On the coupled-cluster equations. Stability analysis and nonstandard correction schemes

Péter R. Surján and Ágnes Szabados

Abstract (i) The coupled-cluster equations being nonlinear, they have to be solved iteratively. An insight into the convergence properties of this iteration can be obtained by analysing the stability of the converged solutions as fixed points. (ii) The usual form of coupled-cluster equations represents an example to the method of moments, with the number of unknown amplitudes being equal to the number of equations. The method of moments generates nonsymmetric equations loosing the variational character of the coupled-cluster method, but enabling efficient evaluation of the matrix elements. Taking higher moments into account, one may obtain more equations than parameters, thus the latter must be determined by minimizing the sum-of-squares of all moments. This leads to additional effort but improved coupled-cluster wave functions and/or energies. (iii) Another way of improving the coupled-cluster method is perturbation theory, which needs special formulations due to the nonsymmetric nature of the formalism. An efficient way to do this is offered by multi-configuration perturbation theory.

28.1 Introduction

Most fundamental equations of physics are linear. The basic equation of quantum mechanics, Schrördinger's equation, is also linear. Nonlinearity in nature usually appears as a consequence of complexity, often caused by dissipative forces, cf. the field equations of condensed matter (fluid mechanics).

Péter R. Surján
Eötvös University, Institute of Chemistry, Laboratory of Theoretical Chemistry, e-mail: surjan@chem.elte.hu

Ágnes Szabados
Eötvös University, Institute of Chemistry, Laboratory of Theoretical Chemistry, e-mail: szabados@chem.elte.hu

The mathematics of linear equations is relatively simple and well developed. Standard variational tools and eigenproblems offer the solution in most cases. In turn, nonlinear equations usually lack standard methods to solve, and dealing with them one often encounters numerical and conceptual difficulties. Here the variational methods are often replaced by the method of moments, and one may attain the solution only via complicated iteration schemes. One easily gets the impression that whenever the physics of the problem permits, one should rely on linear formulations.

A fundamental approximative method of the many-electron problem, however, deals with genuinely nonlinear equations. Coupled-cluster (CC) theory involves exponential wave operators to ensure the extensivity of approximate energies, and generating nonlinearity at the same time. In this review, we emphasize the following issues.

1. The nonlinear nature of the CC equations deserves an analysis of the stability of the solutions. This will offer an insight into the convergence feature of iterative solutions.
2. The CC equations emerge as a trivial application of the method of moments. Improved results can be expected when higher moments, neglected in the standard formulation, are considered.
3. The nonsymmetric nature of the method of moments makes it necessary to develop special perturbation schemes to obtain systematic corrections to a given level of CC theory.

28.2 Stability analysis of iteration schemes

Owing to the size and the nonlinear nature of the CC equations, one uses, almost exclusively, iterative procedures to locate their solutions representing the fixed points of the iteration scheme. To ensure and accelerate convergence, one often applies control parameters or extrapolation techniques such as DIIS[1]. Nonlinearity of the equations represents a complication not only bacause of iterative solutions, but also by generating spurious solutions. Analysis in this line has been carried out Jankovski [2]. Here, we begin this review with a standard tool for analysing convergence features of nonlinear iterative schemes.

The theory of stability matrices and Ljapunov exponents is well established[3]. A general iteration procedure for an m-component vector $\underline{x}$ is given as

$$x_i^{(n+1)} = f_i\left(\underline{x}^{(n)}\right), \quad i = 1, 2, \ldots m. \tag{28.1}$$

Let vector $\underline{a}$ be a fixed point of this iteration:

$$a_i = f_i\left(\underline{a}\right), \quad i = 1, 2, \ldots m. \tag{28.2}$$

Small deviations around this fixed point are written as

$$\underline{x}^{(n)} = \underline{a} + \underline{\xi}^{(n)}. \tag{28.3}$$

Substitution into (28.1) and expanding the m-variable function f into Taylor series up to the first order gives

$$\xi_i^{(n+1)} = \sum_{j=1}^{m} J_{ij} \xi_j^{(n)} + \mathscr{O}(2) \tag{28.4}$$

with the definition of the Jacobian at point $\underline{a}$

$$J_{ij} = \left. \frac{\partial f_i}{\partial x_j} \right|_{\underline{a}} . \tag{28.5}$$

The Jacobi matrix **J** is, in this context, the *stability matrix*. Solution of (28.4) can be looked for in the form

$$\underline{\xi}^{(n)} = e^{\lambda n} \underline{\xi}^{(0)} \tag{28.6}$$

which, after substitution into (28.4) and writing $\lambda = \log \mu$ provides the eigenvalue problem

$$\mathbf{J} \underline{\xi}^{(0)} = \mu \underline{\xi}^{(0)} \tag{28.7}$$

Modes $\underline{\xi}^{(0)}$ are eigenvectors of the stability matrix while the logarithm of the eigenvalues give parameters λ which are related to the Ljapunov exponents of the problem[1].

Analysis of convergence properties of the iteration process (28.1) can be based on the value of the Ljapunov exponents λ, or on their exponentials μ. If all μ-s are positive so all λ-s are real, the procedure converges only if all Ljapunov exponents are negative, that is, if all eigenvalues of the stability matrix satisfy

$$0 < \mu < 1.$$

When one or more exponents are positive, i.e., $\mu > 1$, the iteration will diverge along the corresponding trajectory.

It may happen that one or more eigenvalues of the Ljapunov matrix are negative, generating complex Ljapunov exponents: $\lambda = \lambda_1 + i\lambda_2$. To have a real μ, we require that

$$\operatorname{Im} e^{\lambda} = e^{\lambda_1} \, \sin \lambda_2 \, = 0,$$

which is satisfied by $\lambda_2 = k\pi$ with any integer k. However, $\mu = \cos \lambda_2 \, e^{\lambda_1}$ is negative only for odd k values, thus we may choose $k = 1$. Therefore, the Ljapunov exponent for real, negative μ can be written as

$$\lambda = \log |\mu| + i\pi$$

leading to the convergence condition

---

[1] The more common 'dynamical' definition of Ljapunov exponents works with the limit $n \to \infty$; we adopt here a 'static' definition based on the converged solution $\underline{a}$

$$\mathrm{Re}\ \lambda = \log|\mu| < 0,$$

which requires the moduli of all eigenvalues to be smaller than unity:

$$|\mu| < 1. \tag{28.8}$$

The iterations in this case are not monotonic, but exhibit oscillatory convergence:

$$\underset{\sim}{\xi}^{(n)} \ = \ e^{\lambda n}\underset{\sim}{\xi}^{(0)} \ = \ e^{i\pi n}e^{n\log|\mu|}. \tag{28.9}$$

As the Jacobian is not symmetric, its eigenvalues μ may also become complex. Let us have $\mu = a + ib$. Then, from $e^{\lambda} = a + ib$ we get

$$\left.\begin{aligned} e^{\lambda_1}\cos\lambda_2 &= a \\ e^{\lambda_1}\sin\lambda_2 &= b \end{aligned}\right\}$$

resulting

$$e^{\lambda_1} = \sqrt{a^2 + b^2} \ = \ |\mu|,$$

which leads again to (28.8).

