Theor Chim Acta (1993) 86: 41–81

Theoretica Chimica Acta

Stationary perturbation theory

II. Electron correlation and its effect on properties

Werner Kutzelnigg

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, W-4630 Bochum, Germany

Received June 15, 1992/Accepted September 23, 1992

Summary. After a short recapitulation of the basic concepts of stationary perturbation theory, this is applied to a many-electron Hamiltonian, with or without an external field, given in a Fock space formulation in terms of a finite basis, the exact eigenfunctions of which are the full-CI wave functions. The Lie algebra $\mathscr{L}_c^{(n)}$ of the variational group corresponding to this problem is presented. It has an important subalgebra $\mathscr{L}_c^{(1)}$ of one-particle transformations. Hartree–Fock and coupled Hartree–Fock (also uncoupled Hartree–Fock) as well as MC-SCF and coupled MC-SCF are outlined in this framework. Many-body perturbation theory and Møller–Plesset perturbation theory are derived from the same kind of stationarity condition and a new non-perturbative iteration construction of the full-CI wave function is proposed, the first Newton–Raphson iteration cycle of which is CEPA-0. For the treatment of electron correlation for properties two variants of Møller–Plesset theory referred to as 'coupled' (CMP) and 'uncoupled' (UCMP) are defined, neither of which is fully satisfactory. While CMP satisfies a Brillouin condition, which implies that first order correlation corrections to first- and second-order properties vanish, it does not satisfy a Hellmann–Feynman theorem, i.e. a first order property is *not* the expectation value of the operator associated with the property. Conversely UCMP satisfies a Hellmann–Feynman theorem but no Brillouin theorem. The incompatibility of the two theorems is related to an unbalanced treatment of one-particle- and higher excitations in MP theory. CMP, which is based on coupled Hartree–Fock as uncorrelated reference, appears to have slight advantages over UCMP, but neither variant looks very promising for the evaluation of 2nd order correlation corrections to 2nd-order properties. Then four variants of the perturbation theory of properties with a nonperturbative treatment of electron correlation on CEPA-0 level (but extendable to a higher level) are discussed. While those variants which are the direct counterpart of UCMP and CMP must be discarded, the 'perturbative CEPA-0' derived from a perturbative treatment on full-CI level appears to satisfy all important criteria, in particular it satisfies a Brillouin–Brueckner condition and a Hellmann–Feynman theorem. A simplified version, the 'coupled Brillouin–Brueckner CEPA-0' appears to have essentially the same qualities. It is important to replace the Brillouin condition of MP theory by the Brillouin–Brueckner condition in non-perturbative approaches, especially if one is interested in properties.

Key words: Stationary perturbation theory – Electron correlation – Coupled Møller–Plesset (CMP) – Uncoupled Møller–Plesset (UCMP) – Full CI–Coupled MC-SCF – Many-Body-Perturbation theory (MBPT) – Brillouin-condition – Brillouin–Brueckner-condition – Brueckner orbitals – Hellmann– Feynman theorem

1. Introduction

In paper I of this series [1], henceforth referred to as I, the basic concepts of stationary perturbation theory have been outlined. Unlike in the conventional formulation of perturbation theory it is not assumed that the 'unperturbed' Schrödinger equation is solved exactly, but rather that the energy expectation value is made stationary with respect to a family of infinitesimal unitary transformations that can be described by means of a unitary group $\mathscr{G}$, called the 'variational group', with which a Lie algebra $\mathscr{L}$ is associated. The most important results of I will be recapitulated in Sect. 2.

In the present paper we explicitly consider a many-electron Hamiltonian and we want to describe effects of electron correlation in the framework of stationary perturbation theory, both effects on the energy and on properties.

All considerations in this paper are based on a Hamiltonian formulated in a Fock space, based on a finite-dimensional one-electron basis as given in Sect. 3. The exact eigenfunctions of this Hamiltonian are the full-CI wave functions. The relevant Lie algebra is called $\mathscr{L}_c^{(n)}$ and is also presented in Sect. 3.

There are two main parts of this paper. The first of these (Sects. 6–10) deals with electron correlation in the absence of an external perturbation. After a general introduction to the correlation problem in the stationary context (Sect. 6) we discuss the treatment of electron correlation by perturbation theory (Sect. 7) in terms of a formal parameter λ that is finally set equal to 1. We study the two possibilities of many-body perturbation theory (MBPT) based on the bare nuclear Hamiltonian as unperturbed Hamiltonian, and Møller–Plesset (MP) theory where the unperturbed problem is Hartree–Fock theory. In Sect. 8 an essentially new approach to the iterative non-perturbative construction of the full-CI wave function is presented. This is then (Sect. 9) generalized to the case where the reference function is of MC-SCF type. This is somewhat related to coupled-cluster theory, but strictly in the stationary framework. In Sect. 10 the concept of Brueckner (best-overlap) orbitals in the context of stationary perturbation theory is introduced.

Sections 11–15 then deal with properties, i.e. we include an external perturbation $\mu\Omega$ to the Hamiltonian. An overview of the problem is given in Sect. 11, where also the more academic case of double perturbation theory based on the bare nuclear Hamiltonian is considered. The two variants of MP theory for properties, namely uncoupled Møller–Plesset theory UCMP, based on uncoupled Hartree–Fock (UCHF) and coupled Møller–Plesset theory (CMP), based on coupled Hartree–Fock (CHF) are studied in Sects. 12 and 13, respectively. Both have merits and drawbacks. While UCMP satisfies the Hellmann–Feynmann theorem, but not a Brillouin theorem, the converse is true for CMP. The incompatibility of these two properties is related to an unbalanced treatment of one-particle and two-particle operators in MP theory. Although it is usually preferable to satisfy a Hellmann–Feynman theorem, coupled MP theory appears to be a better choice as compared to UCMP. Anyhow for the most interesting case of 2nd order correlation correction to 2nd order properties double perturbation theory becomes so complicated that non-perturbative approaches look more promising. We then generalize (Sects. 14 and 15) the iterative non-perturbative approach of Sect. 8 to the case where an external perturbation is present. The formalism is somewhat simpler than in double perturbation MP theory, but again it is not trivial to find the appropriate way to formulate a fully satisfactory theory. It is, at least, required to start from a coupled Brueckner–Hartree–Fock theory in order to satisfy a Hellmann–Feynman theorem.

Before we discuss the topics just outlined we say a few words on coupled-Hartree–Fock theory, which is the 'non-correlated' reference for properties (Sect. 4), and on coupled MC-SCF theory (Sect. 5) which is an important intermediate step in the treatment of correlation effects on properties.

In 1955 Löwdin has published three classical papers [2], that are to a large extent the basis of current *ab-initio* quantum chemistry. The central importance of full-CI (not yet baptized so) and of Hartree–Fock as a standard approximation to it has been stressed. The present paper uses still similar concepts. However, it gives more weight to the stationary principle, the Lie-algebraic formulation and the study of properties.

2. Summary of stationary perturbation theory

We require that the expectation value of the Hamiltonian H is stationary with respect to a group $\mathscr{G}$ of unitary transformations:

$$\delta\langle\Psi|U^{-1}HU|\Psi\rangle = 0; \quad \forall U = (U^{\dagger})^{-1} \in \mathscr{G} \tag{2.1}$$

Any such U can be written as:

$$U = \exp X; \quad X = -X^{\dagger} \in \mathscr{L}_r \tag{2.2}$$

where $\mathscr{L}_r$ is the real Lie algebra (of antihermitean operators) associated with the 'variational group' $\mathscr{G}$ [1, 3].

A necessary and sufficient condition for Eq. (2.1) to hold is the (generalized) Brillouin or hypervirial condition:

$$\langle\Psi|[H, X]|\Psi\rangle = 0; \quad \forall X \in \mathscr{L}_r \tag{2.3}$$

The dependence of the wave function Ψ on a perturbation parameter λ (in the case that H depends on λ) is formulated as:

$$\Psi(\lambda) = e^{Y(\lambda)}\Psi_0; \quad Y(\lambda) \in \mathscr{L}_r \tag{2.4}$$

It is essential that the variational group $\mathscr{G}$ is chosen such that with the restrictive assumption (2.4) on $Y(\lambda)$ the dependence of Ψ on λ is well described.

One expands both the energy expectation value and the stationarity conditions in powers of λ, and one uses then the stationarity conditions to simplify the energy expressions. A detailed derivation is given in I [1]. We only consider here the case that H is linear in λ:

$$H(\lambda) = H_0 + \lambda V \tag{2.5a}$$

$$E(\lambda) = \langle\Psi(\lambda)|H(\lambda)|\Psi(\lambda)\rangle = \sum_{k=0}^{\infty} \lambda^k E_k \tag{2.5b}$$

$$Y(\lambda) = \sum_{k=1}^{\infty} \lambda^k Y_k \tag{2.5c}$$

$$C_0 \equiv \langle \Psi_0 | [H_0, X] | \Psi_0 \rangle = 0; \quad \forall X \in \mathscr{L}_r \tag{2.6a}$$

$$C_1 \equiv \langle \Psi_0 | [[H_0, X], Y_1] + [V, X] | \Psi_0 \rangle = 0; \quad \forall X \in \mathscr{L}_r \tag{2.6b}$$

$$C_2 \equiv \langle \Psi_0 | [[H_0, X], Y_2] + \tfrac{1}{2}[[H_0, X], Y_1], Y_1] + [[V, X], Y_1] | \Psi_0 \rangle = 0; \quad \forall X \in \mathscr{L}_r \tag{2.6c}$$

$$E_0 = \langle \Psi_0 | H_0 | \Psi_0 \rangle \tag{2.7a}$$

$$E_1 = \langle \Psi_0 | V | \Psi_0 \rangle \tag{2.7b}$$

$$E_2 = \tfrac{1}{2} \langle \Psi_0 | [V, Y_1] | \Psi_0 \rangle \tag{2.7c}$$

$$E_3 = \tfrac{1}{6} \langle \Psi_0 | [[[H_0, Y_1], Y_1], Y_1] | \Psi_0 \rangle + \tfrac{1}{2} \langle \Psi_0 | [[V, Y_1], Y_1] | \Psi_0 \rangle \tag{2.7d}$$

Equation (2.6a) is a condition for the unperturbed wave function Ψ_0, while Eqs. (2.6b) and (2.6c) are conditions for Y_1 and Y_2 respectively. Equation (2.7b) is recognized as the Hellmann–Feynman theorem. Equations (2.7c,d) demonstrate Wigner's $(2n + 1)$ rule according to which only the Y_k up to Y_n are needed to evaluate E_{2n} and E_{2n+1}.

Equation (2.6b) is equivalent to the requirement that the following functional:

$$F(Y_1) = \langle \Psi_0 | [V, Y_1] + \tfrac{1}{2}[[H_0, Y_1], Y_1] | \Psi_0 \rangle \tag{2.8}$$

is made stationary with respect to variations of Y_1. This is the equivalent of Hylleraas' variational principle in stationary perturbation theory.

