Consequently,

$$E_D^{(2)} = -\langle \Phi | (T_2^{\dagger(1)} F_N T_2^{(1)})_C | \Phi \rangle = \langle \Phi | (W_N T_2^{(1)})_C | \Phi \rangle, \quad (4.108)$$

$$\begin{aligned} E_D^{(3)} &= -\langle \Phi | (T_2^{\dagger(1)} F_N T_2^{(2)})_C | \Phi \rangle = \langle \Phi | (T_2^{\dagger(1)} W_N T_2^{(1)})_C | \Phi \rangle \\ &= \langle \Phi | (W_N T_2^{(2)})_C | \Phi \rangle. \end{aligned} \quad (4.109)$$

and of course also hermitian conjugates of these expressions are of the same value. This conclusion, if generalized, is known as the Wigner  $2n+1$  rule [43] (cf. also [27, 44]), i. e. if one has the  $n$ -th order wave function the  $2n$ -th and  $(2n+1)$ -th order energy can be obtained

$$E^{(2n)} = -\langle \Phi^{(n)} | F_N | \Phi^{(n)} \rangle_C = \langle \Phi^{(n-1)} | W_N | \Phi^{(n)} \rangle_C, \quad (4.110)$$

$$E^{(2n+1)} = -\langle \Phi^{(n)} | F_N | \Phi^{(n+1)} \rangle_C = \langle \Phi^{(n)} | W_N | \Phi^{(n)} \rangle_C. \quad (4.111)$$

Note as well that:

$$E^{(k+l)} = -\langle \Phi^{(k)} | F_N | \Phi^{(l)} \rangle_C = \langle \Phi^{(k)} | W_N | \Phi^{(l-1)} \rangle_C. \quad (4.112)$$

If we have the  $T^{(1)}$  and  $T^{(2)}$  in second iteration of the CC iterative procedure, second and third order energies are obtained from Eqs. 4.108 and 4.109 and fourth order contributions due to single (S), double (D), triple (T) and quadruple (Q) excitations are calculated as:

$$E_D^{(4)} = -\langle \Phi | (T_2^{\dagger(2)} F_N T_2^{(2)})_C | \Phi \rangle, \quad (4.113)$$

$$E_S^{(4)} = -\langle \Phi | (T_1^{\dagger(2)} F_N T_1^{(2)})_C | \Phi \rangle, \quad (4.114)$$

$$E_T^{(4)} = -\langle \Phi | (T_3^{\dagger(2)} F_N T_3^{(2)})_C | \Phi \rangle, \quad (4.115)$$

$$E_Q^{(4)} = \frac{1}{2} \langle \Phi | (T_2^{\dagger(1)} W_N T_2^{(1)} T_2^{(1)})_C | \Phi \rangle. \quad (4.116)$$

Thus, the fourth order MBPT results can be easily obtained as a byproduct of coupled cluster calculations.

## 4.6 Non Iterative Corrections to the CC Energies

From Eqs. 4.96 to 4.104 we learn in which order of the PT the individual excitation clusters appear for the first time. So recognizing that  $T_2$  is first order in  $W_N$ ,  $T_1$  and  $T_3$  are second order,  $T_4$  third order, e.t.c, we can define the following generalized PT structure functions [45]:

$$|\Phi^{[1]}\rangle = T_2 |\Phi\rangle, \quad (4.117)$$

$$|\Phi^{[2]}\rangle = \left( T_1 + T_3 + \frac{1}{2} T_2^2 \right) |\Phi\rangle, \quad (4.118)$$

$$|\Phi^{[3]}\rangle = \left( T_4 + T_1 T_2 + T_3 T_2 + \frac{1}{3!} T_2^3 \right) |\Phi\rangle, \quad (4.119)$$

$$\vdots \quad \quad \quad \vdots$$

The final wave function can be then written as:

$$|\Psi\rangle = e^T |\Phi\rangle = \sum_{k=0}^{\infty} |\Phi^{[k]}\rangle. \quad (4.120)$$