All above cases can be summarized in the condition of convergence:

$$|\mu| < 1$$

Violating this condition, the procedure usually diverges, but on the borderline of convergence and divergence, nonlinear systems may also exhibit chaotic iterations[4]. This is manifested in irregular and stochastic iteration patterns.

28.3 The CC equations

Previously[5], we have applied the above theory to study iteration characteristics of the Bloch equation[6, 7] and the idempotency-conserving density matrix iteration[8, 9]. Here we apply it to the CCSD equations.

Using the CC Ansatz $\Psi = \mathrm{e}^{T}\Phi$, we write the Schrödinger equation as

$$\mathrm{e}^{-T}H\,\mathrm{e}^{T}\Phi \ = \ E\Phi \tag{28.10}$$

where Φ is the reference state, typically the Hartree-Fock solution. The cluster operator for k-fold excitations T_k is formally written as

$$T_k = \frac{1}{k!}\sum_{\mu} {}^{k}t_{\mu}\hat{E}_{\mu}, \tag{28.11}$$

where ${}^{k}t_{\mu}$ are the cluster amplitudes to be determined, and $\hat{E}_{\mu}$-s are excitation operators which, for closed shell systems, can be written as

$$\hat{E}_\mu = E_u^\beta = \sum_\sigma a^\dagger_{\beta,\sigma} a_{u,\sigma} \tag{28.12}$$

for single and

$$\hat{E}_\mu = E_{uv}^{\beta\gamma} = E_u^\beta E_v^\gamma \tag{28.13}$$

for double excitations. From (28.10), the CC equations for the amplitudes emerge as

$$F_\nu \;=\; \langle \Phi | \hat{E}_\nu^\dagger \, \mathrm{e}^{-T} H \, \mathrm{e}^{T} | \Phi \rangle \;=\; 0. \tag{28.14}$$

Substituting the Hamiltonian, the above forms of the excitation operators and evaluating the matrix elements, this equation can be reduced to its orbital form ${}^1F_u^\beta = 0$ and ${}^2F_{uv}^{\beta\gamma} = 0$. The resulting formulae for CCSD were originally tabulated in Ref.[10], and recollected in the Appendix of Ref.[11]. Both ${}^1t$ and ${}^2t$ equations detailed there are of the shape

$$F(\underline{t}) = 0,$$

and there are several ways to recast them into an iterative form $t = f(t)$, e.g. using the scheme

$$t = t + kF(t)$$

where k is a t-independent arbitrary parameter or expression. We consider here the form of the CCSD equations which uses Møller-Plesset-type denominators:

$$t_u^\beta \;=\; t_u^\beta \;-\; \frac{{}^1F_u^\beta}{\varepsilon_\beta - \varepsilon_u - \eta} \tag{28.15}$$

$$t_{uv}^{\beta\gamma} \;= t_{uv}^{\beta\gamma} \;-\; \frac{{}^2F_{uv}^{\beta\gamma}}{\varepsilon_\beta + \varepsilon_\gamma - \varepsilon_u - \varepsilon_v - 2\eta} \tag{28.16}$$

for the ${}^1t$ and ${}^2t$ amplitudes, respectively. Parameter η is introduced merely to control the iteration by damping or accelerating the sequence[2]. In these equations, ε denote orbital energies.

28.4 The stability matrix of the CCSD equation

The stability matrix of the CCSD problem can be obtained by taking the derivatives of the iterative equations with respect to the independent parameters. In our case, the latter are represented by the cluster amplitudes ${}^1t_\mu$ and ${}^2t_\mu$. Since the CCSD equations for these amplitudes are coupled, the stability matrix will be composed of four blocks which we denote by ${}^{11}\mathbf{J}$, ${}^{12}\mathbf{J}$, ${}^{21}\mathbf{J}$, ${}^{22}\mathbf{J}$, respectively, where ${}^{mn}\mathbf{J}$ stands for the derivatives of the m-equations with respect to the ${}^nt$ amplitudes. The four blocks of the stability matrix are obtained as

---

[2] Negative (positive) values for η will damp (accelerate) the iteration by increasing (decreasing) the denominators, respectively.

$$\begin{aligned}
\mathbf{J}^{11}_{u\beta,l\lambda} &= \delta_{ul}\delta_{\beta\lambda} - \frac{\dfrac{\partial\ {}^{1}F_{u\beta}}{\partial t_{l}^{\lambda}}}{\varepsilon_\beta - \varepsilon_u - \eta} \\
\mathbf{J}^{12}_{u\beta,lm\lambda\mu} &= -\frac{\dfrac{\partial\ {}^{1}F_{u\beta}}{\partial t_{lm}^{\lambda\mu}}}{\varepsilon_\beta - \varepsilon_u - \eta} \\
\mathbf{J}^{21}_{uv\beta\gamma,l\lambda} &= -\frac{\dfrac{\partial\ {}^{2}F_{uv\beta\gamma}}{\partial t_{l}^{\lambda}}}{\varepsilon_\beta + \varepsilon_\gamma - \varepsilon_u - \varepsilon_v - 2\eta} \\
\mathbf{J}^{22}_{uv\beta\gamma,lm\lambda\mu} &= \delta_{ul}\delta_{vm}\delta_{\beta\lambda}\delta_{\gamma\mu} - \frac{\dfrac{\partial\ {}^{2}F_{uv\beta\gamma}}{\partial t_{lm}^{\lambda\mu}}}{\varepsilon_\beta + \varepsilon_\gamma - \varepsilon_u - \varepsilon_v - 2\eta}.
\end{aligned} \tag{28.17}$$

The explicit results for these derivatives are listed in Ref.[11].

The above equations show that the stability matrix $\mathbf{J}$ is in close connection to the Jacobi matrix $\mathbf{J}^0$ of the CCSD equations. The structure of Eqs. (28.17) is

$$J_{ab} = \delta_{ab} - \frac{J^0_{ab}}{\Delta_{ab} - \eta} \tag{28.18}$$

where Δ_{ab} stands for the energy denominator in the iterative equations corresponding to the $a \rightarrow b$ excitation. Using Eq.(28.14), we obtain

$$J^o_{ab} = \frac{\partial F_a}{\partial t_b} \; = \; \langle \Phi | \hat{E}_a^\dagger \, \mathrm{e}^{-T} [H, \hat{E}_b] \, \mathrm{e}^{T} | \Phi \rangle$$

The eigenvalue equation of this matrix is known as the EOM-CC equation for excitation energies[12, 13, 14, 15, 16].

It is important to emphasize the following two features of the above analysis:

- It is valid only if no extrapolation techniques like DIIS are used to govern the iteration. While the DIIS method is known to be highly efficient in accelerating convergence, it does not change the *nature* of the fixed point. If, by a wrongly chosen value for η, the fixed point of the iteration becomes repellent, neither DIIS nor any other extrapolation technique can be expected to ensure convergence.
- The Ljapunov exponents as defined above characterize the nature of the fixed point, rather than the process of the iteration. We always assume that the initial amplitudes lie in the sufficient proximity of the fixed point.