The condition (2.6a) can be satisfied in the following way. We choose a trial function Φ_0 and make the ansatz

$$\Psi_0 = e^{\sigma} \Phi_0; \qquad \sigma = \sum_k c_k X_k; \quad X_k \in \mathscr{L}_r \tag{2.9}$$

Insertion of (2.9) into (2.6a) leads to a non-linear system for the expansion coefficients c_k. A possibility to solve this non-linear system consists of a Newton–Raphson type iterative scheme, in which one constructs a set of c_k from the linearized system, uses these to construct an approximate σ and an improved wave function from (2.9), after which one continues in the same way. For details see I. Sects. 2 and 6 [1].

For the evaluation of E_2 and E_3 the essential step is the construction of Y_1 from (2.6b). In view of (2.4) and (2.5a) we expand Y_1 in a basis $\{X_k\}$ of $\mathscr{L}_r$

$$Y_1 = \sum_k b_k X_k; \quad X_k \in \mathscr{L}_r \tag{2.10}$$

Insertion of (2.10) into (2.6b) leads to a linear system of equations

$$\sum_l H_{kl} b_l + V_k = 0 \tag{2.11a}$$

with

$$H_{kl} = \langle \Psi_0 | [[H_0, X_k], X_l] | \Psi_0 \rangle \tag{2.12a}$$

$$V_k = \langle \Psi_0 | [V, X_k] | \Psi_0 \rangle \tag{2.13}$$

The matrix H_{kl} is called the Hessean of the unperturbed problem. In I. Sect. 6 it has been discussed how to avoid that the Hessean becomes singular. This essentially amounts to choosing a subalgebra of $\mathscr{L}_c$, excluding operators for which Eq. (2.3) is trivially fulfilled.

It is often convenient to expand Y_1 in a basis $\{X_k\}$ of the complex Lie algebra $\mathscr{L}_c$ rather than the real subalgebra $\mathscr{L}_r$ of antihermitean operators. In this case a generalization of the definition (2.12a) is necessary that is valid for other than antihermitean X_k. This is:

$$H_{kl} = \langle \Psi_0 | [X_k^\dagger, [H_0, X_l]] | \Psi_0 \rangle \tag{2.12b}$$

If one defines V_k still by Eq. (2.13), the linear system of Eqs. (2.11a) has to be replaced by:

$$\sum_l H_{kl} b_l + V_k^* = 0 \tag{2.11b}$$

For $X_k \in \mathscr{L}_r$, Eqs. (2.12b) and (2.11b) reduce to Eqs. (2.12a) and (2.11a), respectively.

In the same framework double (and multiple) perturbation theory can be formulated. For

$$H(\lambda, \mu) = H_0 + \lambda V_{10} + \mu V_{01} \tag{2.14a}$$

$$Y(\lambda, \mu) = \sum_{k,l=0}^{\infty} \lambda^k \mu^l Y_{kl}; \quad (k = l = 0 \text{ excluded}) \tag{2.14b}$$

$$E = \sum_{k,l=0}^{\infty} \lambda^k \mu^l E_{kl} \tag{2.14c}$$

one gets (2.6a) and (2.7a) as before and

$$C_{10} \equiv \langle \Psi_0 | [[H_0, X], Y_{10}] + [V_{10}, X] | \Psi_0 \rangle = 0 \tag{2.15a}$$

$$C_{01} \equiv \langle \Psi_0 | [[H_0, X], Y_{01}] + [V_{01}, X] | \Psi_0 \rangle = 0 \tag{2.15b}$$

$$E_{10} = \langle \Psi_0 | V_{10} | \Psi_0 \rangle \tag{2.16a}$$

$$E_{01} = \langle \Psi_0 | V_{01} | \Psi_0 \rangle \tag{2.16b}$$

$$E_{11} = \langle \Psi_0 | [V_{10}, Y_{01}] | \Psi_0 \rangle = \langle \Psi_0 | [V_{01}, Y_{10}] | \Psi_0 \rangle \tag{2.16c}$$

$$\begin{aligned} E_{21} &= \tfrac{1}{2} \langle \Psi_0 | 2[[V_{10}, Y_{10}], Y_{01}] + [[V_{01}, Y_{10}], Y_{10}] + [[[H_0, Y_{10}], Y_{10}], Y_{01}] | \Psi_0 \rangle \\ &= \tfrac{1}{2} \langle \Psi_0 | 2[[V_{10}, Y_{01}], Y_{10}] + [[V_{01}, Y_{10}], Y_{10}] + [[[H_0, Y_{01}], Y_{10}], Y_{10}] | \Psi_0 \rangle \end{aligned} \tag{2.16d}$$

For other E_{kl} see I. Sect. 5 [1]. One recognizes (2.16c) as Dalgarno's exchange theorem [4].

In double perturbation theory one will often be interested in a Brillouin theorem or a Hellmann–Feynman theorem with respect to a single parameter. Consider

$$\begin{aligned} \lim_{\lambda \to 0} \langle \Psi | [H, X] | \Psi \rangle &= \lim_{\lambda \to 0} \langle \Psi_0 | e^{-Y} [H_0 + \mu V_{01}, X] e^{Y} | \Psi_0 \rangle \\ &= \langle \Psi_0 | [H_0, X] + \mu \{ [V_{01}, X] + [[X_0, X], Y_{01}] \} + O(\mu^2) | \Psi_0 \rangle = 0 \end{aligned} \tag{2.17}$$

This Brillouin theorem means that the coefficients of all powers of μ vanish for $\lambda = 0$, which is in fact the case in view of (2.6a) and (2.15a). Of course (2.17) holds up to that order in μ that one has considered explicitly. Let us now take a Hellmann–Feynman theorem, but with respect to the other parameter

$$\begin{aligned} \lim_{\mu \to 0} \frac{\partial E}{\partial \mu} &= E_{01} + \lambda E_{11} + \lambda^2 E_{21} + \cdots = \lim_{\mu \to 0} \langle \Psi | \frac{\partial H}{\partial \mu} | \Psi \rangle = \lim_{\mu \to 0} \langle \Psi_0 | e^{-Y} V_{01} e^{Y} | \Psi_0 \rangle \\ &= \langle \Psi_0 | V_{01} + \lambda [V_{01}, Y_{10}] + \lambda^2 \{ [V_{01}, Y_{20}] + \tfrac{1}{2} [[V_{01}, Y_{10}], Y_{10}] \} + O(\lambda^3) | \Psi_0 \rangle \end{aligned} \tag{2.18}$$

If we collect powers of μ and note that (2.16b,c,d) are satisfied, we realize that (2.18) does, in fact, hold. In the case of (2.16d) one must realize that this can alternatively be written as

$$E_{21} = \langle \Psi_0 | [V_{01}, Y_{20}] + \tfrac{1}{2}[[V_{01}, Y_{10}], Y_{10}] | \Psi_0 \rangle \tag{2.19}$$

which one verifies by first using (2.15b) for $X = Y_{20}$, then the Jacobi identity and (2.6a) and finally vanishing of C_{20} constructed in analogy to (2.15a,b).

One of the reasons why one prefers schemes which satisfy a Hellmann–Feynman theorem is that the error of a 1st order property is quadratic in the error of the wave function, otherwise only linear.

3. The variational group of full CI and its subgroups

Starting point for all aspects of this paper is the Fock space Hamiltonian [5, 6]

$$H = h_q^p a_p^q + \tfrac{1}{2} g_{rs}^{pq} a_{pq}^{rs} \tag{3.1}$$

written here in a tensor notation with the Einstein summation convention over repeated indices implied. In order to write H in the form (3.1) one must first specify a (necessarily finite) orthonormal one electron spin-orbital basis $\{\chi_p\}$. In terms of this basis, matrix elements h_q^p, g_{rs}^{pq} and creation and annihilation operators $a^p \equiv a_p^\dagger$ and a_q are defined and in terms of these excitation operators a_p^q, a_{pq}^{rs}.

$$h_q^p = \langle \chi_q | h | \chi_p \rangle \tag{3.2a}$$

$$g_{rs}^{pq} = \langle \chi_r(1) \chi_s(2) | r_{12}^{-1} | \chi_p(1) \chi_q(2) \rangle \tag{3.2b}$$

$$a_p^q = a^q a_p; \qquad a_{pq}^{rs} = a^r a^s a_q a_p \tag{3.2c}$$

The exact eigenfunctions of the Hamiltonian (3.1) are the 'full-CI' wave functions in the given one-electron basis $\{\chi_p\}$.

We shall consider essentially three variational groups and their associated Lie algebras.

(a) The group $\mathscr{G}^{(1)}$ of unitary one-particle transformations within the given spin-orbital basis, generated by the elements

$$a_p^q \in \mathscr{L}_c^{(1)} \tag{3.3a}$$

For spinfree problems it is advantageous to use instead the group of one-particle transformations between spinfree orbitals and the associated Lie-algebra with elements

$$E_P^Q = a_{P\alpha}^{Q\alpha} + a_{P\beta}^{Q\beta} \tag{3.3b}$$

(b) The group $\mathscr{G}^{(\infty)}$ of all unitary transformations of arbitrary particle rank generated by the elements

$$a_p^q,\ a_{pq}^{rs},\ a_{pqr}^{stu} \text{ etc.} \in \mathscr{L}_c^{(\infty)} \tag{3.4a}$$

(c) The group $\mathscr{G}^{(n)}$ of unitary transformations of n-electron wave function with fixed n. It can be generated by the $\mathscr{L}_c^{(n)}$ with elements Z_k such that

$$Z_k = Z_k' P_n = P_n Z_k'; \quad Z_k' \in \mathscr{L}_c^{(\infty)} \tag{3.5a}$$

where P_n is the projector to n-particle states. Since the Z_k' given by (3.4a) are particle-number conserving, they commute with P_k, which implies that

$$[Z_k, Z_l] = [Z_k', Z_l'] P_n \tag{3.5b}$$

If Z'_k has an excitation rank larger than n, the corresponding Z_k vanishes. The Lie algebra $\mathscr{L}_c^{(n)}$ hence breaks off at operators of particle rank larger than n. We shall not write the projector P_n explicitly, but assume whenever we write $Z \in \mathscr{L}_c^{(n)}$ that a Fock space operator (3.4a) is multiplied by P_n. Since, at the end we evaluate expectation values with an n-electron wave function, the P_n can be omitted.

Of course, $\mathscr{L}_c^{(1)}$ is a subalgebra of $\mathscr{L}_c^{(\infty)}$. This also means that the Lie-algebra of elements of $\mathscr{L}_c^{(1)}$ multiplied by P_n (*not* by P_1) is a subalgebra of $\mathscr{L}_c^{(n)}$. Since this 'modified' $\mathscr{L}_c^{(1)}$ leads to the same matrix elements as the original $\mathscr{L}_c^{(1)}$ we shall say that $\mathscr{L}_c^{(1)}$ is a subalgebra of $\mathscr{L}_c^{(n)}$.