Unlike in the previous section, now each  $T_n$  operator can contain corrections which can be of infinite order in  $W_N$ . The precise definition of the function  $|\Phi^{[k]}\rangle$  will depend upon the defining equation for the amplitudes in  $T_n$ . Expansion of Eq. 4.105 in terms of generalized perturbation functions leads to similar expression as in the ordinary PT (cf. Eqs. 4.110-4.112):

$$\Delta E = \sum_{k=2}^{\infty} E^{[k]}, \quad (4.121)$$

$$\begin{aligned} E^{[2n]} &= \langle \Phi^{[n-1]} | W_N | \Phi^{[n]} \rangle_C \\ &+ \sum_{k=1-n}^{n-1} \left( \langle \Phi^{[n+k]} | F_N | \Phi^{[n-k]} \rangle + \langle \Phi^{[n+k]} | W_N | \Phi^{[n-k-1]} \rangle \right)_C, \end{aligned} \quad (4.122)$$

$$\begin{aligned} E^{[2n+1]} &= \langle \Phi^{[n]} | W_N | \Phi^{[n]} \rangle_C \\ &+ \sum_{k=1-n}^{n-1} \left( \langle \Phi^{[n+k+1]} | F_N | \Phi^{[n-k]} \rangle + \langle \Phi^{[n+k+1]} | W_N | \Phi^{[n-k-1]} \rangle \right)_C, \end{aligned} \quad (4.123)$$

If the cluster operator  $T$  is not complete, terms under sums do not completely cancel as in the ordinary PT; but, depending on defining equations for the amplitudes, those terms are canceled to a certain order. We will shortly come back to this point.

From Eqs. 4.117-4.123 it is clear that the energy of L-CCD, CCD and even CCSD is only correct to *3rd* order in the sense of PT, because the missing  $T_3$  prevents the wave function to be more accurate than *1st* order in  $W_N$ . Since treating the  $T_3$  operator in each iteration within the CC iterative procedure is much more costly than the lower versions of CC approaches, it proved to be efficient to correct the calculated (CCD or CCSD) energies in such a way that the energy were complete to fourth order in the sense of MBPT. For CCD it would mean adding of contributions from Eqs. 4.114 and 4.115, while for CCSD just from Eq. 4.115. The infinite order nature of CC is, however, much more resembled if the  $T_3$  (and for CCD also  $T_1$ ) amplitudes are calculated from Eqs. 4.99 (and 4.97), using generalized first order  $T_2$ , i. e. using  $T_2$  amplitudes which come out after the iterative procedure is finished. Those  $T_3$  (and  $T_1$  for CCD) are then used to calculate *non iterative* corrections via fourth order like energy expressions Eq. 4.115 (Eq. 4.114 for CCD) yielding what is known as the  $T(\text{CCSD})$  [31] correction to CCSD, or  $ST(\text{CCD})$  [46] to CCD. In similar manner one can correct for higher order contributions. Fifth order non iterative corrections to CCSD and variants of CCSDT have been also reported [47, 48], and even a method which is correct to sixth order [49].

Above mentioned corrections aimed at completing the energy to be correct to a certain order in the sense of MBPT. Another logical possibility is offered by an attempt to achieve the accuracy of a higher version in the CC hierarchy, calculating the lower variant and correcting the energy in a non-iterative way for some missing energy terms which are, to a given order in the sense of the generalized perturbation structure, present within the higher version. Hence, analyzing CCSDT-1 and CCSD+T(CCSD) (recently also denoted as CCSD[T] [50]) one realizes that up-to fifth order terms, they differ by a single fifth order term, namely:

$$E_{ST}^{[5]} = \langle \Phi | (T_1^\dagger W_N T_3)_C | \Phi \rangle, \quad (4.124)$$

which can be easily calculated as soon as  $T_3$  amplitudes needed for T(CCSD) are available. Correcting the CCSD[T] energy for the latter term provides the well established CCSD(T) method [51]. Similar correction is done within QCISD(T).

## 4.7 Alternative CC Hierarchies

Besides what we call the standard CC approach, there are number of other possibilities to solve the coupled cluster problem. As soon as the cluster operator were complete, the solution would be unique, irrespective of the approach. The standard

approach based on Eqs. 4.79, 4.80 has a great advantage that we solve for just one set of amplitudes (i. e. just  $T$  but not  $T^\dagger$ ), and all expansions are finite. The price to pay for this is that the resulting amplitudes do not minimize the energy functional. This complicates somewhat the calculation of energy derivatives with respect to some external perturbations (i. e. properties). E. g. if the energy is a function of many parameters (among them  $T$  and  $T^\dagger$ ) then