Numerical illustrations of the above theory published in Ref.[11] seem to suggest that a sufficiently large value of the damping parameter η always ensures convergence. We cannot provide a formal proof for this statement. Even if one finds a sufficiently large value for η so that all Ljapunov exponents are negative, this means only that the fixed point of the iteration is attractive. Starting from a set of amplitudes far from the converged ones, the iteration may diverge or may converge to another fixed point. However, we have not yet treated any cases in which convergence was impossible to achieve. (Note that a convergent iteration does not mean that one has reached the desired state.)

In concluding, the nature of the fixed points of the CCSD equations can be analysed by finding the eigenvalue of the stability matrix which has the largest absolute value. The logarithm of the modulus of the eigenvalue (admitting complex solutions) can be regarded as the Ljapunov exponent of the problem which should be negative to ensure convergence. The nature of the fixed points can be effectively controlled by denominator shifts which may either damp or accelerate the iteration. Connection between the stability matrix $\mathbf{J}$ and the Jacobian $\mathbf{J}^0$ reveals an interesting relation between excitation energies (eigenvalues of $\mathbf{J}^0$) and the Ljapunov exponents (eigenvalues of $\mathbf{J}$).

28.5 Improving CC results: method of moments

In this section, following an idea by Jankowski et al.[17], we investigate a possibility to improve the cluster amplitudes. Emphasizing the close connection between CC theory and the theory of moments, CC amplitudes are determined by minimizing the sum-of-squares of the moments of the Hamiltonian constructed with the CCSD trial function. In addition to standard moments with SD excitations, we include all moments with excitations higher than doubles in an unconstrained minimization. This procedure is computationally demanding thus it does not lead to any practical method. To circumvent the computational difficulties of amplitude minimization, Jankowski et al.[17] substituted the exponential of the cluster operator by the linearized CCD (LCCD) Ansatz, reducing thereby the moment optimization problem to a solution of a linear system of equations. In this work, keeping the nonlinear e^T Ansatz, we rely on a numerical minimization procedure.

Since the publication of the original work [17], many papers have appeared exploring higher moments in CC theory[18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35]. A common feature of these works is that they, with a given CC Ansatz, use the higher moments either to modify the amplitude equations or derive (noniterative) corrections to the CC energy. For a theoretical comparison of various CC Ansätze, see [36].

Owing to the recent success of utilizing higher moments in CC theory, it appears to be interesting to revisit the problem of moment optimization with the simplest, traditional CCSD Ansatz.

28.5.1 The method of moments

The method of moments has been introduced in quantum chemistry a long time ago[37, 38, 39]. It can be summarized as follows. If Ψ is an exact eigenvector of H, and E is the associated eigenvalue, any quantities of the form

$$m_\nu = \langle f_\nu | H - E | \Psi \rangle \quad \nu \; = \; 1,2,\ldots p \tag{28.19}$$

are identically zero for arbitrary well-behaved testing functions f_ν. Quantities m_ν are the moments of the Hamiltonian. If the exact wave function is substituted by an approximate wave function Φ, the above moments are not necessarily zero. Considering Φ to depend on the parameter set $t_\mu, \;\; \mu = 1,2,\ldots n$, the best trial function Φ can be determined by minimizing the functional for a given set of testing functions f_ν,

$$M = \sum_{\nu=1}^{p} m_\nu^2 \tag{28.20}$$

with respect to parameters t_μ in Φ. The necessary condition for this minimum is

$$\frac{\partial M}{\partial t_\mu} \; = \; 2 \sum_{\nu=1}^{p} m_\nu \, \frac{\partial m_\nu}{\partial t_\mu} \; = \; 0 \qquad \mu = 1,2,\ldots n \tag{28.21}$$

which constitutes an $n \times n$ system of equations for the n unknown parameters t_μ. One may show that in the special case where (i) the testing functions f_ν are obtained from a common bra generator function depending on the same set of parameters as the ket function Φ, and (ii) the bra generator and the ket trial function are the same, then the method of moments reduces to the variational method.

The accuracy of the method of moments depends on (i) the number of parameters in the trial ket function, n, and (ii) the number of moments considered, p. In general, p can be much larger than n. The method has a trivial variant when $p = n$ which corresponds to the standard CCSD equations. In this case the functional M can be set exactly zero, since the equations

$$m_\nu = 0$$

can be solved for all ν. This is, of course, not the exact solution if the set of testing functions f_ν-s is restricted.

28.5.2 CC theory and the method of moments

The standard CC method accomplishes the trivial version of the method of moments. To see this, we introduce the similarity-transformed Hamiltonian

$$\overline{H} = e^{-T} H e^{T},$$

and choose the bra test functions as $\langle f_\nu| = \langle 0|\, T_\nu^\dagger$. Therefore, all moments

$$m_\nu \;=\; \langle 0|\, T_\nu^\dagger\, (\overline{H} - E)|0\rangle \;=\; \langle 0|\, T_\nu^\dagger\, \overline{H}|0\rangle \;=\; 0 \qquad (28.22)$$

are required to vanish for $\nu = 1,2,\ldots n$. Here we used the orthogonality of the test functions to the reference state to get rid of the energy.

Having n amplitudes e.g. in a CCSD wave function, we eliminate only the moments which are projected by single and double substitutions $\langle 0|T_\nu^\dagger$. However, both the exponential wave operator e^T and the Hamiltonian H, when acting on the reference state $|0\rangle$, generate moments which correspond to triple, quadruple, etc. excitations. These moments are completely neglected in standard amplitude equations, making CCSD theory nonexact, as many of these moments are nonzero in general.

Stimulated by the theory of moments, one may ask what happens when trying to achieve a balance between the moments emerging from SD and higher excitations. That is, what kind of CCSD amplitudes result by minimizing the functional

$$M = \sum_{\nu=1}^{p} m_\nu^2 \;=\; \sum_{\nu=1}^{p} \big(\langle 0|\, T_\nu^\dagger\, \overline{H}\, |0\rangle\big)^2 \qquad (28.23)$$

with respect to the CCSD amplitudes t_μ, when

$$\nu = 1,2,\ldots p \qquad (p > n)$$

runs over S, D, T, Q, … excitations. In the following section a numerical procedure is outlined that is used to perform an unconstrained minimization of the sum of squared moments.

28.5.3 Amplitude optimization

Solution of the problem represented by the minimization of functional M in Eq.(28.23) can be performed in a general way by numerical optimization techniques. To this end, derivation of the gradients (derivatives of M with respect to the amplitudes) is highly desirable. Components of the gradient vector g

$$g_\lambda \;=\; \frac{\partial M}{\partial t_\lambda}$$

were defined in Eq.(28.21). Using the shorthand

$$\langle \nu| \;=\; \langle 0|T_\nu^\dagger$$

and the decomposition

$$T = \sum_\mu t_\mu T_\mu,$$

the derivatives of the individual moments read as

$$\begin{aligned}\frac{\partial m_\nu}{\partial t_\lambda} &= \langle \nu | \frac{\partial e^{-T}}{\partial t_\lambda} H\, e^T |0\rangle \,+\, \langle \nu | e^{-T} H \,\frac{\partial}{\partial e^T t_\lambda} |0\rangle \\ &= \langle \nu | e^{-T} [H, T_\lambda] e^T |0\rangle \end{aligned} \tag{28.24}$$

where the square bracket stands for the commutator.