Since the full-CI wave function can be obtained from any other wave function expressible in the same basis by means of a unitary transformation

$$U = e^{\sigma}; \quad \sigma = -\sigma^{\dagger} \in \mathscr{L}_c^{(n)} \tag{3.6}$$

the stationarity condition satisfied by the full-CI wave function is

$$\langle \Psi | [H, Z] | \Psi \rangle = 0; \quad \forall Z \in \mathscr{L}_c^{(n)} \tag{3.7}$$

Since $\mathscr{L}_c^{(1)}$ is a subalgebra of $\mathscr{L}_c^{(n)}$, we see that a necessary though, of course, not sufficient condition for a full-CI wave function is that

$$\langle \Psi | [H, X] | \Psi \rangle = 0; \quad \forall X \in \mathscr{L}_c^{(1)} \tag{3.8}$$

In fact (3.8) is not only fulfilled by the full-CI wave function, but also by a Hartree–Fock or an MC-SCF wave function.

It is generally easier to find a wave function that satisfies (3.8) than (3.7). It therefore makes sense to try first to satisfy (3.8) before in a second step one tries to improve Ψ such that it satisfies (3.7) as well. Often one will be able to satisfy (3.8) exactly, but (3.7) only approximately. This leads then to *some unbalance between the treatment of one-particle transformations and transformations of higher-particle rank*, which is very characteristic for many-electron theory. This unbalance causes problems in particular if in addition to the electron interaction another perturbation is present, i.e. in the theory of correlation corrections to properties (see Sects. 12, 13).

Of course, for spin-independent problems we may replace the operators (3.4a) by their spinfree counterparts [5]

$$E_P^Q, \; E_{PQ}^{RS}, \; E_{PQR}^{STU}, \text{ etc.} \tag{3.4b}$$

with E_P^Q defined by (3.3b) and the higher order excitation operator analogously. Here we shall only argue on spin-orbital level, so we don't need special names for the Lie algebras generated by (3.3b) or (3.4b) respectively.

The Lie algebra $\mathscr{L}_c^{(n)}$ generates the variational group of full CI. One important subalgebra has already been mentioned, namely that $\mathscr{L}_c^{(1)}$ of all one-particle excitations. Other subalgebras of $\mathscr{L}_c^{(n)}$ are generated by the operators of type (3.4a) defined for a subspace of the chosen space of spin-orbitals. In MC-SCF theory the subalgebra corresponding to excitations within the space of active spin-orbitals plays an important role (see Sect. 5 and I. Sect. 8).

We are especially interested in properties. To describe them we need an additional term $\mu\Omega$ to the Hamiltonian and we require that (3.7) holds for

$$H(\mu) = H + \mu\Omega \tag{3.9}$$

We shall henceforth use the symbol H for the full Hamiltonian (including the electron interaction) in the absence of the external field. The Hamiltonian including $\mu\Omega$ will be referred to as $H(\mu)$.

4. Coupled Hartree–Fock and uncoupled Hartree–Fock

Like Hartree–Fock theory is by definition the best treatment of a many-electron system in the framework of an independent-particle model – i.e. with complete neglect of electron correlation effects, coupled Hartree–Fock is in the same sense the best independent particle method for properties and hence the reference with respect to which correlation effects are defined.

A short outline of coupled Hartree–Fock has therefore to precede the treatment of correlation effects.

The main interest in this section is not to give a new derivation of Hartree–Fock and coupled Hartree–Fock theory, but rather to show in some detail how one arrives from the general Lie algebraic formulation of the variation principle and of stationary perturbation theory, by explicit insertions of a basis of the relevant Lie algebra, at working equations that can be used for computations. This is rather easily done in the present context, but can hardly be worked out in full length in the more complicated situations of later sections.

A formulation of Hartree–Fock theory essentially in the spirit of the present approach has probably first been given by Hinze and Roothaan [10] as well as Levy and Berthier [11].

What we want is to satisfy (3.8) for $H(\mu)$ given by (3.9) and H by (3.1). The formalism of I. Sect. 4 applies directly. We search for an unperturbed wave function Ψ_0 and a perturbing operator $Y(\mu)$ such that the μ-dependent trial wave function is of the form (2.4, 2.5c). We call the external perturbation $\mu\Omega$, because λV will be reserved for the electron interaction as perturbation. The Hartree–Fock wave function is determined by the condition (2.6a) and the 1st order operator Y_1 by (2.6b). The lowest orders of the energy are given by (2.7).

The reference wave function which we optimize by applying a one-particle transformation is usually a single Slater determinant, but we can also consider the more general case that it is a linear combination of Slater determinants with fixed coefficients (e.g. by symmetry).

Explicitly we get for (2.7a) (subscripts on E count orders in μ)

$$E_0 = h^p_q \gamma^q_p + \tfrac{1}{2} g^{pq}_{rs} \Gamma^{rs}_{pq} \tag{4.1}$$

where γ and Γ are the (unperturbed) one- and two-particles density matrices respectively

$$\gamma^q_p = \langle \Psi_0 | a^q_p | \Psi_0 \rangle; \qquad \Gamma^{pq}_{rs} = \langle \Psi_0 | a^{pq}_{rs} | \Psi_0 \rangle \tag{4.2}$$

The Brillouin condition (2.6a) becomes

$$\langle \Psi_0 | [H, a^t_u] | \Psi_0 \rangle = h^t_q \gamma^q_u - \gamma^t_q h^q_u + g^{tp}_{rs} \Gamma^{rs}_{up} \Gamma^{tp}_{rs} g^{rs}_{up} = 0 \tag{4.3}$$

In the special case of Ψ_0 a single Slater determinant, considerable simplifications arise

$$\gamma^p_q = \delta^p_q n_q; \qquad \Gamma^{pq}_{rs} = (\delta^p_r \delta^q_s - \delta^p_s \delta^q_r) n_p n_q; \quad n_p = 0 \text{ or } 1 \tag{4.4}$$

$$E_0 = h^p_p n_p + \tfrac{1}{2} \bar{g}^{pq}_{pq} n_p n_q \tag{4.5}$$

$$\bar{g}^{tq}_{rs} = g^{tq}_{rs} - g^{qt}_{rs} \tag{4.6}$$

provided, of course, that the basis functions χ_p used in (3.1, 2) are chosen so that the spin orbitals occupied in Ψ_0 are a subset of the basis. The stationary condition (4.3) can then be formulated as

$$f^s_r (n_r - n_s) = 0 \tag{4.7a}$$

with the matrix elements of the Fock operator

$$f^s_r = h^s_r + \bar{g}^{st}_{rt} n_t \tag{4.7b}$$

and n_p the occupation number (equal to 0 or 1) or the spin orbital ψ_p.

An alternative formulation [7, 8] is

$$[\underline{\underline{f}}, \underline{\underline{\gamma}}] = 0 \tag{4.7c}$$

i.e. that the matrix representation of the Fock operator (4.7b) commutes with the one-particle density matrix (4.2). $\underline{\underline{f}}$ and $\underline{\underline{\gamma}}$ must hence have the same eigenfunctions. This justifies an iterative method, where one starts with some $\underline{\underline{\gamma}}$, constructs from it a new $\underline{\underline{f}}$, from its eigenfunctions a new $\underline{\underline{\gamma}}$ etc. until self-consistency, as is usually done in closed-shell Hartree–Fock calculations.

We shall henceforth use the labels $i, j, k, \ldots$ for occupied spin orbitals ($n_i = 1$) and $a, b, c, \ldots$ for unoccupied ones ($n_a = 0$) $p, q, r, \ldots$ for arbitrary ones (in Sect. 5 we further introduce $x, y, z, \ldots$ for partially occupied spin orbitals). Then we can write the Brillouin condition (4.7a) as

$$f^i_a = f^a_i = 0 \tag{4.7d}$$

Since occupied and unoccupied spin orbitals don't interact, we can choose a set of occupied and unoccupied orbitals such that

$$f^p_q = \varepsilon_p \delta^p_q \tag{4.7e}$$

i.e. that the Fock operator is diagonal. We shall make this choice.

We now consider a Hamiltonian including the external perturbation $\mu\Omega$, i.e. we switch to coupled Hartree–Fock theory (CHF). The first order energy (in powers of μ) is in view of (2.7b)

$$E_1 = \langle \Psi_0 | \Omega | \Psi_0 \rangle \tag{4.8a}$$

We shall assume that Ω is a one-electron operator (for this choice coupled-Hartree–Fock-approximation is usually applied)

$$\Omega = \omega^p_q a^q_p \tag{4.9a}$$

$$E_1 = \omega^p_p n_p = \omega^i_i \tag{4.9b}$$

In order to get E_2 and E_3, the linear system (2.11b) has to be solved, with H_{kl} given by (2.12b) and V_k by (2.13). The basis operators X_k are now of the form a^q_p, i.e. each index k has to be replaced by a pair of indices p, q. We take $X_k = a^w_v$; $X_l = a^t_u$, then (2.11b) becomes

$$H^{vt}_{wu} b^u_t + \Omega^v_w = 0 \tag{4.10}$$

with

$$\Omega^v_w = -\langle \Psi_0 | [\Omega, a^v_w] | \Psi_0 \rangle = \omega^v_q \gamma^q_w - \omega^p_u \gamma^v_p = \omega^v_w (n_v - n_w) \tag{4.11}$$

$$\begin{aligned} H^{vt}_{wu} = \langle \Psi_0 | [a^v_w, [H, a^t_u]] | \Psi_0 \rangle &= h^t_w \gamma^v_u - \delta^v_u h^t_q \gamma^q_w - \delta^t_w h^q_u \gamma^v_q + h^v_u \gamma^t_w + g^{tp}_{ws} \Gamma^{rs}_{up} \\ &\quad + g^{tp}_{rw} \Gamma^{rv}_{up} - \delta^v_u g^{tp}_{rs} \Gamma^{rs}_{wp} - g^{tv}_{rs} \Gamma^{rs}_{uw} - g^{sr}_{wu} \Gamma^{vt}_{sr} - \delta^t_w g^{sr}_{pu} \Gamma^{pv}_{sr} + g^{vr}_{pu} \Gamma^{pt}_{wr} + g^{sr}_{pu} \Gamma^{pt}_{sw} \end{aligned} \tag{4.12a}$$

One sees that

$$H^{vt}_{wu} = H^{tv}_{uw}; \qquad H^{wu}_{vt} = (H^{vt}_{wu})^* \tag{4.13}$$

For the special case that Ψ_0 is a single Slater determinant, the expression (4.12a) simplifies greatly. One can derive this by inserting (4.4) into (4.12a), but it is easier to proceed as shown in the appendix by using the particle-hole formalism

$$H^{vt}_{wu} = (\delta^{v}_{u} f^{t}_{w} - \delta^{t}_{w} f^{v}_{u})(n_v - n_w) + \bar{g}^{vt}_{wu}(n_u - n_t)(n_v - n_w) \qquad (4.12b)$$