$$\Delta E = \Delta E(T, T^\dagger, P_1, \dots) \quad (4.125)$$

$$\begin{aligned} \left(\frac{d\Delta E}{d\alpha}\right)_{\alpha=0} &= \frac{\partial\Delta E}{\partial T} \left(\frac{dT}{d\alpha}\right)_{\alpha=0} + \frac{\partial\Delta E}{\partial T^\dagger} \left(\frac{dT^\dagger}{d\alpha}\right)_{\alpha=0} \\ &+ \frac{\partial\Delta E}{\partial P_1} \left(\frac{dP_1}{d\alpha}\right)_{\alpha=0} + \dots, \end{aligned} \quad (4.126)$$

and if both  $T$  and  $T^\dagger$  would minimize the energy functional the first two terms in Eq. 4.126 vanish, which is the case in variational approaches. Among others, alternative approaches are aimed in achieving the variational (or stationary) property at least to a higher accuracy than the standard methods do.

If the operator  $S$  in the CC ansatz is taken such that  $S^\dagger = -S$  then we have (cf. Eqs. 4.75-4.78):

$$(e^{S^\dagger} H_N e^S)|\Phi\rangle = (H_N e^S)_C |\Phi\rangle = \Delta E |\Phi\rangle. \quad (4.127)$$

The Eq. 4.127 can be again as in Eq. 4.79,4.80 projected to various subspaces and in such a way the amplitudes of  $S$  can be evaluated. The approach is both variational and it preserves the connected CC structure, i. e. the extensivity. The above property is fulfilled for  $S = \sigma = T - T^\dagger$ . Since the cluster operator is unitary, the theory suggested by Kutzelnigg [25]<sup>3</sup> is known as the Unitary Coupled Cluster Theory (UCC) [52, 53, 54, 55]. Unfortunately, the expansion Eq. 4.127 is infinite (even if projected onto the subspace of any configurations) and one is again forced to additional approximations.

Another possibility is starting from the symmetric energy functional Eq. 4.105, terminating the expansion Eq. 4.121 after  $n$ -th order terms in generalized PT sense. If such a truncated functional were directly differentiated with respect to amplitudes of  $T^\dagger$ , the size extensivity would be lost. Nevertheless, terms that would be non extensive are canceled out at least to this given order  $n$ , and are omitted. The resulting truncated energy functional is, however, again non-symmetric. Hierarchy of CC approaches based on such a functional is denoted as XCC(n) (X -Expectation value CC theory) [56]. Symmetrizing the resulting functional first, one obtains the SXCC(n) (symmetrized XCC) hierarchy which is similar to corresponding hierarchy based upon UCC (i. e. UCC(n)) [45, 57].

Another alternative possibilities are provided by projecting the SE Eq. 4.74 from

---

<sup>3</sup>Notation  $\sigma$  for this operator was introduced in this paper



the left first onto some (not only linear) approximations of  $\langle\Phi|\exp(T^\dagger)$  and again differentiating the resulting energy functional. To this category belong variational CC (VCC) alternatives suggested by Kutzelnigg [7]; but also the so called Extended CC (ECC) suggested by Arponen [58] in which the Eq. 4.74 is first projected onto  $\langle\Phi|(\exp(T^\dagger)\exp(T))_C$ .

Alternative CC approaches allow to account for contributions from higher excitations than it is within the established standard methods. Namely, if taken to be accurate to fifth order in the sense of PT, the contribution from the connected  $T_4$  clusters is considered, still avoiding its explicit evaluation [56, 59]. An analysis of the alternative CC methods has been given by Szalay et. al. [60].

Very recently, Kutzelnigg critically analyzed various possibilities of *a posteriori* corrections using energy expressions from alternative CC approaches, and suggested to evaluate such corrections at least to check the reliability of the results obtained by traditional CC approach [61].

## 4.8 Explicitly Correlated CC Approach

As mentioned in the **Introduction**, with extension of the conventional orbital basis the CI expansion converges to its exact limit extremely slowly, due to the improper description of the correlation cusp. That the explicit inclusion of the inter-electronic coordinates  $r_{ij}$  into the wave function is a powerful means to speed up the convergence has been known since Hylleraas suggested such an approach in his study of the He atom [62], and has been definitely confirmed in the calculation by Kolos and Wolniewicz of the potential curve of the hydrogen molecule [63]. Explicit inclusion of terms linear in  $r_{ij}$  (and/or odd powers of  $r_{ij}$ ) as it is done in 'Hylleraas CI' [64] leads to the appearance of 'difficult' three and four-electron integrals, which makes the computation prohibitive except for very small molecules. An alternative approach with explicit  $r_{ij}$ -dependence as well, the so called Gaussian geminal method [10], has the advantage that no 'difficult' integrals arise, but the correlation cusp is not strictly taken care of, and a sophisticated optimization of non-linear parameters is needed [65].