Evaluation of the matrix elements occurring in (28.24), in principle, can be performed by standard many-body techniques or by applying the more recent automated implementation philosophy[40, 41, 42, 43, 44, 45]. The only non-usual ingredient of these formulae is that state $\langle \nu |$ denotes not only an SD, but also higher excited configurations. Fortunately, 6-fold excitations are the highest which may contribute[17]. One can also restrict $\langle \nu |$ at a selected lower maximal excitation level, quadruples for example. The need for higher excitations makes the evaluation of matrix elements quite involved, and the string based algorithm[40] discussed also in the present Volume can offer an enormous help in this.

Having the gradients, one can invoke one of the standard optimization routines to get optimized amplitudes. We have examined the performance of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) procedure[46] and the optimally conditioned (OC) method by Davidon[47]. They were found to perform similarly. As starting parameters for the optimization, standard CCSD amplitudes are convenient to use in most cases.

28.5.4 Numerical results

The model studies reported so far[48] were obtained not exactly with the above formulae, but with the energy-dependent form of the CC equations. Due to the complexity of the problem, small systems in very small basis sets were considered. Two-electron systems are of course excluded, since for them CCSD is equivalent to full CI.

Consider first a simple four-electron example, the Be atom. Table 28.1 presents standard and moment-optimized CCSD energies in comparison with full CI. The expectation value of the Hamiltonian computed with the actual wave function is also tabulated. To measure the accuracy of the wave function, we give the norm of the

$$|r\rangle = (H - \langle H\rangle)\,|\Psi\rangle$$

residual vector.

Table 28.1 clearly shows that balancing SD and higher moments, results in a slight energy loss. However, the residual norm indicates that overall accuracy of the wave function improves. For Be, this improvement is quite small, it is some 8 % in the minimal basis, only 1% in the split shell basis, and less than 4 % in the polarized basis set. Similar experience was gained on other small systems.

Table 28.1 Effect of moment-optimized amplitudes on the Be atom. E(CCSD) corresponds to the standard CC energy formula, $\langle H\rangle$ is the expectation value of the Hamiltonian computed by the given wave function, $\sqrt{\langle r|r\rangle}$ is the norm of the residual vector.

basis set	level	E(CCSD)	$\langle H\rangle$	$\sqrt{\langle r\|r\rangle}$
STO-6G				
	HF		-14.503361	0.16097
	CCSD	-14.556084	-14.556086	0.00510
	opt-CCSD	-14.556076	-14.556086	0.00472
	FCI		-14.556089	0.0
6-31G				
	HF		-14.566764	0.17153
	CCSD	-14.613518	-14.613518	0.01399
	opt-CCSD	-14.613085	-14.613515	0.01383
	FCI		-14.613545	0.0
6-31G**				
	HF		-14.566944	0.26110
	CCSD	-14.616483	-14.616487	0.03588
	opt-CCSD	-14.614763	-14.616413	0.03462
	FCI		-14.616635	0.0

The slight improvement shown by the residual norm may become important in special cases. Dissociation curves computed by breaking more than one bond at a time were found to constitute such an example. We report here the results for the nitrogen molecule, where a triple bond rupture is monitored. Figure 1 shows the case of the N_2 molecule in 6-31G** basis, with the four lowest canonical molecular orbitals (MO) kept frozen. The standard CCSD curve is quite pathologic in this case, an effect well known from previous studies[19, 25, 27, 30, 49, 50, 51, 52]. When determining CCSD amplitudes from moment minimization, the erratic behavior of the curve is greatly improved, and the moment optimized CCSD potential curve gets much closer to full CI. The curves can be compared to those showing the expectation value with the CCSD wave function, depicted also in Fig. 1. With standard CCSD wavefunction the expectation value gives an acceptable estimation till about $R \sim 2$ Å. When computing the expectation value with moment-optimized CCSD, one gets a curve that is the closest to full CI.

It is notable that in all basis sets collected in Table 28.1 we see an increase of the CCSD energy (often getting away from the exact value). At the same time the wavefunction improves in the residual norm sense. This fact is tolerable, regarding the non-variational character of the CC Ansatz. However, it is interesting that not only the moment-like CC energy, but also the expectation value of the Hamiltonian may increase (cf. Fig. 1.). This clearly indicates that we are still far form the region of validity of the Eckart theorem[53] which, for non-degenerate states would require the simultaneous improvement of the energy and the wave function.

To gain a better insight into the effect of moment optimization, we have collected the moments larger than 0.005 in absolute value for water in Table 28.2. First

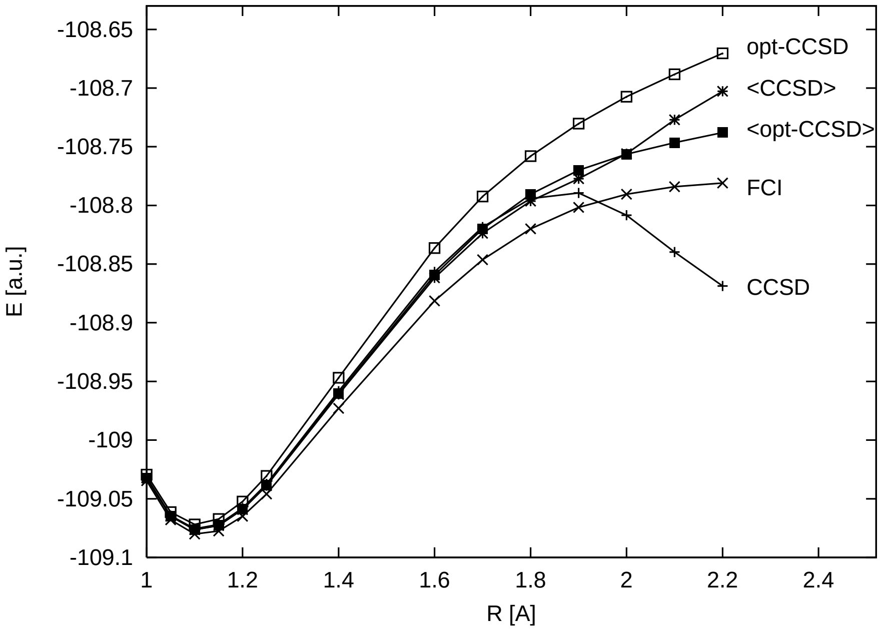

Fig. 28.1 Potential curve of the nitrogen molecule in split-valence polarized basis set. Methods used are standard and moment-optimized CCSD. Expectation values taken with both wavefunctions are also depicted. For abbreviations see text.

few lines of the Table show various double excitations. The associated moments are exactly zero in standard CCSD, as they should be. The last three lines correspond to quadruple excitations, among which the 4545 → 6767 frontier excitation has an enormous moment, almost 0.25. Certainly, in this small basis set appearance of this huge moment is responsible for the failure of CCSD at this geometry. Upon minimizing all moments democratically, we find that this large moment is greatly reduced, at the price of generating small moments in the SD space. Upon minimization the other two quadruple moments also diminish by an order of magnitude.