For the choice (4.7e) this reduces to

$$H^{vt}_{wu} = \delta^{v}_{u}\delta^{t}_{w}(\varepsilon_t - \varepsilon_u)(n_u - n_t) + \bar{g}^{vt}_{wu}(n_u - n_t)(n_v - n_w) \qquad (4.12c)$$

or with the convention *i*, *j*, *k* for occupied spin orbitals and *a*, *b*, *c* for empty ones, the only nonvanishing elements are:

$$H^{ja}_{bi} = \delta^{j}_{i}\delta^{a}_{b}(\varepsilon_a - \varepsilon_i) + \bar{g}^{ja}_{bi} \qquad (4.14a)$$

$$H^{ba}_{ji} = -\bar{g}^{ba}_{ji} \qquad (4.14b)$$

For the 'overlap matrix' Δ, that plays a role in the RPA approach (see I. Sect. 6.3) one gets

$$\Delta^{vt}_{wu} = \langle \Psi_0 | [a^{v}_{w}, a^{t}_{u}] | \Psi_0 \rangle = \delta^{t}_{w}\gamma^{v}_{u} - \delta^{v}_{u}\gamma^{t}_{w} = \delta^{t}_{w}\delta^{v}_{u}(n_u - n_t) \qquad (4.15a)$$

$$\Delta^{ia}_{ai} = 1; \qquad \Delta^{aa}_{ii} = -1 \qquad (4.15b)$$

The linear system of Eqs. (4.10) becomes for the general case

$$H^{ja}_{bi} b^{i}_{a} + H^{ji}_{ba} b^{a}_{i} + \Omega^{j}_{b} = 0 \qquad (4.16a)$$

$$H^{ba}_{ji} b^{i}_{a} + H^{ba}_{ji} b^{a}_{i} + \Omega^{b}_{j} = 0 \qquad (4.16b)$$

Noting that $\Omega^{b}_{j} = -(\Omega^{j}_{b})^{*}$ and using (4.11) one sees immediately that

$$b^{a}_{i} = -(b^{i}_{a})^{*} \qquad (4.17)$$

and that it is sufficient to solve only one of the two Eqs. (4.16). Inserting the explicit expressions (4.12c) and (4.11) into (4.16a) we get

$$(\varepsilon_j - \varepsilon_b) b^{j}_{b} = \omega^{j}_{b} + \bar{g}^{ja}_{bi} b^{i}_{a} - \bar{g}^{ji}_{ba} b^{a}_{i} \qquad (4.18)$$

It is convenient to solve this system iteratively, ignoring the coupling via $\bar{g}$ on the r.h.s. in the first iteration and using the $n-t$ iteration of the b on the r.h.s. in (4.18) to get those of the $(n+1)$st iteration. Complete neglect of these coupling terms defines 'uncoupled Hartree–Fock'. We come back to this approximation at the end of this section.

When self-consistency is reached for the system (4.18), the second order energy is obtained according to (2.7c) as

$$E_2 = \tfrac{1}{2}\langle \Psi_0 | [\Omega, Y_1] | \Psi_0 \rangle = \tfrac{1}{2}\omega^{p}_{q} b^{r}_{p} \gamma^{q}_{r} - \tfrac{1}{2}\omega^{p}_{r} b^{r}_{s} \gamma^{s}_{p} = \mathrm{Re}\, \omega^{a}_{i} b^{i}_{a} \qquad (4.19)$$

We have used here the formulation in terms of spin-orbitals, the spinfree formulation (to be preferred in practice) is only slightly more complicated.

In I. Sect. 7 it was pointed that there is some difference in the formalism for real and for imaginary perturbations. If Ω is real (imaginary), also Y is real (imaginary) and this means that

$$b^{i}_{a} = b^{a}_{i} \quad \text{for } \Omega \text{ real} \qquad (4.20a)$$

$$b^{i}_{a} = -b^{a}_{i} \quad \text{for } \Omega \text{ imaginary} \qquad (4.20b)$$

One can then write (4.18) as

$$(\varepsilon_j - \varepsilon_b) b^{j}_{b} = \omega^{j}_{b} + (\bar{g}^{ja}_{bi} - \bar{g}^{ji}_{ba}) b^{i}_{a} \quad \text{for } \Omega \text{ real} \qquad (4.21a)$$

$$(\varepsilon_j - \varepsilon_b) b^{j}_{b} = \omega^{j}_{b} + (\bar{g}^{ja}_{bi} + \bar{g}^{ji}_{ba}) b^{i}_{a} \quad \text{for } \Omega \text{ imaginary} \qquad (4.21b)$$

For the third order energy we get from (2.7d) a sum of the expectation values of a triple and a double commutator. The expectation value of the triple commutator vanishes because for an Y_1 of the form

$$Y_1 = b^a_i a^i_a + b^i_a a^a_i \tag{4.22}$$

no fully contracted diagrams for the H given by (3.1) can be constructed. So we are left with the double commutator (see the appendix)

$$\begin{aligned} E_3 &= \tfrac{1}{2}\langle \Psi_0 | [[\Omega, Y_1], Y_1] | \Psi_0 \rangle = \tfrac{1}{2}\{\omega^j_i b^a_j b^i_a - \omega^a_b b^b_i b^i_a - \omega^b_a b^j_b b^a_i + \omega^i_j b^j_a b^a_i\} \\ &= \mathrm{Re}\{\omega^j_i b^a_j b^i_a - \omega^a_b b^b_i b^i_a\} \end{aligned} \tag{4.23}$$

An important feature of closed-shell Hartree–Fock theory is that there is a *one-electron operator* of which the Hartree–Fock function Φ_{HF} is eigenfunction with the eigenvalue E_{HF}, namely

$$H_{HF} = f^p_q a^q_p - \tfrac{1}{2}\bar{g}^{ij}_{ij} \tag{4.24}$$

This will turn out important in later sections where the Hamiltonian H will be decomposed as (with $\tilde{V}$ sometimes called 'fluctuation potential')

$$H = H_{HF} + \tilde{V} \tag{4.25}$$

The relation between H_{HF} and H is particularly transparent if we formulate both in the particle-hole formalism (see the appendix) in which excitation operators carry a tilde

$$H = E_{HF} + f^p_q \tilde{a}^q_p + \tfrac{1}{2} g^{pq}_{rs} \tilde{a}^{rs}_{pq} \tag{4.26a}$$

$$E_{HF} = h^i_i + \tfrac{1}{2}\bar{g}^{ij}_{ij} = f^i_i - \tfrac{1}{2}\bar{g}^{ij}_{ij} \tag{4.26b}$$

$$H_{HF} = f^p_q \tilde{a}^q_p + f^i_i - \tfrac{1}{2}\bar{g}^{ij}_{ij} \tag{4.26c}$$

$$\begin{aligned} \tilde{V} &= E_{HF} - f^i_i + \tfrac{1}{2}\bar{g}^{ij}_{ij} + \tfrac{1}{2}\bar{g}^{pq}_{rs}\tilde{a}^{rs}_{pq} = \tfrac{1}{2}\bar{g}^{pq}_{rs}\tilde{a}^{rs}_{pq} \\ &= \tfrac{1}{2}\bar{g}^{pq}_{rs} a^{rs}_{pq} + (f^p_q - h^p_q) a^p_q + \tfrac{1}{2}\bar{g}^{ij}_{ji} \end{aligned} \tag{4.26d}$$

If one considers that H_{HF} treats the average electron interaction, one can argue that one may replace the correct μ-dependent Hamiltonian (3.9) by the Hamiltonian

$$H_{HF} + \mu\Omega \tag{4.27}$$

in order to describe the property Ω with an average electron interaction taken care of. This philosophy leads to uncoupled Hartree–Fock theory (UCHF). Note that the abbreviation UHF cannot be used since this is usually interpreted as 'unrestricted Hartree–Fock'. The drawback of this approach is that stationarity of $\langle \Psi_0 | H_{HF} + \mu\Omega | \Psi_0 \rangle$ as function of μ does not imply stationarity of $\langle \Psi_0 | H + \mu\Omega | \Psi_0 \rangle$. This is only implied in the limit $\mu \to 0$. One hence ignores that the presence of $\mu\Omega$ has an effect on the electron interaction, even the averaged one.

As to an early discussion of CHF and UCHF see [9].

5. MC-SCF and coupled MC-SCF

In the MC-SCF theory we have to consider two variational groups $\mathcal{G}_1$ and $\mathcal{G}_2$, both of which are subgroups of the full-CI group $\mathcal{G}^{(n)}$. The Lie algebra $\mathcal{L}_1$ is

$\mathscr{L}_c^{(1)}$ with elements (3.3a) as in Hartree–Fock theory while the Lie algebra $\mathscr{L}_2$ of $\mathscr{G}_2$ is that of the shift operators

$$P_{\mu\nu} = |\phi_\mu\rangle\langle\phi_\nu| \tag{5.1}$$

corresponding to CI in the active space. In the case of CAS-SCF – which we mainly consider – an equivalent basis of this Lie algebra is

$$\{a_x^y, a_{xy}^{zu}, \ldots\}P_\mu \tag{5.2}$$

i.e. a basis of the same type as $\mathscr{L}_c^{(n)}$ but defined in terms of the subset of active spin orbitals with occupation number between 0 and 1, labelled $x, y, z, \ldots$. This basis is not given at the outset, but only in terms of the stationary condition for the unperturbed wave function. We follow the formalism of I. Sect. 8, where it is important to define the order in which operators of the two groups are applied. For mixed commutators one must always first commute H with a element of $\mathscr{L}_1 = \mathscr{L}_c^{(1)}$ then of $\mathscr{L}_2$ with basis (5.1) or (5.2). The MC-SCF wave function Ψ_0 must satisfy the stationarity condition (2.3) for operators X both in $\mathscr{L}_1$ and $\mathscr{L}_2$. For an iterative construction of Ψ_0 (see I. Sect. 8) we start from a trial function Φ and formulate Ψ_0 as

$$\Psi_0 = e^\sigma \Phi; \quad \sigma = \sum_k c_k^{(1)} X_k^{(1)} + \sum_k c_k^{(2)} X_k^{(2)}; \quad X_k^{(1)} \in \mathscr{L}_1, X_k^{(2)} \in \mathscr{L}_2 \tag{5.3}$$

The linear system to be solved in the first Newton–Raphson cycle for the determination of the $c_k^{(j)}$ is then

$$\sum_l (H_{kl}^{(11)} c_l^{(1)} + H_{kl}^{(12)} c_l^{(2)}) + H_k^{(1)} = 0 \tag{5.4a}$$

$$\sum_l (H_{kl}^{(21)} c_l^{(1)} + H_{kl}^{(22)} c_l^{(2)}) + H_k^{(2)} = 0 \tag{5.4b}$$

with

$$H_k^{(j)} = \langle\Phi|[H, X_k^{(j)}]|\Phi\rangle; \quad X_k^{(j)} \in \mathscr{L}_j; j = 1, 2 \tag{5.5}$$

$$H_{kl}^{(jj)} = \langle\Phi|[[H, X_k^{(j)}], X_l^{(j)}]|\Phi\rangle; \quad X_k^{(1)}, X_l^{(1)} \in \mathscr{L}_1; \quad X_k^{(2)}, X_l^{(2)} \in \mathscr{L}_2 \tag{5.6a}$$

$$H_{kl}^{(12)} = \langle\Phi|[[H, X_k^{(1)}], X_l^{(2)}]|\Phi\rangle = H_{kl}^{(21)}; \quad X_k^{(1)} \in \mathscr{L}_1, X_k^{(2)} \in \mathscr{L}_2 \tag{5.6b}$$

Note that in mixed double commutators, H must first be commuted with an element of $\mathscr{L}_1$.