In 1985 Kutzelnigg suggested a theory in which explicit  $r_{ij}$ -dependent terms in the wave function expansion are present, but where the evaluation of 'difficult' integrals is avoided [66, 67, 68]. In this 'R12-method' the linear  $r_{ij}$ -dependent terms are considered in the final wave function via inclusion of pair functions like

$$|\Phi_{ij}\rangle = \sum_{\alpha>\beta} \bar{R}_{\alpha\beta}^{ij} \tilde{a}_{ij}^{\alpha\beta} |\Phi\rangle, \quad (4.128)$$

with  $\alpha, \beta$  denoting virtual orbitals which belong to the complete basis, and

$$\sum_{\alpha>\beta} \bar{R}_{\alpha\beta}^{ij} \tilde{a}_{ij}^{\alpha\beta} = \left( \sum_{\alpha>\beta} \bar{r}_{\alpha\beta}^{ij} \tilde{a}_{ij}^{\alpha\beta} - \sum_{a>b} \bar{r}_{ab}^{ij} \tilde{a}_{ij}^{ab} \right) \quad (4.129)$$

Ansatz Eq. 4.128 actually corresponds to acting of  $r_{12}$  operator on the reference function and a subsequent projection onto the orthogonal complement of the conventional (finite basis) configuration space:

$$\sum_{i>j} |\Phi_{ij}\rangle \approx \left( \sum_{\substack{\kappa>\lambda \\ i>j}} \tilde{a}_{ij}^{\kappa\lambda} |\Phi\rangle \langle \Phi | \tilde{a}_{\kappa\lambda}^{ij} - \sum_{\substack{r>s \\ i>j}} \tilde{a}_{ij}^{rs} |\Phi\rangle \langle \Phi | \tilde{a}_{rs}^{ij} \right) r_{12} |\Phi\rangle. \quad (4.130)$$

Terms with triple and higher excitations vanish because the  $r_{12}$  is a two particle operator. Terms due to single excited determinants are not included in Eq. 4.130. This is, of course, an approximation. Nevertheless, for basis sets saturated at the level of lower angular momenta those terms vanish. Summations over complete basis set which arise in the subsequent derivation of working equations are treated via completeness insertions, and products which sufficiently rapidly converge to zero with extension of the basis set are omitted. This has been called the "standard approximation" (SA) [67].

Without going more into detail, let us recall that the R12 ansatz improves the convergence towards the exact limit with extension of the basis set from  $\sim \ell^{-3}$  to  $\sim \ell^{-5}$ , where  $\ell$  corresponds to function with the highest angular momentum quantum number. Ansatz Eq. 4.128 has been implemented within MBPT(2) [68], MBPT(3) [69], CID and CEPA [70].

The action of  $r_{12}$  operator on  $|\Phi\rangle$  actually means that a pair of orbitals  $i, j$  is replaced by the pair of (the same) orbitals  $i, j$  multiplied by the inter-electronic coordinate. In a sense, pairs of orbitals multiplied by the inter-electronic coordinate make an additional virtual space of pair functions. Thus additional determinants that extend the conventional configuration space can be created substituting arbitrary pair of orbitals by a pair function from this "new" virtual space. Because ansatz Eq. 4.129 only includes certain substitutions (i. e.  $|i(1)\rangle|j(2)\rangle \rightarrow r_{12}|i(1)\rangle|j(2)\rangle$ ), the method is not invariant with respect to unitary transformations within the occupied orbitals. An invariant approach within the MBPT(2) was suggested by Klopper [71].

The ansatz Eq. 4.130 can be generalized which is conveniently accomplished within the CC theory [72]. The resulting CC-R12 can be alternatively called *explicitly correlated* CC method.