In summarizing, we have investigated the quality of the CCSD wave function and the associated energy upon performing an unconstrained optimization of the sum of moment squares with respect to CCSD amplitudes. The numbers presented above permit us to draw two general conclusions: (i) in each case studied the CCSD energy raises when computed with moment-optimized amplitudes, while (ii) the wave function improves in a residual norm sense, if compared with standard CCSD. This observation is quite significant, particularly when CCSD fails due to its non-variational behavior, like the cases of stretched multiple bonds.

Table 28.2 Collection of the largest moments for the water molecule at stretched geometry, R_{OH}=2.5 Å, before and after optimization. Excitations are identified by the serial numbers of the MOs involved. MOs 4 and 5 are the symmetric and antisymmetric combinations of OH bonding orbitals, 6 and 7 are their virtual counterparts. MO 2 is the σ lone pair on the oxygen atom.

excitation	standard CCSD	moment-optimized CCSD
$45 \rightarrow 67$ (α,α)	0.0000	0.0197
$45 \rightarrow 67$ (α,β)	0.0000	-0.0146
$55 \rightarrow 66$	0.0000	0.0249
$55 \rightarrow 77$	0.0000	-0.0178
$44 \rightarrow 66$	0.0000	-0.0179
$44 \rightarrow 77$	0.0000	0.0239
$4545 \rightarrow 6767$	0.2461	0.0130
$2525 \rightarrow 6767$	0.0169	0.0007
$2424 \rightarrow 6767$	0.0183	0.0023

28.6 Improving CC results: perturbation theory

The usual formulation of Rayleigh-Schrödinger perturbation theory (PT) relies on the availability of the full set of right- and left eigenvectors of a zero order Hamiltonian. In multi-reference perturbation theory this requirement is relaxed, and perturbed wave functions are often expanded in auxiliary bases. In most formulations, the symmetry between bra and ket vectors is kept, however. In coupled cluster theory, it is appropriate to abandon this latter constraint too, owing to the essentially non-symmetric (moment-like) nature of the formalism. A general perturbative tool that can be routinely applied even is such cases is the so-called multi-configuration perturbation theory (MCPT) [54, 55, 56]. A short review of this formalism as applied to a coupled cluster wavefunction is presented below.

28.6.1 Nonsymmetric PT formulation

Consider a multiconfigurational reference state $|\mathrm{CC}\rangle$ in intermediate normalization:

$$|\mathrm{CC}\rangle \;=\; |\mathrm{HF}\rangle \;+\; \sum_{k=1} d_k\,|k\rangle\;,$$

where $|k\rangle$ denotes determinants obtained by applying single, double etc. excitations to the Fermi vacuum $|\mathrm{HF}\rangle$. Determinant $|k\rangle$ with $k = 0$ will be identified as the Fermi vacuum. Coefficients d_k can be obtained by conversion of CC amplitudes to CI coefficients[57], e.g.

$$d_{ij}^{ab} \;=\; t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a$$

if determinant k is associated with excitation $i, j \rightarrow a, b$.

A skew-projector $\hat{O}$ can be taken down in terms of the coupled-cluster wavefunction and the Fermi-vacuum in the following form

$$\hat{O} \; = \; |\mathrm{CC}\rangle\langle\mathrm{HF}| \, .$$

It is easily seen that operator $\hat{O}$ is idempotent and leaves $|\mathrm{CC}\rangle$ and $\langle\mathrm{HF}|$ intact. The projector orthogonal and complementary to $\hat{O}$ can be defined as

$$\hat{P} \; = \; \hat{1} - \hat{O}$$

It is also trivially verified that $\hat{P}|\mathrm{CC}\rangle$ results zero as well as $\langle\mathrm{HF}|\hat{P}$ while

$$\hat{P}|k\rangle \; = \; |k\rangle$$

for $k \geq 1$. A spectral form of projector $\hat{P}$ can be constructed, starting from the representation of the identity

$$\hat{1} \; = \; \sum_{k=0} |k\rangle\langle k|$$

and writing

$$\hat{P} \; = \; \hat{1}\hat{P} \; = \; \sum_{k=1} |k\rangle\langle k|\hat{P} \tag{28.25}$$

where $k = 0$ is omitted from the sum since $\langle\mathrm{HF}|\hat{P} = 0$. Expression (28.25) is in fact a biorthogonal spectral resolution of operator $\hat{P}$ formulated in terms of direct space vectors $|k\rangle$ and reciprocal space vectors

$$\langle\tilde{k}| \; = \; \langle k|\hat{P} \; = \; \langle k| \; - \; d_k\langle\mathrm{HF}| \, .$$

It is indeed straightforward to see, that vectors $|\mathrm{CC}\rangle$ and $|k\rangle$ for $k \geq 1$ form a biorthogonal set with vectors $\langle\mathrm{HF}|$ and $\langle\tilde{k}|$, satisfying the conditions

$$\begin{array}{rl} \langle\mathrm{HF}|\mathrm{CC}\rangle = 1 & \quad \langle\mathrm{HF}|l\rangle = 0 \\ \langle\tilde{k}|\mathrm{CC}\rangle = 0 & \quad \langle\tilde{k}|l\rangle = \delta_{kl} \end{array}$$

with $k, l \geq 1$. In terms of the above direct and reciprocal space vectors a non-Hermitean zero-order Hamiltonian can be taken down in the form

$$\hat{H}^0 \; = \; E_{\mathrm{CC}}\,|\mathrm{CC}\rangle\langle\mathrm{HF}| \; + \; \sum_{k=1} E_k\,|k\rangle\langle\tilde{k}| \tag{28.26}$$

where $E_{\mathrm{CC}} = \langle\mathrm{HF}|\hat{H}|\mathrm{CC}\rangle$ is the coupled-cluster energy. The excited state energies E_k are parameters of the theory. They are in principle arbitrary quantities that define $\hat{H}^0$, i.e. the *partitioning*. The above definition of $\hat{H}^0$ possesses the properties

$$\hat{H}^0\,|\mathrm{CC}\rangle \; = \; E_{\mathrm{CC}}\,|\mathrm{CC}\rangle \tag{28.27}$$

and

$$\hat{H}^0\,|k\rangle \; = \; E_k\,|k\rangle \, .$$

One can similarly see, that left eigenvectors of operator (28.26) are vectors $\langle\tilde{k}|$ and the principal determinant $\langle\mathrm{HF}|$.