In terms of Φ and the coefficients $c_k^{(i)}$ obtained in the first iteration cycle from (5.4) one constructs by the recipe (5.3) a better approximation Φ_1, inserts this for Φ in (5.5, 6) to get new coefficients from (5.4) and proceeds until self consistency. If Φ was sufficiently close to the desired Ψ_0, this scheme converges quadratically.

Having so solved the unperturbed problem one can switch to coupled MC-SCF, i.e. solve (2.6b) for Y_1, with Y_1 expanded in analogy to (5.3) as

$$Y_1 = \sum_k b_k^{(1)} X_k^{(1)} + \sum_k b_k^{(2)} X_k^{(2)}; \quad X_k^{(1)} \in \mathscr{L}_1; X_k^{(2)} \in \mathscr{L}_2 \tag{5.7}$$

One gets a system of equations very much like (5.4), just with $H_k^{(1)}$ and $H_k^{(2)}$ replaced by $\Omega_k^{(1)}$ and $\Omega_k^{(2)}$, and Φ in (5.3) replaced by Ψ_0. We can write this in block form as

$$\begin{pmatrix} \underline{\underline{H}}^{(11)} & \underline{\underline{H}}^{(12)} \\ \underline{\underline{H}}^{(12)\dagger} & \underline{\underline{H}}^{(22)} \end{pmatrix} \begin{pmatrix} \vec{b}^{(1)} \\ \vec{b}^{(2)} \end{pmatrix} + \begin{pmatrix} \vec{\Omega}^{(1)} \\ \vec{\Omega}^{(2)} \end{pmatrix} = 0 \tag{5.8}$$

with

$$\Omega_k^{(j)} = \langle\Psi_0|[\Omega, X_k^{(j)}]|\Psi_0\rangle; \quad X_k^{(j)} \in \mathscr{L}_j \tag{5.9}$$

For the second order energy we obtain

$$E^{(2)} = \mathrm{Re}\{\vec{b}^{(1)}\vec{\Omega}^{(1)} + \vec{b}^{(2)}\vec{\Omega}^{(2)}\} \tag{5.10}$$

It consists of an orbital-rotation- and a CI component.

The formalism just outlined is very compact. It becomes more lengthy if one inserts the explicit expressions for the matrix elements (5.5) and (5.6). For the matrix elements (5.6a) of the Hessean for $j = 1$ we can use the general expression (4.12a), of course, without the simplifications that only hold for a single Slater-determinant reference function.

For the matrix elements (5.6a) with $j = 2$ one gets

$$\begin{aligned} \langle \Phi | [[H, P_{\mu\nu}], P_{\varrho\sigma}] | \Phi \rangle &= \delta_{\nu\varrho} \langle \Phi | H | \phi_\mu \rangle \langle \phi_\sigma | \Phi \rangle - \langle \Phi | \phi_\mu \rangle \langle \phi_\nu | H | \phi_\varrho \rangle \langle \phi_\sigma | \Phi \rangle \\ &\quad - \langle \Phi | \phi_\varrho \rangle \langle \phi_\sigma | H | \phi_\mu \rangle \langle \phi_\nu | \Phi \rangle + \delta_{\sigma\mu} \langle \Phi | \phi_\varrho \rangle \langle \phi_\nu | H | \Phi \rangle \end{aligned} \tag{5.11}$$

A considerable simplification arises (see I. Appendix B) if the ϕ_μ are so chosen that $\langle \phi_\mu | H | \phi_\nu \rangle$ is diagonal and that Φ is identical with one of the ϕ_μ. The matrix elements of type (5.6b)

$$\langle \Phi | [[H, a^p_q], P_{\varrho\sigma}] | \Phi \rangle = \langle \Phi | [H, a^p_q] | \phi_\varrho \rangle \langle \phi_\sigma | \Phi \rangle - \langle \Phi | \phi_\varrho \rangle \langle \Phi_\sigma | [H, a^p_q] | \Phi \rangle \tag{5.12}$$

can be evaluated by means of (4.3) just with the density matrix elements γ^t_q, Γ^{tp}_{rs} replaced by the corresponding transition density matrix elements.

Most MC-SCF methods in current use [12, 13] follow essentially the general scheme outlined here, while there is relatively little discussion on coupled MC-SCF in the literature [14].

6. Iterative construction of a full-CI wave function

In Sects. 6 and 10 we study a Hamiltonian H in the absence of an external perturbation. We use the notation that Φ is an uncorrelated wave function and Ψ the full-CI wave function or an approximation to it. We call the perturbation operator σ rather than Y, because Y will be reserved for the response to an external perturbation $\mu\Omega$.

We know that a full-CI wave function satisfies the Brillouin condition

$$\langle \Psi | [H, Z] | \Psi \rangle = 0; \quad \forall Z \in \mathscr{L}^{(n)}_c \tag{6.1}$$

We remember that the Hartree–Fock wave function Φ_{HF} satisfies

$$\langle \Phi_{HF} | [H, X] | \Phi_{HF} \rangle = 0; \quad \forall Z \in \mathscr{L}^{(1)}_c \tag{6.2}$$

and that it can be constructed iteratively, starting from a trial function and choosing the ansatz (2.9) with $\sigma \in \mathscr{L}^{(1)}_r$. This suggests to proceed for (6.1) in a similar way. We choose a trial function Φ and construct a new function [7, 8, 15, 16]

$$\Psi = e^\sigma \Phi; \quad \sigma = \sum_k c_k Z_k \tag{6.3}$$

and determine the coefficients c_k such that the expectation value $\langle \Psi | H | \Psi \rangle$ is made stationary, i.e. that (6.1) is satisfied.

$$\begin{aligned} 0 &= \langle \Phi | e^{-\sigma} [H, Z_l] e^\sigma | \Phi \rangle \\ &= \langle \Phi | [H, Z_l] + [[H, Z_l], \sigma] + \tfrac{1}{2} [[[H, Z_l], \sigma], \sigma] + \cdots | \Phi \rangle \end{aligned} \tag{6.4}$$

Insertion of the expansion (6.3) into (6.4) leads to a non-linear system of equations for the c_k. If the trial function Φ was sufficiently well chosen, $\|\sigma\|$ will be small and the $|c_k|$ will be small and it will be a good approximation to linearize the system (6.4), i.e. to solve [7, 8, 15, 16]

$$\langle \Phi | [H, Z_k] | \Phi \rangle + \sum_l c_l \langle \Phi | [[H, Z_k], Z_l] | \Phi \rangle = 0 \tag{6.5}$$

Using these $c_l = c_l^{(1)}$ we construct via (6.3) a new trial function $\Phi^{(1)}$ and we solve the system (6.5), with Φ replaced by $\Phi^{(1)}$. If the scheme converges, it does so quadratically and finally approaches the full-CI function Ψ.

There is a significant difference with respect to quadratically convergent SCF or MC-SCF (see Sect. 5), that can be formulated in the same way. In SCF Φ is a single Slater determinant, and $e^{\sigma}\Phi$ will be a single Slater determinant again (in MC-SCF the number of determinants in the expansion will remain unchanged), while now starting from a single Slater determinant Φ, the new function $\Phi^{(1)}$ will be a complicated linear combination of determinants. Before we develop a general strategy for the iterative construction of the full-CI wave function we consider some questions of general interest.

The first question is the choice of the reference function Φ. One possibility (a) is to take Φ as the appropriate eigenfunction of the bare nuclear Hamiltonian (this is, by the way, often done in Hartree–Fock theory as *iteration start*). The other one (b) consists in choosing the Hartree–Fock wave function Φ_{HF} as reference.

The difference between the two choices is obvious. In case (a) one never discriminates between $X \in \mathscr{L}^{(1)}$ and $Z \notin \mathscr{L}^{(1)}$, one treats all $Z \in \mathscr{L}^{(n)}$ on the same level, in case (b) one first takes only $X \in \mathscr{L}^{(1)}$ until (6.2) is satisfied and one worries only then about the $Z \notin \mathscr{L}^{(1)}$. Since – as we shall see – the treatment of $Z \notin \mathscr{L}^{(1)}$ is usually possible only approximatively, choice (b) has the advantage that at least single excitations are – to some extent – treated exactly, but at the same time the drawback that single and higher excitations are taken care of in an unbalanced way.

There is also a third possibility, namely to choose the 'best-overlap' or Brueckner determinant as reference. This does not make much sense as starting point for perturbation theory, but has advantages in the nonperturbative context. For details see Sect. 10.

Another choice that one has to make, is whether one wants to treat the electron interaction perturbatively or in a nonperturbative way. Perturbation theory has the advantage that the ordering in powers of the perturbation parameter simplifies the resulting equations to the extent that one does not have to worry about additional simplifications, as is necessary in the nonperturbative approach. On the other hand, the perturbation expansion may not converge and low-order perturbation theory may lead to poor results.

Since perturbation theory is somewhat simpler we discuss this first (Sects. 7 and 8), before we take up the idea of a nonperturbative approach (Sect. 9).

7. Unitary formulation of many-body perturbation theory (MBPT) and Møller–Plesset perturbation theory (MP)

It is not the aim of this section to give a new derivation of many-body perturbation theory (MBPT), but rather to demonstrate that even MBPT (in a finite one-electron basis) fits into the general framework of stationary perturbation theory and that its derivation in this context is relatively simple. Most important for us is the relation to the nonperturbative approach in the same context (Sect. 8). The present derivation of MBPT shares with that from coupled cluster theory [17, 18] that the Lie algebraic structure makes the connectedness of the expansion (keyword linked-cluster theorem) immediately obvious and that diagrams are not needed for a general presentation of the theory, although they are useful for the illustration of particular terms. As to a recent historical-critical review on MBPT the reader is referred to [19].

We want to treat the λ-dependent Hamiltonian

$$H_\lambda = H_0 + \lambda V \tag{7.1}$$

$$H_0 = h^p_q a^q_p; \qquad V = \tfrac{1}{2} g^{pq}_{rs} a^{rs}_{pq} \tag{7.2}$$

by means of stationary perturbation theory. The perturbation parameter λ is only formal and will be set equal to 1 at the end, but it serves to define orders in perturbation theory.

The variational group is obviously that of full CI (see Sect. 3) with the Lie algebra $\mathscr{L}^{(n)}_c$, if we consider n-particle states. The *unperturbed* Brillouin condition (2.6a) is trivially fulfilled for Z operators of higher particle excitation rank than 1. We need only consider the nontrivial conditions

$$\langle \Psi_0 | [H_0, a^q_p] | \Psi_0 \rangle = h^q_s \gamma^s_p - h^r_p \gamma^q_r = 0 \tag{7.3}$$

Obviously (7.3) implies that $\underline{\underline{h}}$ and $\underline{\underline{\gamma}}$ commute, i.e. have common eigenfunctions. We hence first diagonalize $\underline{\underline{h}}$, i.e., the matrix representation of the one-electron Hamiltonian H_0 in the given basis $\{\chi_p\}$

$$h^q_p c^p_r = e_r c^q_r; \qquad \psi_r = c^p_r \chi_p \tag{7.4}$$

We then construct the Slater determinant Ψ_0 from the ψ_r corresponding to the lowest one-electron energies e_r in the sense of the aufbau principle and express the Hamiltonian in terms of the ψ_r and the corresponding excitation operators such that

$$H_0 = e_p a^p_p \tag{7.5}$$

For the 0th and 1st order energy (in orders of λ) we get (summation over i)

$$E_0 = e_i \tag{7.6a}$$

$$E_1 = \langle \Psi_0 | V | \Psi_0 \rangle = \tfrac{1}{2} \bar{g}^{ij}_{ij} \tag{7.6b}$$

Next we have to satisfy the condition (2.6b) for σ_1 (as to the choice of the symbol σ rather than Y see the remark on the beginning of Sect. 6), or the linear system of equations (2.11b) for σ_1 expanded in the form (2.10), with H_{kl} given by (2.12b) and V_k by (2.13). Since V is a two electron operator the V_k vanish for X_k an excitation operator of excitation rank other than 1 or 2.