The operator "S" in Eq. 4.42 need not merely be restricted to a conventional type of excitation operator (like  $T$ ). In CC-R12  $S$  contains the conventional global excitation operator; but also a nonstandard excitation operator that creates the

**Table 3.**  
CCSD(T) and CCSD(T)-R12 energy convergence towards  
the basis limit. Energies in  $E_h$ .

Basis	size	SCF	CCSD(T)	CCSD(T)-R12
<i>LiH</i> <sup>a</sup>				
<i>sp</i> <sup>b</sup>	68	-7.987103	-0.075891	-.082598
<i>spd</i>	128	-7.987331	-.081165	-.083066
<i>spdf</i>	198	-7.987339	-.082177	-.083136
<i>H<sub>2</sub>O</i> <sup>c</sup>				
cc-pCV5Z(uc)	272	-76.067126	-0.364029	
+ <i>sp</i> <sup>d</sup>	304	-76.067176	-0.364058	
+ <i>spdfg</i>	371	-76.067176	-0.364662	
+ <i>spdfghi</i>	435	-76.067179	-0.365896	
+ <i>spdfghi</i> +D	492	-76.067401	-0.366343	
O:g/H:f <sup>e</sup>	291	-76.067345		-0.369819
O:h/H:g <sup>e</sup>	320	-76.067350		-0.369989
cc-pV6Z(uc)	351	-76.067404		-0.369401
+Steep <sup>f</sup>	374	-76.067419		-0.369891

<sup>a</sup> Values taken from ref. [73, 74].

<sup>b</sup> Subsets of Li:11s8p6d5f/H:9s8p6d5f basis set

<sup>c</sup> Values taken from ref. [79].

<sup>d</sup> Correlation consistent basis cc-pCV5Z(uc) [79] augmented by subsets of *1p1d3f2g1h* and further diffuse functions (D)

<sup>e</sup> O:h≡15s9p7d5f3g1h, H:g≡9s7p5d3f1g, O:g,H:f are subsets

<sup>f</sup> Steep functions added to cc-pV6Z(uc) [79]

above mentioned pair functions. In order the method were invariant, we have to permit any pair of functions to be replaced by any other  $r_{12}$ -multiplied pair of spin-orbitals, i. e.  $|k(1)\rangle|l(2)\rangle \rightarrow r_{12}|i(1)\rangle|j(2)\rangle$ . This is achieved if we choose  $S = T + \mathcal{R}$  [72] with

$$\mathcal{R} = \sum_{\substack{i>j \\ k>l}} c_{ij}^{kl} \mathcal{R}_{kl}^{ij} \quad (4.131)$$

$$\mathcal{R}_{kl}^{ij} = \sum_{\alpha>\beta} \bar{R}_{\alpha\beta}^{ij} \tilde{a}_{kl}^{\alpha\beta} \quad (4.132)$$

Note that the operator  $\mathcal{R}$  essentially contains a replacement operator of the same form as  $T_2$ . However, now the summation over the given basis  $a, b$  is replaced by the orthogonal complement to the latter within the complete basis  $\alpha, \beta$ . This implies

that  $\mathcal{R}$  commutes with  $T$  which is the necessary condition in order to be plugged in the exponential together with  $T$ . Moreover, the choice of  $\mathcal{R}$  guarantees the orthogonality of the resulting functions to  $|\Phi\rangle$  and consequently the intermediate normalization Eq. 4.45. Note that the operator  $\mathcal{R}$  is actually a doubly connected product of two operators, namely

$$\mathcal{R} = (\mathcal{C}\mathcal{R}_o)_{CC}, \quad (4.133)$$

$$\mathcal{C} = \sum_{\substack{i>j \\ k>l}} c_{kl}^{ij} \tilde{a}_{ij}^{kl}, \quad (4.134)$$

$$\mathcal{R}_o = \sum_{\substack{\alpha>\beta \\ k>l}} \bar{R}_{\alpha\beta}^{kl} \tilde{a}_{kl}^{\alpha\beta}. \quad (4.135)$$

We have exceptionally used subscript "CC" to stress that the two operators are always doubly connected. Hence,

$$\mathcal{R} = (\mathcal{C}\mathcal{R}_o)_{CC} = \sum_{i>j} \sum_{k>l} \sum_{\alpha>\beta} c_{kl}^{ij} \bar{R}_{\alpha\beta}^{kl} \tilde{a}_{ij}^{\alpha\beta}. \quad (4.136)$$