Perturbation theory using a biorthogonal vector set as constructed above differs from usual Rayleigh-Schrödinger PT only in that reciprocal (biorthogonal) functions are used in the bra vectors of each matrix element. (For applications of biorthogonal PT in the theory of chemical bond and intermolecular interactions, see Refs. [58, 59, 60].) Accordingly, for the first order energy correction of the ground state E_{CC} we get:

$$E^{(1)} \;=\; \langle\mathrm{HF}|\hat{W}|\mathrm{CC}\rangle \;=\; 0$$

having utilized Eq.(28.27). The second order looks

$$E^{(2)} = -\sum_{k=1} \frac{\langle\mathrm{HF}|\hat{W}|k\rangle\langle\tilde{k}|\hat{W}|\mathrm{CC}\rangle}{E_k - E_{\mathrm{CC}}} \tag{28.28}$$

with $\hat{W} = \hat{H} - \hat{H}^0$. Higher order PT corrections can be put down similarly.

Before discussing the question of partitioning, let us analyse the numerator of the second order energy (28.28). One readily sees that $\hat{W}$ can be substituted for $\hat{H}$ in the matrix elements, regarding that $|\mathrm{CC}\rangle$ and $\langle\mathrm{HF}|$ are eigenvectors to $\hat{H}^0$ from the right and left, respectively. Matrix element $\langle\mathrm{HF}|\hat{H}|k\rangle$ ensures that at most doubly excited determinants can contribute to the expression, while matrix element

$$\langle\tilde{k}|\hat{H}|\mathrm{CC}\rangle \;=\; \langle k|\hat{H} - E_{\mathrm{CC}}|\mathrm{CC}\rangle = 0 \tag{28.29}$$

sets all the terms zero as a result of the converged CCSD equations. The first nonzero energy correction of the theory is therefore the third order term

$$E^{(3)} = \sum_{kl} \frac{\langle\mathrm{HF}|\hat{W}|k\rangle\langle\tilde{k}|\hat{W}|l\rangle\langle\tilde{l}|\hat{W}|\mathrm{CC}\rangle}{(E_k - E_{\mathrm{CC}})\,(E_l - E_{\mathrm{CC}})}\;.$$

Energy corrections obtained by the above theory are tabulated up to order 5 for the HF molecule in 6-31G* basis set in Table 28.3. Partitioning used in this calculation is given by e.g.

$$E_k \;=\; E_{\mathrm{CC}} + \varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j \tag{28.30}$$

if determinant k arises upon excitation $i,j \longrightarrow a,b$ acting on the Fermi vacuum. One particle energies ε_p are Hartree-Fock orbital energies. When willing to describe a situation where single-reference theories fail, e.g. bond breaking, it is wise to make a different choice for one particle energies. In this case ionization potentials and electron affinities have been successfully used instead of Koopmans values[61, 62, 63].

Inspecting data of Table 28.3 one sees that the third order correction represents a considerable improvement upon the CCSD energy. As the covalent bond is stretched, the third order energy deteriorates, which is a consequence of the partitioning choice. The error increases from about a tenth of a milliHartree around equilibrium distance to about one milliHartree at twice the equilibrium length. At

$2.0R_e$ the third order approximation is worse than the CCSD(T) result. Regarding higher orders, it is notable that correction of order 4 does not represent any improvement, and the PT series manifestly oscillates around the exact value.

Table 28.3 PT energy errors $E - E_{\mathrm{FCI}}$ [mH] for the H-F molecule in 6-31G* basis set, frozen core approximation (R_e=0.917 Å). Reference function is a CCSD wavefunction. Full CI energies in Hartree are -100.188421 (R_e), -100.108796 ($1.5R_e$), -100.034004 ($2.0R_e$).

	geometry		
order	R_e	$1.5R_e$	$2.0R_e$
nonsymmetric theory			
0	2.146	4.313	9.474
3	-0.074	0.146	1.386
4	0.312	1.108	3.773
5	-0.080	-0.155	-0.170
symmetric theory			
0	1.93	3.939	8.437
2	0.279	0.636	1.553
3	0.160	0.427	1.121
CCSD(T)	0.457	0.840	0.275

28.6.2 Symmetric PT formulation

A somewhat better behaving PT series is obtained if turning to a more symmetrical PT framework. This is however computationally demanding as it starts with a symmetrical projector of the form

$$\hat{Q} = \frac{|\mathrm{CC}\rangle\langle\mathrm{CC}|}{\langle\mathrm{CC}|\mathrm{CC}\rangle}$$

where the norm of the coupled-cluster wavefunction is not computable for anything but small model problems. If pursuing the theory even with such computational limitations in mind, one takes down the orthogonal and complementary projector to $\hat{Q}$

$$\hat{R} = \hat{1} - \hat{Q}\,.$$

To find a spectral form of projector $\hat{R}$ let us evaluate the product $\hat{P}\hat{R}$, which leads to

$$\begin{aligned}\hat{P}\hat{R} &= \left(\hat{1} - |\mathrm{CC}\rangle\langle\mathrm{HF}|\right)\left(\hat{1} - \frac{|\mathrm{CC}\rangle\langle\mathrm{CC}|}{\langle\mathrm{CC}|\mathrm{CC}\rangle}\right)\\ &= \hat{1} - |\mathrm{CC}\rangle\langle\mathrm{HF}| = \hat{P}\,,\end{aligned}$$

while if multiplying the two projectors in reverse order, one gets

$$\hat{R}\hat{P} \;=\; \hat{R}\,.$$

Let us then write projector $\hat{R}$ as

$$\hat{R} \;=\; \hat{R}\hat{1}\hat{P} \;=\; \sum_{k=1} \hat{R}|k\rangle\langle k|\hat{P} \tag{28.31}$$

where again the term $k = 0$ is omitted from the sum, because $\langle k|\hat{P} = 0$. Just as in the previous case, one can observe the right hand side of expression (28.31) being of the form of a biorthogonal spectral resolution with direct space vectors

$$|k'\rangle \;=\; \hat{R}|k\rangle$$

and reciprocal vectors

$$\langle\tilde{k}'| \;=\; \langle k|\hat{P}\,.$$

In fact it can be verified that biorthogonal relations

$$\begin{aligned} \langle \mathrm{CC}|\mathrm{CC}\rangle/\langle \mathrm{CC}|\mathrm{CC}\rangle = 1 \qquad & \langle \mathrm{CC}|l'\rangle = 0 \\ \langle \tilde{k}'|\mathrm{CC}\rangle = 0 \qquad & \langle \tilde{k}'|l'\rangle = \delta_{kl} \end{aligned}$$

indeed hold for $k,l \geq 1$.

Zero-order Hamiltonian in terms of these direct and reciprocal space vectors is now formulated as

$$\hat{H}^0 \;=\; \overline{E}_{\mathrm{CC}}\,\frac{|\mathrm{CC}\rangle\langle \mathrm{CC}|}{\langle \mathrm{CC}|\mathrm{CC}\rangle} \;+\; \sum_{k=1} E_k\,|k'\rangle\langle\tilde{k}'| \tag{28.32}$$

with $\overline{E}_{\mathrm{CC}} = \langle \mathrm{CC}|\hat{H}|\mathrm{CC}\rangle/\langle \mathrm{CC}|\mathrm{CC}\rangle$. Using biorthogonal PT the energy terms arising from this zero-order choice are

$$E^{(1)} \;=\; \langle \mathrm{CC}|\hat{W}|\mathrm{CC}\rangle/\langle \mathrm{CC}|\mathrm{CC}\rangle \;=\; 0$$

$$E^{(2)} = -\frac{1}{\langle \mathrm{CC}|\mathrm{CC}\rangle}\sum_{k=1}\frac{\langle \mathrm{CC}|\hat{W}|k'\rangle\langle\tilde{k}'|\hat{W}|\mathrm{CC}\rangle}{E_k - E_{\mathrm{CC}}} \tag{28.33}$$

with $\hat{W} = \hat{H} - \hat{H}^0$. Higher orders can be constructed analogously.