It is convenient to introduce the particle-hole formalism [20] (see the appendix) with

$$H_0 = e_p \tilde{a}^p_p + e_i \tag{7.7a}$$

$$V = g^{pq}_{rs} \tilde{a}^{rs}_{pq} + \bar{g}^{ip}_{iq} \tilde{a}^q_p - \tfrac{1}{2} g^{ij}_{ji} \tag{7.7b}$$

Then

$$V_q^p = \tfrac{1}{2} g_{tu}^{rs} \langle \Psi_0 | [\tilde{a}_{rs}^{tu}, \tilde{a}_q^p] | \Psi_0 \rangle + \bar{g}_{is}^{ir} \langle \Psi_0 | [\tilde{a}_r^s, \tilde{a}_q^p] | \Psi_0 \rangle = \bar{g}_{iq}^{ip} (n_p - n_q) \tag{7.8a}$$

$$\begin{aligned} V_{rs}^{pq} &= \tfrac{1}{2} g_{vw}^{tu} \langle \Psi_0 | [\tilde{a}_{tu}^{vw}, \tilde{a}_{rs}^{pq}] | \Psi_0 \rangle + \bar{g}_{iu}^{it} \langle \Psi_0 | [\tilde{a}_t^u, \tilde{a}_{rs}^{pq}] | \Psi_0 \rangle \\ &= \{ g_{rs}^{pq} (\delta_t^p \delta_u^q \delta_r^v \delta_s^w + \delta_u^p \delta_r^q \delta_u^v \delta_t^w) - g_{sr}^{pq} (\delta_t^p \delta_u^q \delta_u^v \delta_t^w + \delta_u^p \delta_r^q \delta_r^v \delta_s^w) \} \\ &\quad \times \{ n_r n_s (1 - n_p)(1 - n_q) - n_p n_q (1 - n_p)(1 - n_s) \} \end{aligned} \tag{7.8b}$$

The only nonvanishing matrix elements of type (7,8) are

$$V_i^a = \bar{g}_{ij}^{aj}; \qquad V_a^i = -\bar{g}_{aj}^{ij} \tag{7.9a}$$

$$V_{ij}^{ab} = \bar{g}_{ij}^{ab}; \qquad V_{ab}^{ij} = -\bar{g}_{ab}^{ij} \tag{7.9b}$$

To construct the matrix elements of the Hessean in terms of the corresponding basic operators we note that

$$[H_0, \tilde{a}_p^q] = (e_q - e_p) a_p^q; \qquad [H_0, \tilde{a}_{pq}^{rs}] = (e_r + e_s - e_p - e_q) \tilde{a}_{pq}^{rs} \tag{7.10}$$

$$\langle \Psi_0 | [\tilde{a}_q^p, [H_0, \tilde{a}_s^r]] | \Psi_0 \rangle = (e_r - e_s)(n_s - n_r) \delta_q^r \delta_s^p \tag{7.11a}$$

$$\begin{aligned} \langle \Psi_0 | [\tilde{a}_{rs}^{pq}, [H_0, \tilde{a}_{vw}^{tu}]] | \Psi_0 \rangle &= (e_t + e_u - e_v - e_w)(\delta_r^t \delta_s^u - \delta_s^t \delta_r^q)(\delta_v^p \delta_w^q - \delta_w^p \delta_v^p) \\ &\quad \times \{ n_v n_w (1 - n_t)(1 - n_u) - n_t n_u (1 - n_v)(1 - n_w) \} \end{aligned} \tag{7.11b}$$

Couplings between single and double excitations vanish.

The only non-vanishing matrix elements of the Hessean are hence

$$\langle \Psi_0 | [\tilde{a}_a^j, [H_0, \tilde{a}_j^a]] | \Psi_0 \rangle = e_a - e_j \tag{7.12a}$$

$$\langle \Psi_0 | [\tilde{a}_j^a, [H_0, \tilde{a}_a^j]] | \Psi_0 \rangle = e_a - e_j \tag{7.12b}$$

$$\langle \Psi_0 | [\tilde{a}_{ij}^{ab}, [H_0, \tilde{a}_{ab}^{ij}]] | \Psi_0 \rangle = e_a + e_b - e_i - e_j \tag{7.12c}$$

$$\langle \Psi_0 | [\tilde{a}_{ab}^{ij}, [H_0, \tilde{a}_{ij}^{ab}]] | \Psi_0 \rangle = e_a + e_b - e_i - e_j \tag{7.12d}$$

Since the Hessean is diagonal, the solution of the linear system (2.11b) is trivial

$$b_i^a = (e_a - e_i)^{-1} \bar{g}_{ij}^{aj} = -(b_a^i)^* \tag{7.13a}$$

$$b_{ij}^{ab} = (e_a + e_b - e_i - e_j)^{-1} \bar{g}_{ij}^{ab} = -(b_{ab}^{ij})^* \tag{7.13b}$$

In the expansion of σ one must, of course, take care that

$$\tilde{a}_{ij}^{ab} = -\tilde{a}_{ji}^{ab} = -\tilde{a}_{ij}^{ba} = \tilde{a}_{ji}^{ba} \tag{7.13c}$$

i.e. that a non-redundant basis is given for $i < j$, $a < b$. For the 2nd order energy one gets

$$E_2 = (e_i - e_a)^{-1} |\bar{g}_{ij}^{aj}|^2 + \sum_{i<j} \sum_{a<b} (e_i + e_j - e_a - e_b)^{-1} |\bar{g}_{ab}^{ij}|^2 \tag{7.14}$$

Higher orders of σ and E are obtained in a straightforward way. The results are in agreement with what one gets from the traditional derivation of MBPT, which is usually not done in a stationary context, but under the assumption that the unperturbed problem is solved exactly. The same diagrams can be used to represent the results. However, the use of diagrams is neither necessary to formulate nor to understand the theory. The Lie-algebraic structure of the theory guarantees that only connected diagrams contribute.

To second order in σ, i.e. for σ_2 triple excitations appear, which are needed for the evaluation of E_4 and E_5.

The main message of the first part of this section is that MBPT or more precisely the perturbation expansion of full CI can be derived in the framework of stationary perturbation theory and satisfies hence all theorems that hold in this framework.

We now want to switch from classical MBPT to MP. We no longer treat the entire electron interaction by perturbation theory, but we take a part of it to infinite order in λ, namely that part which can be described by a one-particle transformation, for which a treatment to infinite order is rather simple. So rather than to require that the Brillouin condition

$$\langle \Psi(\lambda) | [H_0 + \lambda V, Z] | \Psi(\lambda) \rangle = 0; \quad \forall Z \in \mathscr{L}_c^{(n)} \tag{7.15}$$

holds order by order in λ, we now try first to satisfy

$$\langle \Psi_0 | [H_0 + V, X] | \Psi_0 \rangle = 0; \quad \forall X \in \mathscr{L}_c^{(1)} \tag{7.16}$$

in a non-expanded form, before we satisfy the counterpart of (7.15) for the $Z \in \mathscr{L}_c^{(n)}$ that are not elements of $\mathscr{L}_c^{(1)}$. Equation (7.16) defines, of course, Hartree–Fock theory, i.e. we can use the results of Sect. 4.

The distinction between MBPT and MP made here is not shared by all authors in the field, but if there is a point in using the term MP as an alternative to MBPT, this only makes sense if MP is regarded as based on Hartree–Fock as 0th order reference and MBPT on the bare nuclear Hamiltonian [19]. There is a need to distinguish these two situations, while there is no need at all to discriminate – as is sometimes done – between a formulation in terms of diagrams (MBPT) or without diagrams (MP), since the theory is the same whether or not one uses diagrams.

In order to treat that part of the electron interaction, not taken care of in (7.16) – which is usually referred to as electron correlation –, we can use an important property of closed-shell Hartree–Fock theory, namely that there is a one-electron Hamiltonian H_{HF} defined by (4.26c), that has the Hartree–Fock wave function as eigenfunction and the Hartree–Fock energy E_{HF} as eigenvalue.

Let Φ_{HF} be the Hartree–Fock wave function (a single Slater determinant). The 0th order energy E_0 is, of course, equal to the Hartree–Fock energy E_{HF} given by (4.26b). The 1st order energy vanishes

$$E_1 = \langle \Phi_{HF} | \tilde{V} | \Phi_{HF} \rangle = 0 \tag{7.17}$$

For the construction of E_2 and E_3 we need σ_1 that is constructed in an analogous way as in the beginning of this section.

Since $\tilde{V}$, in the particle-hole formulation, is a pure two-electron operator, only two electron operators contribute to σ_1. We get

$$Y_1 = \sum_{i<j} \sum_{a<b} (b_{ij}^{ab} \tilde{a}_{ab}^{ij} + b_{ab}^{ij} \tilde{a}_{ij}^{ab}) \tag{7.18}$$

$$b_{ij}^{ab} = (\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j)^{-1} \bar{g}_{ij}^{ab} = -(b_{ab}^{ji})^* \tag{7.19}$$

$$E_2 = \tfrac{1}{4}(\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b)^{-1} |\bar{g}_{ab}^{ij}|^2 \tag{7.20}$$

For the evaluation of E_3 we start from (2.7d). The expectation value of the triple commutator vanishes, such that

$$E_3 = \tfrac{1}{2} \langle \Phi | [[V, Y_1], Y_1] | \Phi \rangle = \tfrac{1}{8} b_{cd}^{ij} \bar{g}_{ab}^{cd} b_{ij}^{ab} + \tfrac{1}{8} b_{ab}^{kl} \bar{g}_{kl}^{ij} b_{ij}^{ab} - b_{cb}^{ik} \bar{g}_{ak}^{cj} b_{ij}^{ab} \tag{7.21}$$

Again the final results are formally the same as in traditional theory, which indicates that MP can be justified in the framework of stationary perturbation theory.

As in the traditional formulation of MP, both single and triple excitations contribute to σ_2 as well as to E_4 and E_5.

8. Stationary unitary coupled-cluster theory

We now want to take up the ideas of Sect. 6 to construct the full CI wave function iteratively (rather than by perturbation theory as in Sect. 7).

It is convenient to perform first a Hartree–Fock calculation i.e. to construct a wave function which satisfies (6.2). We then search for a wave function Ψ given by (6.3) and we approximate the solution of (6.4) by that of (6.5).