Since now we have additional "excitation" space, the working CC equations change accordingly:

$$\langle \Phi | (H_N e^{T+\mathcal{R}})_C | \Phi \rangle = \Delta E, \quad (4.137)$$

$$\langle \Phi | \tilde{a}_{ab\dots}^{ij\dots} (H_N e^{T+\mathcal{R}})_C | \Phi \rangle = 0, \quad (4.138)$$

$$\langle \Phi | (\mathcal{R}_{ij}^{kl})^\dagger (H_N e^{T+\mathcal{R}})_C | \Phi \rangle = 0. \quad (4.139)$$

More details including the use of diagrammatic technique in this particular method are given elsewhere [72]. Recent calculations on CCSD-R12 [50] and CCSDT1-R12 [73, 75, 76, 77, 78, 79, 80] level prove the ability of the CC-R12 to provide extremely accurate energies and other properties. Two examples given in Tables 3 evidently illustrate the much faster convergence of the CC-R12 values with extension the one electron basis set towards the limit.

## 4.9 Acknowledgments

The author acknowledges a continuous support of the research in this field by the Slovak grant agency VEGA (project No. 4227). Author thanks V. Malkin and I. Mayer for careful reading of the manuscript and for valuable comments.

# Bibliography

- [1] S. F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970)
- [2] R.N. Hill, *J. Chem. Phys.* **126**, 1015 (1962); W. Kutzelnigg and J.D. Morgan jr., *J. Chem. Phys.* **96**, 4484 (1992)
- [3] T. Kato, *Commun. Pure Appl. Math.* **10**, 15 (1957)
- [4] R. J. Bartlett, in: D. A. Yarkony (ed.) *Modern Electronic Structure Theory*, Part I, p. 1047, World Scientific, Singapore (1995)
- [5] J. Paldus, in: G. I. Malli (ed.), *Relativistic and Electron Correlation Effects in Molecules and Solids*, p. 207, Plenum Press, New York, (1994)
- [6] P. Taylor, *Lect. Notes Chem.* **64**, 1 (1994)
- [7] W. Kutzelnigg, *Theoret. Chim. Acta* **80**, 349 (1991)
- [8] R. J. Barlett, *J. Phys. Chem.* **93** 1697 (1988)
- [9] M. Urban, I. Černušák, V. Kellö, J. Noga, in: S. Wilson (ed.), *Methods in Computational Chemistry*, Vol. 1., p. 117 Plenum Press, New York (1987), and references therein
- [10] R.J. Bartlett (ed.), *Recent Advances in Computational Chemistry Vol. 3 Recent Advances in Coupled-Cluster Methods*, World Scientific, Singapore (1997)
- [11] P. Surján, "Second Quantized Approach to Quantum Chemistry. An Elementary Introduction" Springer, Berlin (1989)
- [12] J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966); *Adv. Chem. Phys.* **14**, 35 (1969).
- [13] G C. Wick, *Phys. Rev.* **80**, 268 (1951)
- [14] J. E. Meyer and M. G. Meyer, *Statistical Mechanics*, Willey, New York, 1940

- [15] M. Gell-Mann and F. Low, Phys. Rev **106**, 364 (1951)
- [16] J. Hubbard, Proc. Roy. Soc. (London) A, **240**, 539 (1957); *ibid* **243**, 336 (1957); *ibid* **244**, 199 (1957)
- [17] A. Hugenholtz, Physica, **23**, 481 (1957); *ibid* **23**, 533 (1957)
- [18] F. Coester, Nucl. Phys. **7**, 421 (1958); F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).
- [19] see e. g. W. H. Louisell, "Radiation and Noise in Quantum Electronics", Sec. 3.2, McGraw-Hill, New York (1964)
- [20] J. Paldus, J. Čížek, I. Shavitt, Phys. Rev. A **5**, 50 (1972)
- [21] R. J. Bartlett, G. D. Purvis, Int. J. Quant. Chem. **14**, 561 (1978)
- [22] J. A. Pople, R. Krishnan, H. B. Schlegel, J. Binkley, Int. J. Quant. Chem. **14**, 545 (1978)
- [23] P. R. Taylor, G. B. Bacskay, A. C. Hurley, N. S. Hush, J. Chem. Phys. **69**, 1971 (1978)
- [24] G.D. Purvis and R.J. Bartlett, J. Chem. Phys **76**, 1910 (1982)
- [25] W. Kutzelnigg, in: H. F. Schaefer III (ed.), *Modern Theoretical Chemistry Vol. 3: Methods of Electronic Structure Theory*, p. 129, Plenum Press, New York, (1977)
- [26] R. Ahlrichs, Comput. Phys. Commun. **17**, 31 (1979)
- [27] R. J. Bartlett and D. M. Silver, in: J. L. Calais, O. Goscinski, J. Linderberg, and Y. Öhrn (eds.), *Quantum Science*, Plenum Press, New York (1976)
- [28] J. A. Pople, M. Head-Gordon, and K. Raghavachari, J. Chem. Phys. **87**, 5968 (1987)
- [29] J. Paldus, J. Čížek, and B. Jeziorski, J. Chem. Phys. **90**, 4356 (1989)
- [30] Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, J. Chem. Phys. **81**, 5906 (1984).
- [31] M. Urban, J. Noga, S. J. Cole, and R. J. Bartlett, J. Chem. Phys. **83**, 404 (1985).
- [32] J. Noga, R. J. Bartlett, and M. Urban, Chem. Phys. Lett. **134**, 126 (1987).
- [33] J. Noga and R. Bartlett, J. Chem. Phys. **86**, 7071 (1987); **89**, 3401(E) (1988).