Inspecting second order expression (28.33), one immediately sees, that matrix element $\langle\tilde{k}'|\hat{W}|\mathrm{CC}\rangle$ in the numerator is zero for singly or doubly excited determinants, for the reason (28.29). However, at difference with the previous PT formulation, the states contributing to $E^{(2)}$ are not restricted to singles and doubles. In this latter approach every determinant that may interact with the coupled-cluster function via the Hamiltonian contributes to the second order correction. This may involve the full configuration space and therefore it is highly impractical.

Note, that this approach is symmetric only as what concerns projector $\hat{Q}$. Projector $\hat{R}$ is still represented in the form of a biorthogonal spectral resolution, i.e.

direct and reciprocal vectors are different. When computing the illustrative examples shown in Table 28.3, we neglected primes and tildes in expression (28.33), and also in the third order formula. Hence the presented numbers should be considered only approximate. Second and third order energies shown in the table are obtained by the partitioning of Eq. (28.30). The results indicate that in the symmetric formulation the errors slowly but systematically decrease with increasing order. Note however that the first nonvanishing correction is better in the case of the nonsymmetric theory. The difference between second order by the symmetric theory and third order by the nonsymmetric approach is diminished as the interatomic distance is enlarged.

28.6.3 Connected moment expansion

The marked difference between symmetric and moment-like formulation of corrections to the coupled-cluster wavefunction has been observed in a different context also[64]. In this study the Horn-Weinstein function

$$f(\alpha) = \frac{\langle \mathrm{CC}|\hat{H}e^{-\alpha\hat{H}}|\mathrm{CC}\rangle}{\langle \mathrm{CC}|e^{-\alpha\hat{H}}|\mathrm{CC}\rangle}$$

and its momentum-type analogue

$$\tilde{f}(\alpha) = \frac{\langle \mathrm{HF}|\hat{H}e^{-\alpha\hat{H}}e^{\hat{T}}|\mathrm{HF}\rangle}{\langle \mathrm{HF}|e^{-\alpha\hat{H}}e^{\hat{T}}|\mathrm{HF}\rangle}$$

constitute the basis of the approximations. Both of the above functions tend to the ground state eigenvalue of $\hat{H}$ as $\alpha \to \infty$. Derivative of f with respect to α evaluated at $\alpha = 0$ are related to the so-called connected moments of the Hamiltonian

$$\begin{aligned} f(0) = I_1 &= \langle\hat{H}\rangle/\mathcal{N} \\ -\left.\frac{\mathrm{d}f}{\mathrm{d}\alpha}\right|_{\alpha=0} = I_2 &= \langle\hat{H}^2\rangle/\mathcal{N} - \langle\hat{H}\rangle^2/\mathcal{N}^2 \\ -\left.\frac{\mathrm{d}^2 f}{\mathrm{d}\alpha^2}\right|_{\alpha=0} = I_3 &= \langle\hat{H}^3\rangle/\mathcal{N} - 3\langle\hat{H}^2\rangle\langle\hat{H}\rangle/\mathcal{N}^2 + 2\langle\hat{H}\rangle^3/\mathcal{N}^3 \end{aligned}$$

and similarly for higher derivatives. Here, we used $\langle \mathrm{CC}|\hat{H}^k|\mathrm{CC}\rangle = \langle\hat{H}^k\rangle$ and the squared norm of the coupled cluster function is denoted by $\mathcal{N}$. The connected moments can be used for constructing a successive approximation to the ground state energy e.g. in the form[65]

$$E = I_1 - I_2^2/I_3 - \ldots \tag{28.34}$$

The analogous treatment of function $\tilde{f}$ leads to another infinite series for the ground state energy,

$$E \;=\; \tilde{I}_1 - \tilde{I}_2^2/\tilde{I}_3 - \ldots \tag{28.35}$$

which is considerably cheaper to compute as the coupled-cluster function never figures in the bra vectors of the modified connected moments

$$\begin{aligned} \tilde{f}(0) = \tilde{I}_1 \;&=\; \langle e^{-\hat{T}}\hat{H}e^{\hat{T}}\rangle \;=\; \langle\overline{H}\rangle \\ -\;\frac{\mathrm{d}\tilde{f}}{\mathrm{d}\alpha}\bigg|_{\alpha=0} = \tilde{I}_2 \;&=\; \langle\overline{H}^2\rangle \;-\; \langle\overline{H}\rangle^2 \\ -\;\frac{\mathrm{d}^2 f}{\mathrm{d}\alpha^2}\bigg|_{\alpha=0} = I_3 \;&=\; \langle\overline{H}^3\rangle \;-\; 3\langle\overline{H}^2\rangle\langle\overline{H}\rangle \;+\; 2\langle\overline{H}\rangle^3 \end{aligned}$$

where we used $\overline{H} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$ and $\langle \mathrm{HF}|\overline{H}^k|\mathrm{HF}\rangle = \langle\overline{H}^k\rangle$. In analogy with the PT formulation, it can be shown that the second connected moment in the nonsymmetric theory is exactly zero due to the fulfillment of the coupled-cluster equations. As it was found in Ref.[64], only the impractical symmetric expansion (28.34) offers good corrections, while the accuracy of the nonsymmetric expansion (28.35) is insufficient.

Regarding the coupled cluster wavefunction as a multiconfigurational reference, systematic improvement is possible to obtain by the MCPT approach, either following the symmetric or the nonsymmetric formulation. Noniterative corrections can be also derived by means of connected moments. Both schemes show that numbers obtained from a symmetric version are more reliable, though these methods can not be used in practice. The third (first nonvanishing) order of the nonsymmetric MCPT seems to offer a reliable and practical tool.

Acknowledgements The authors thank D. Mukherjee, Z. Rolik and P. Szakács for extensive discussions. Partial support from grants OTKA K-81588, K-81590. NI-67702 and Tet-Ind04/2006 is acknowledged.