Let us rewrite the linear system (6.5) as

$$\sum_l H_{kl} c_l = -D_k \tag{8.1}$$

Since this is an iteration start there is no point in solving (8.1) exactly. It is recommended to consider only those lines in (8.1) where $D_k \neq 0$. Then, of course, one must also limit the expansion (6.3) to these k such that the Hessean $\underline{\underline{H}}$ remains a quadratic (and regular) matrix. Obviously $D_k = 0$ if Z_k is an excitation operator of particle rank higher than two (since the Hamiltonian does not contain higher than two particle operators). But $D_k = 0$ also holds for one-particle operators due to the Brillouin condition. The Z_k to be considered are hence two particle excitation operators of the form

$$a_{ij}^{ab}, a_{ab}^{ij} \tag{8.2}$$

We get the corresponding coefficients c_{ab}^{ij} and c_{ij}^{ab} from the linear system of equations (8.1) or (6.5) that is well known as the CEPA-0 system. It amounts to making the truncated energy expression

$$E = \langle \Phi | H + [H, \sigma] + \tfrac{1}{2}[[H, \sigma], \sigma] | \Phi \rangle \tag{8.3}$$

stationary with respect to variation of σ (expanded in operators of type 8.2) [11]. When stationarity is achieved, an alternative expression equivalent to (8.3) is

$$E = \langle \Phi | H | \Phi \rangle + \tfrac{1}{2} \langle \Phi | [H, \sigma] | \Phi \rangle \tag{8.4}$$

CEPA-0 is known under many names (for details see [21]), the oldest is probably L-CPMET [22] (for linearized coupled-pair-many-electron theory). That CEPA-0 is the first non-trivial approximation to a unitary coupled-cluster theory has probably been first shown in [15].

CEPA-0 is often a good approximation (at least in cases where Φ is not near-degenerate with some other state) and usually much better than 2nd order perturbation theory. In fact (8.3) can also be derived via a partial summation of certain terms of MP to infinite order.

To discuss the accuracy obtained in this first iteration and in subsequent iterations we define the norm $S = \|\sigma\|$ of the correlation operator σ. The dominant c_k obtained as solution of (8.1) are hence of $O(S)$. In the next iteration cycle contributions of higher order in S will be obtained.

The σ constructed in the first iteration cycle will be called $\sigma^{(1)}$. In a straightforward quadratically convergent iteration scheme the next step would be to replace Φ in (6.4) by

$$\Phi^{(1)} = e^{\sigma^{(1)}} \Phi; \quad \sigma^{(1)} = \sum_k c_k^{(1)} Z_k \tag{8.5}$$

It is perfectly in the spirit of a Newton–Raphson scheme to linearize this system to (6.5) – of course with Φ replaced by $\Phi^{(1)}$ and c_l by $c_l^{(2)}$. However, the Hausdorff expansion in powers of $\sigma^{(1)}$ should not be truncated, which causes problems since this expansion is infinite. In Hartree–Fock theory one is much better off because the counterpart of $\Phi^{(1)}$ can easily be obtained from Φ by a simple orbital rotation.

Fortunately there is an argument which allows us to truncate the Hausdorff expansion in powers of $\sigma^{(1)}$ in (6.4) without leaving the context of a quadratically convergent scheme. The basis of this argument is the observation that expectation values of odd-order commutators of H with double excitation operators Z_D of type (8.1)

$$\langle\Phi|[H, Z_D]|\Phi\rangle; \qquad \langle\Phi|[[[H, Z_D], Z'_D], Z''_D]|\Phi\rangle \tag{8.6a}$$

are usually significantly smaller than even-order commutators

$$\langle\Phi|[[H, Z_D], Z'_D]|\Phi\rangle; \qquad \langle\Phi|[[[[H, Z'_D], Z'_D], Z''_D], Z'''_D]|\Phi\rangle \tag{8.6b}$$

In fact for the one-electron part H_0 of H, expectation values of type (8.6a) vanish. Of the two-electron part only exchange-type integrals contribute to these matrix elements. This argument is valid as long as no excited configuration is near-degenerate with the leading configuration, because then expectation values (8.6b) which involve energy differences, can become small. This case has to be avoided (we come back to this later). On the same argument also the validity of perturbation theory is based.

The norm S of the correlation operator σ is dominated by the contribution $\sigma_D^{(1)}$ of two-particle operators to σ obtained in the first iteration cycle (we now use the subscript D for double-excitation operators).

$$S = \|\sigma\| \approx \|\sigma_D^{(1)}\| \tag{8.7}$$

Both the second and the third terms in (8.3) are of $O(S^2)$ with respect to the reference energy $\langle\Phi|H|\Phi\rangle$, which becomes obvious if we remember that (8.3) is equivalent to (8.4) for the σ constructed from the CEPA-0 equations.

Consider now the matrix elements required in the 2nd iteration cycle

$$D_k^{(2)} = \langle\Phi|[H, Z_k] + [[H, Z_k]\,\sigma^{(1)}] + \tfrac{1}{2}[[[H, Z_k], \sigma^{(1)}], \sigma^{(1)}] + \cdots|\Phi\rangle \tag{8.8a}$$

$$\begin{aligned} H_{kl}^{(2)} = {} & \tfrac{1}{2}\langle\Phi|[[H, Z_k], Z_l] + [[[H, Z_k], Z_l], \sigma^{(1)}] + \cdots|\Phi\rangle \\ & + k \text{ and } l \text{ interchanged} \end{aligned} \tag{8.8b}$$

Let us first consider the case that Z_k and Z_l are two-particle operators. As a result of stationarity to 1st order (the CEPA-0 equations) the sum of the first two terms in (8.8a) vanishes, and the leading term is the third one. This means that the update $\sigma_D^{(2)}$ of σ_D is smaller by $O(S^2)$ than the leading term $\sigma_D^{(1)}$ of σ_D. To get $\sigma_D^{(2)}$ correct to $O(S^3)$ it is sufficient to take the leading term and the next term (which is of the same order) of (8.8a) i.e.

$$D_k^{(2)} = \langle\Phi|\tfrac{1}{2}[[H, Z_k], \sigma^{(1)}], \sigma^{(1)}] + \tfrac{1}{6}[[[H, Z_k], \sigma^{(1)}], \sigma^{(1)}], \sigma^{(1)}]|\Phi\rangle \tag{8.9}$$

and only the first term of (8.8b).

This second iteration cycle – so far limited to two particle operators $\sigma \approx \sigma_D$ – corresponds to making the Hausdorff expansion to 4th order in σ stationary with respect to variation of a $\sigma = \sigma_D$.

Of course, in the 2nd iteration cycle also 1-particle (σ_S), 3-particle (σ_T) and even 4-particle operators (σ_Q) contribute, since they have non-vanishing $D_k^{(2)}$.

The importance of σ_S and σ_T is comparable to that of the update of σ_D as far as the energy is concerned, while for σ itself and hence for properties σ_S is even more important.

If we consider (8.8a) for a 1-particle or 3-particle operator, we see that the 1st term vanishes, but not the 2nd term and this means D_k is of $O(S^2)$. The leading contributions to σ_S and σ_T (which start with $\sigma_S^{(2)}$ andf $\sigma_T^{(2)}$) are of $O(S^2)$, i.e. one order of S smaller than σ_D.

To get the σ_S and σ_T to the accuracy consistent with the 2nd iteration cycle it is sufficient to take the leading (i.e. the 2nd) term in (8.8a) and the 1st term in (8.8b). Since in the 2nd iteration cycle one gets σ_D correct to $O(S^3)$ one might wonder whether one should include more terms in (8.8) to get σ_S and σ_T to the same accuracy. This is not necessary, since the argument related to (8.6) has to be somewhat modified. If one single (or triple) excitation is involved expectation values like

$$\langle\Phi|[[H, Z_S], Z_D]|\Phi\rangle; \qquad \langle\Phi|[[[[H, Z_S], Z_D], Z_D], Z_D|\Phi\rangle \tag{8.10a}$$

are usually smaller by one order of S than

$$\langle\Phi|[[[H, Z_S], Z_D], Z_D]|\Phi\rangle \quad \text{etc.} \tag{8.10b}$$

This means that inclusion of the next term in (8.8a) would improve σ_S or σ_T by something of $O(S^4)$ which is beyond the accuracy achieved for σ_D in the 2nd cycle.

As far as the contributions to the energy are concerned, the leading terms $\sigma_S^{(1)}$ and $\sigma_T^{(1)}$ of σ_S and σ_T have an effect of the same order as the update $\sigma_D^{(2)}$ of σ_D, namely of $O(S^4)$, since σ_S and σ_T don't interact directly with Φ.

For σ_Q in the 2nd iteration cycle the first two terms in (8.8a) vanish, hence the leading contribution is of $O(S^3)$ like $\sigma_D^{(2)}$, so it should be included if we want σ to this accuracy. Of course, the contribution to the energy is only of $O(S^5)$, which justifies to neglect σ_Q in the 2nd iteration cycle.

To have a quadratically convergent scheme it is sufficient in the first iteration cycle to truncate the Hausdorff expansion of the energy at terms quadratic in σ and to limit σ to σ_D, in the 2nd cycle one has to include up to terms of 4th order in σ and to include up to triple excitations, i.e. to consider σ_S, σ_D, σ_T, and in the 3rd cycle terms to 8th order in σ etc.

If we are interested in σ (or via σ in properties) rather than in E, it may be recommended to include only σ_S and σ_T (for one-electron properties only σ_S) in the 2nd iteration cycle.

While the 2nd iteration still looks feasible, the 3rd iteration appears hopeless. One must rather try to achieve that the convergence is so fast that stopping after the 2nd or even the 1st iteration leads to sufficiently accurate result. This is the case if S is sufficiently small.

Let e.g. $S \approx 0.1$, then the correlation energy E_{corr} is $\sim 1\%$ of E_0, the error of the 1st iteration (CEPA-0) is $\sim 1\%$ of E_{corr} or $\sim 0.01\%$ of E_0, the error of the second iteration 0.01% of E_{corr} or $\sim 10^{-6}$ E_0. This is certainly satisfactory. However, if $S \approx 0.3$, we get an error of $\sim 10\%$ of E_{corr} in the 1st iteration and $\sim 1\%$ of E_{corr} in the second iteration. This is at the limit of what is tolerable.

In typical cases of only dynamic correlation there are many double excitations with comparable, but small coefficients. Then the present scheme should work. Non-dynamical correlation is characterized by a single or a few double excitations with relatively large coefficients. It is then recommended to deal with these excitations not in an iterative way, but rather in a closed form to infinite order, i.e. by full CI in a limited basis.

A possible recipe is hence as follows. One first performs a CEPA-0 calculation and determines the natural orbitals of the CEPA-0 wave function. One checks whether the occupation numbers of all virtual orbitals are sufficiently small. In this case one either stops or goes to the 2nd iteration. If there are weakly occupied NOs with too large occupation number one performs next a CAS-SCF calculation *with these orbitals as active*. Then one performs a new iterative coupled cluster calculation based on MC-SCF as described in Sect. 11.

There may be cases where the first iteration cycle, i.e., CEPA-0 is not accurate enough, but where the 2nd cycle is too tedious.