- [34] G. E. Scuseria, H. F. Schaefer III, Chem. Phys. Lett. **152**, 382 (1988)
- [35] V. Kvasnička, V. Laurinc, S. Biskupič, Phys. Rep. **90**, 159 (1982); V. Kvasnička, V. Laurinc, S. Biskupič and M. Haring, Adv. Chem. Phys. **52**, 181 (1983)
- [36] S. A. Kucharski and R. J. Bartlett, Chem. Phys. Lett. **158**, 550 (1989).
- [37] N. Oliphant and L. Adamowicz, J. Chem. Phys. **95**, 6645 (1991).
- [38] J. C. Slater, Phys. Rev. **38**, 1109 (1931)
- [39] J. Paldus and J. Čížek, Advan. Quantum. Chem. **9**, 105 (1975)
- [40] F. E. Harris, H. J. Monkhorst, D. L. Freeman, *Algebraic and Diagrammatic Methods in Many-Fermion Theory*, Oxford University Press, New York, Oxford (1992)
- [41] C. Møller, M. S. Plesset, Phys. Rev. **46**, 618 (1934)
- [42] B. O. Roos and P. E. M. Siegbahn, in: H. F. Schaefer III (ed.), *Modern Theoretical Chemistry Vol. 3: Methods of Electronic Structure Theory*, p. 277, Plenum Press, New York, (1977)
- [43] E. P. Wigner, Math. Naturw. Anz. (Budapest) **53** 477 (1935)
- [44] V. Kvasnička, V. Laurinc, S. Biskupič, Czech. J. Phys. **B31**, 41 (1981)
- [45] R. J. Bartlett, S. A. Kucharski, J. Noga, J. D. Watts, G. W. Trucks: Lecture Notes in Chemistry **52**, 125, ed. U. Kaldor, Springer, Berlin (1989)
- [46] K. Raghavachari, J. Chem. Phys. **82**, 5761 (1985)
- [47] R. J. Bartlett, J. D. Watts, S. A. Kucharski, and J. Noga Chem. Phys. Lett. **165**, 513 (1990).
- [48] K. Raghavachari, J. A. Pople, E. S. Replogle, and M. Head-Gordon, J. Phys. Chem., **94**, 5579 (1990).
- [49] S. A. Kucharski and R. J. Bartlett, Chem. Phys. Lett. **206**, 574 (1993).
- [50] J. Noga, W. Kutzelnigg, and W. Klopper, Chem. Phys. Letters **199**, 497 (1992)
- [51] K. Raghavachari, G.W.Trucks, J.A. Pople, and M. Head-Gordon, Chem. Phys. Lett., **157**, 479 (1989).