References

1. G.E. Scuseria, T.J. Lee, H.F.S. III, Chem. Phys. Letters **130**, 236 (1986)
2. K. Kowalski, K. Jankowski, Phys. Rev. Letters **81**, 1195 (1998)
3. Granino A. Korn and Thresa M. Korn, *Mathematical Handbook for Scientists and Engineers* (McGraw Hill, New York, 1968)
4. E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, 1993)
5. P. Szakács, P.R. Surján, J. Math. Chem. **43**, 314 (2008)
6. I. Lindgren, J. Morrison, *Atomic Many-Body Theory* (Springer, Berlin, 1986)

7. P.O. Löwdin, Int. J. Quantum Chem. **72**, 370 (1999)
8. D. Kőhalmi, Á. Szabados, P.R. Surján, Phys. Rev. Letters **95**, 13002 (2005)
9. Z. Szekeres, P.G. Mezey, P.R. Surján, Chem. Phys. Letters **424**, 420 (2006)
10. G. E. Scuseria and C. L. Janssen and H. F. Schaefer, J. Chem. Phys. **89**, 7382 (1988)
11. P. Szakács, P.R. Surján, Int. J. Quantum Chem. **000**, 000 (2008)
12. J.Geertsen, M. Rittby, R.J. Bartlett, Chem. Phys. Letters **164**, 57 (1989)
13. J.F. Stanton, R.J. Bartlett, J. Chem. Phys. **98**, 7029 (1993)
14. L. Meissner, R.J. Bartlett, J. Chem. Phys. **102**, 7490 (1995)
15. O. Christiansen, H. Koch, P. Jørgensen, Chem. Phys. Letters **243**, 409 (1995)
16. H. Koch, O. Christiansen, P. Jørgensen, A.M.S. de Meras, T. Helgaker, J. Chem. Phys. **106**, 1808 (1997)
17. K. Jankowski, J. Paldus, P. Piecuch, Theor. Chim. Acta **80**, 223 (1991)
18. X. Li, J. Paldus, J. Chem. Phys. **115**, 5759 (2001)
19. X. Li, J. Paldus, J. Chem. Phys. **115**, 5774 (2001)
20. X. Li, J. Paldus, J. Chem. Phys. **117**, 1941 (2002)
21. X. Li, J. Paldus, J. Chem. Phys. **118**, 2470 (2003)
22. J. Paldus, X. Li, Collection of Czechoslovak Chemical Communications **68**, 554 (2003)
23. J. Paldus, X. Li, International Journal of Modern Physics B **17**, 5379 (2003)
24. K. Kowalski, P. Piecuch, Mol. Phys. **102**, 2425 (2004)
25. K. Kowalski, P.D. Fan, P. Piecuch, Mol. Phys. **103**, 2191 (2005)
26. K. Kowalski, P. Piecuch, J. Chem. Phys. **113**, 18 (2000)
27. K. Kowalski, P. Piecuch, J. Chem. Phys. **113**, 5644 (2000)
28. P. Piecuch, K. Kowalski, I.S.O. Pimienta, M.J. McGuire, Int. Reviews in Physical Chemistry **21**, 527 (2002)
29. I.S.O. Pimienta, K. Kowalski, P. Piecuch, J. Chem. Phys. **119**, 2951 (2003)
30. P. Piecuch, K. Kowalski, I.S.O. Pimienta, P.D. Fan, M. Lodriguito, M.J. McGuire, S.A. Kucharski, T. Kuś, M. Musiał, Theor. Chem. Acc. **112**, 349 (2004)
31. P. Piecuch, M. Włoch, J.R. Gour, A. Kinal, Chem. Phys. Letters **418**, 463 (2005)
32. P. Piecuch, M. Włoch, J. Chem. Phys. **123**, 224105 (2005)
33. I. Ozkan, A. Kinal, M. Balci, J. Phys. Chem. A **108**, 507 (2004)
34. M.J. McGuire, P. Piecuch, J. Am. Chem. Soc. **127**, 2608 (2005)
35. C.J. Cramer, M. Włoch, P. Piecuch, C. Puzzarini, L. Gagliardi, J. Phys. Chem. A **110**, 1991 (2006)
36. P.G. Szalay, M. Nooijen, R.J. Bartlett, J. Chem. Phys. **103**, 281 (1995)
37. E. Szondy, T. Szondy, Acta Phys.Hung. **20**, 253 (1966)
38. M. Hegyi, M. Mezei, T. Szondy, Theor. Chim. Acta **21**, 168 (1971)
39. K. Ladányi, V. Lengyel, T. Szondy, Theor. Chim. Acta **21**, 176 (1971)
40. M. Kállay, P.R. Surján, J. Chem. Phys. **115**, 2945 (2001)
41. S. Hirata, P.D. Fan, A.A. Auer, M. Nooijen, P. Piecuch, J. Chem. Phys. **121**, 12197 (2004)
42. A. Hartono, A. Sibiryakov, M. Nooijen, G. Baumgartner, D.E. Bernholdt, S. Hirata, C.C. Lam, R.M. Pitzer, J. Ramanujam, P. Sadayappan, Lecture Notes in Computer Science **3514**, 155 (2005)
43. S. Hirata, J. Phys. Chem. A **107**, 9887 (2003)
44. S. Hirata, J. Chem. Phys. **121**, 51 (2004)
45. P. Piecuch, S. Hirata, K. Kowalski, P.D. Fan, T.L. Windus, Int. J. Quantum Chem. **106**, 79 (2006)
46. R. Fletcher, Computer J. **13**, 317 (1970)
47. W.C. Davidon, Mathematical Programming **9**, 1 (1975)
48. Z. Rolik, D.K. Á. Szabados, P.R. Surján, J. Mol. Struct. (THEOCHEM) **768**, 1723 (2006)
49. T.V. Voorhis, M. Head-Gordon, Chem. Phys. Letters **317**, 575 (2000)
50. T.V. Voorhis, M. Head-Gordon, J. Chem. Phys. **115**(17), 7814 (2001)
51. P. Piecuch, V. Spirko, A.E. Kondo, J. Paldus, J. Chem. Phys. **104**, 4699 (1996)
52. W.D. Laidig, P. Saxe, R.J. Bartlett, J. Chem. Phys. **86**, 887 (1987)
53. C.E. Eckart, Phys. Rev. **36**, 878 (1930)
54. P.R. Surján, Á. Szabados, Z. Szekeres, Int. J. Quantum Chem. **90**, 1309 (2002)

55. Z. Rolik, Á. Szabados, P.R. Surján, J. Chem. Phys. **119**, 1922 (2003)
56. Á. Szabados, Z. Rolik, G. Tóth, P.R. Surján, J. Chem. Phys. **122**, 114104 (2005)
57. I. Mayer, *Simple Theorems, Proofs, and Derivations in Quantum Chemistry* (Kluwer, New York, 2003)
58. P. R. Surján, I. Mayer, I. Lukovits, Phys. Rev. A **32**, 748 (1985)
59. P.R. Surján, I. Mayer, I. Lukovits, Chem. Phys. Letters **119**, 538 (1985)
60. P.R. Surján, I. Mayer, J. Mol. Struct. (THEOCHEM) **226**, 47 (1991)
61. A. Zaitevskii, J.P. Malrieu, Chem. Phys. Letters **233**, 597 (1995)
62. P.R. Surján, Á. Szabados, Int. J. Quantum Chem. **69**, 713 (1998)
63. Y. Mochizuki, Chem. Phys. Letters **472**, 143 (2009)
64. J. Noga, Á. Szabados and P. R. Surján, Int. J. Mol. Sci. **3**, 508 (2002)
65. J. Cioslowski, Phys. Rev. Lett. **58**, 83 (1987)