- [52] S. Koch and W. Kutzelnigg, *Theor. Chim. Acta* **59**, 397 (1981); W. Kutzelnigg, *J. Chem. Phys.* **77**, 3081 (1982); *ibid* **80**, 822 (1984); W. Kutzelnigg and S. Koch, *J. Chem. Phys.* **79**, 4315 (1983)
- [53] S. Pal, M. Durga Prasad, and D. Mukherjee, *Theor. Chim. Acta* **62**, 523 (1983); *ibid* **66**, 311 (1984)
- [54] K. Tanaka and H. Terashima, *Chem. Phys. Lett.* **106**, 558 (1984)
- [55] M. Hoffman and J. Simons, *J. Chem. Phys.* **88**, 993 (1988); *Chem. Phys. Lett.* **142**, 451 (1987)
- [56] R. J. Bartlett and J. Noga, *Chem. Phys. Lett.* **150**, 29 (1988)
- [57] R. J. Bartlett, S. A. Kucharski, and J. Noga, *Chem. Phys. Lett.* **154**, 133 (1989); J. D. Watts, G. W. Trucks and R. J. Bartlett, *Chem. Phys. Lett.* **157**, 359 (1989)
- [58] J. Arponen, *Annals Phys.* **151**, 311 (1983); J. Arponen, R. F. Bishop, and E. Pajanne, *Phys. Rev.* **A36**, 2519 (1987); *ibid* **A36**, 2539 (1987); R. F. Bishop, J. Arponen, and E. Pajanne, *Lecture Notes in Chemistry* **50**, ed. D. Mukherjee, p. 79, Springer, Berlin (1989); R. F. Bishop and J. Arponen, *Int J. Quant. Chem.* **S24**, 197 (1990)
- [59] J. Noga, S. A. Kucharski, R. J. Bartlett, *J. Chem. Phys.* **90**, 3399 (1989)
- [60] P. G. Szalay and R. J. Bartlett, *Int. J. Quant. Chem.* **S26**, 85 (1992); P. G. Szalay, M. Nooijen and R. J. Bartlett, *J. Chem. Phys.* **103** 281 (1995)
- [61] W. Kutzelnigg, *Molec. Phys.* **94**, 65 (1998)
- [62] E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929)
- [63] W. Kolos, L. Wolniewicz, *J. Chem. Phys.* **41**, 3663 (1964); *J. Chem. Phys.* **43**, 2429 (1965); *Phys. Rev. Lett.* **20**, 243 (1968);
- [64] J. S. Sims, S. Hagström, *Phys. Rev. A* **4**, 908 (1971); *J. Chem. Phys.* **55**, 4699 (1971); D. Frye, A. Preiskorn, G.C. Lie and E. Clementi, in *Modern Techniques in Computational Chemistry*. MOTTECC-90, E. Clementi ed., Escom, Leiden (1990)
- [65] K. Szalewicz, B. Jeziorski, H.J. Monkhorst and J.G. Zabolitzky, *J. Chem. Phys.* **78**, 1420 (1983); **79**, 5543 (1983); S. A. Alexander, H.J. Monkhorst, K. Szalewicz, *J. Chem. Phys.* **85**, 5821 (1986); **87**, 3976 (1987)
- [66] W. Kutzelnigg, *Theoret. Chim. Acta* **68**, 445 (1985);



- [67] W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991)
- [68] W. Klopper and W. Kutzelnigg, *Chem. Phys. Letters* **134**, 17 (1986)
- [69] W. Klopper and W. Kutzelnigg, *J. Chem. Phys.* **94**, 2020 (1991)
- [70] W. Klopper, R. Röhse and W. Kutzelnigg, *Chem. Phys. Letters* **178**, 455 (1991)
- [71] W. Klopper, *Chem. Phys. Letters* **186**, 583 (1991)
- [72] J. Noga and W. Kutzelnigg, *J. Chem. Phys.* **101**, 7738 (1994)  
J. Noga, W. Klopper, W. Kutzelnigg in: *Recent Advances in Computational Chemistry Vol. 3 Recent Advances in Coupled-Cluster Methods*, R.J. Bartlett ed., World Scientific, Singapore, (1997) pp. 1-48
- [73] J. Noga, D. Tunega, W. Klopper, and W. Kutzelnigg, *J. Chem. Phys.* **103**, 309 (1995)
- [74] D. Tunega, J. Noga, *Theoret. Chem. Acc.* , in press (1998)
- [75] W. Klopper, J. Noga, *J. Chem. Phys.* **103**, 6127 (1995)
- [76] D. Tunega, J. Noga, W. Klopper, *Chem. Phys. Lett.* **269**, 435 (1997)
- [77] H. Müller, W. Kutzelnigg, J. Noga, W. Klopper, *J. Chem. Phys.* **106**, 1863 (1997)
- [78] H. Müller, W. Kutzelnigg, J. Noga, *Molec. Phys.* **92**, 535 (1997)
- [79] T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997)
- [80] W. Klopper, M. Quack, M.A. Suhm, *Molec. Phys* **94**, 105 (1998)