While the first iteration cycle gives the energy correct to $O(S^2)$, the energy after the 2nd iteration cycle is correct to $O(S^4)$. One might want to have a scheme where the energy is correct to $O(S^3)$. It does not appear possible to achieve this 'intermediate accuracy' in the framework of a Newton–Raphson type iterative scheme towards the construction of the full CI wave function. On the other hand methods in current use like the CC-D, or CC-SD approaches (coupled-cluster with double excitations or with single and double excitations) [22, 23] or their simplifications like CEPA (coupled electron pair approximation) [24] or 'quadratic CI' [25] achieve just this type of accuracy. However, these methods are not related to a stationarity condition as considered in the present paper. That even CC methods satisfy some kind of stationarity condition has recently been analyzed in detail [21, 26].

To switch to conventional coupled cluster theory would, to a large extent, destroy the present philosophy. However, one can try to sum certain important contributions to infinite order in σ_D in such a way that this leads only to a slight modification of the CEPA-0 functional. It has been shown that on these lines the CPF (coupled pair functional) method [27] can be justified [21].

Note that σ_S and σ_T are relatively unimportant for the energy, but very important for properties (at least σ_S), so that a strategy which is good for the energy need not be good for properties (see Sects. 11, 14, 15).

An interesting possibility to formulate a theory correct to $O(S^3)$ – and to other arbitrary orders – is inspired by the unitary coupled-cluster approach UCC(k) of Bartlett and Noga [28]. Here one also makes a truncated Hausdorff expansion of the stationary with respect to variation of σ. However, in the last term one replaces H by H_{HF}. So the functionals for the 2nd, 3rd and fourth order in σ are respectively

$$\langle \Phi | H + [H, \sigma] + \tfrac{1}{2}[[H_{HF}, \sigma], \sigma] | \Phi \rangle \tag{8.11a}$$

$$\langle \Phi | H + [H, \sigma] + \tfrac{1}{2}[[H, \sigma], \sigma] + \tfrac{1}{6}[[[H_{HF}, \sigma], \sigma], \sigma] | \Phi \rangle \tag{8.11b}$$

$$\langle \Phi | H + [H, \sigma] + \tfrac{1}{2}[[H, \sigma], \sigma] + \tfrac{1}{6}[[[H, \sigma], \sigma], \sigma] + \tfrac{1}{24}[[[[H_{HF}, \sigma], \sigma], \sigma], \sigma] | \Phi \rangle \tag{8.11c}$$

This hierarchy of approximations is reminiscent of MP perturbation theory, since one partitions H into H_{HF} and $\tilde{V}$ and truncates the expansion of H_{HF} at the higher order than that of $\tilde{V}$. In fact (8.11a) is just the Hylleraas functional for MP theory to 2nd order. (8.11b) is practically identical with CEPA-0. For σ limited to σ_D the triple commutator vanishes and (8.11b) is CEPA-0; if σ_S and/or σ_T are included there are some very small contributions from the triple commutator.

Obviously (8.11c) is the desired functional, which is correct to $O(S^3)$, i.e. intermediate between the first and the second iteration in our iterative approach.

9. Brueckner orbitals and the Brillouin–Brueckner condition

In Møller–Plesset perturbation theory (MP) the Brillouin condition of Hartree–Fock theory plays a central role (see Sect. 7). It guarantees (a) that the first-order correction E_1 to the energy vanishes and that (b) to the first-order correlation operator σ_1 only double excitations contribute and that hence (c) only double excitations contribute to the second and third order E_2 and E_3.

Although the non-perturbative treatment of Sect. 8 was also based on a Hartree–Fock reference function, no rigorous consequences resulted from this fact, only plausibility arguments, e.g. that double excitations dominate and that in the first iteration cycle it is sufficient to replace σ by σ_D. What we have used is that with a Hartree–Fock reference function the norm S of $\|\sigma\|$ is sufficiently small, such that a Newton–Raphson type iterative scheme is possible, and that in some cases even the first iteration cycle may be sufficient. There is, however, no theorem which tells us that in this respect the Hartree–Fock reference is the best choice.

A well-known alternative to Hartree–Fock is to take that particular one-particle reference function Φ that has maximum overlap with the exact wave function Ψ (for both Φ and Ψ normalized to unity). The formal solution to this problem has been given long ago [29, 30]. We repeat it here in our formalism. We apply an infinitesimal unitary one-particle transformation to Φ

$$\Phi \rightarrow \tilde{\Phi} = e^X \Phi; \quad X \in \mathscr{L}_r^{(1)} \tag{9.1}$$

and require that the overlap integral $\langle \Phi | \Psi \rangle$ is stationary with respect to the transformation (9.1). This is the case if

$$\langle \Phi | X | \Psi \rangle = \langle \Phi | X e^{\sigma} | \Phi \rangle = 0; \quad \forall X \in \mathscr{L}_c^{(1)} \tag{9.2}$$

Since X is a one-particle excitation operator, this means that single excitations $X\Phi$ with respect to Φ are not contained in Ψ, i.e. that the operator $U = e^{\sigma}$ which transforms Φ to Ψ does not contain single excitations.

The single Slater-dominant wave function Φ which satisfies (9.2) is called the 'best-overlap' or 'Brueckner' determinant [29], and the spin-orbitals from which Φ is build up are called Brueckner orbitals.

Before we investigate how to construct the Brueckner orbitals, we have to ask whether we really want to satisfy (9.2). In fact (9.2) involves the full wave operator e^{σ} which is non-separable. We know, however, that *many-body theory should be formulated in terms of separable quantities only*. If one thinks a little bit, one finds out that what one wants is that the correlation operator σ does not contain single excitations, i.e.

$$\langle \Phi | X\sigma | \Phi \rangle = 0, \quad \forall X \in \mathscr{L}_c^{(1)} \tag{9.3}$$

The two conditions (9.2) and (9.3) are not equivalent. Consider e.g. that σ consists of two-particle and three-particle operators. Then σ^2 contains also single excitations as products of 3 particle excitations and two-particle deexcitations.

Now the question arises why it has not yet been observed that (9.3) is a better condition than (9.2). The answer is that the concept of a Brueckner determinant has mainly been used in connection with wave functions $\Psi = e^S \Phi$ in intermediate rather than unitary normalization. In this case S consists only of excitation and no deexcitation operators and (9.3) with σ replaced by S implies (9.2) with σ replaced by S. On the other hand stationarity conditions like (9.3) fit much better into the context of a stationary approach than a non-stationary one as in the case of the intermediate normalization.

If one knows σ, it is relatively easy to construct the Brueckner determinant by the requirement (9.3). Of course, if one knows the exact σ, there is no real need for this and if σ is only approximate, this may not be very helpful. Let σ, e.g., be determined from CEPA-0, such that it contains no single excitations. In this case (9.3) is satisfied for Φ_{HF} and we can't improve this.

Another more constructive approach is as follows. We take the CEPA-0 functional that we want to minimize, but we do not only want to minimize it with respect to variations of σ but also with respect to variations of the reference function. We argue that the σ should be expressed in terms of the orbitals occupied or unoccupied in the improved reference function i.e. that e^{Y} should be applied to $e^{\sigma}\Phi$ rather than to Φ, we hence take the functional

$$\langle\Phi|\, e^{-Y}H\,e^{Y}+[e^{-Y}H\,e^{Y},\sigma]+\tfrac{1}{2}[[e^{-Y}H\,e^{Y},\sigma],\sigma]\,|\Phi\rangle \tag{9.4}$$

For an optimized Φ we have $Y=0$ and the stationarity condition with respect to variation of Y is

$$\langle\Phi|[H,X]+[[H,X],\sigma]+\tfrac{1}{2}[[[H,X],\sigma],\sigma]|\Phi\rangle=0;\quad \forall X\in\mathscr{L}_c^{(1)} \tag{9.5}$$

We call this the Brillouin–Brueckner condition (more precisely: to 2nd order in σ – or at CEPA-0 level). For an iterative solution of (9.5) we linearize the stationarity condition for the functional (9.4) after expansion in powers of Y to

$$\begin{aligned}&\langle\Phi|[H,X]+[[[H,X],Y]+[[H,X]+[H,X],Y],\sigma]\\&+\tfrac{1}{2}[[[H,X]+[[H,X],Y],\sigma],\sigma]|\Phi\rangle=0\end{aligned} \tag{9.6}$$

The nice thing is that whenever one has constructed Y, one can immediately absorb it into a new Φ. Condition for stationarity with respect to variation of σ is

$$\langle\Phi|[H,Z]+\tfrac{1}{2}[[H,\sigma],Z]+\tfrac{1}{2}[[H,Z],\sigma]|\Phi\rangle=0 \tag{9.7}$$

In the very beginning we neither know Y nor σ, so we set them equal to 0, and we only have an arbitrary reference function Φ. For $\sigma=0$, (9.5) becomes just the Brillouin theorem of Hartree–Fock theory, i.e. we determine Φ as Hartree–Fock function. We enter with Φ into (9.7), which is the CEPA-0 system discussed in Sect. 8 from which we get σ. With this σ we enter (9.6) and get a new Y and from this a new Φ and so forth.

Remember that in the discussion of CEPA-0 in Sect. 8 we have argued that single excitations don't contribute to σ in the 1st iteration cycle because for these in (6.5) the first term is zero due to the Brillouin theorem. Now we can argue that due to the Brillouin–Brueckner condition (9.5) the matrix element D_k given by (8.8a) is for $X\in\mathscr{L}_c^{(1)}$ of $O(S^3)$ and σ_S does not contribute in the 2nd iteration cycle. So for Φ determined by (9.4) single excitations in σ can be neglected entirely in the 2nd iteration cycle. They have been absorbed into the reference function.

If we don't truncate the Hausdorff expansion in powers of σ, condition (9.5) implies that σ does not contain single particle excitations, i.e. $\sigma_S=0$.

10. Stationary unitary coupled-cluster theory based on a multiconfiguration reference function

We can apply the same philosophy as in Sects. 6 and 8, but now with respect to a reference function of MC-SCF type. We hence want to satisfy (6.1) with Φ in

(6.2) an MC-SCF function. For the sake of simplicity we consider that it is of CAS-SCF type. We label doubly occupied orbitals as i, j, k, active orbitals as x, y, z and unoccupied orbitals as a, b, c. Our Φ now not only satisfies the Brillouin-condition (5.4a) but also the CI condition (5.4b).

We want to solve (6.4) and again we linearize to (6.5), i.e. we take a CEPA-0 like 1st iteration.

Like in the case of a closed-shell reference state we can argue that the dominating contributions are those for which the first matrix element does not vanish. Due to the Brillouin theorem single excitations described by a_i^a, a_i^x, a_x^a give vanishing matrix elements. The CI condition guarantees that fully internal excitations like a_x^y, a_{xy}^{zu} etc. don't contribute. Also triple and higher excitations involving at least three occupied or three virtual orbitals like a_{ijk}^{abx} or a_{ixy}^{abc} are negligible. What we have to consider are (a) genuine double excitations of the type

$$a_{ij}^{ab},\, a_{ij}^{ax},\, a_{ij}^{xy},\, a_{ix}^{ab},\, a_{xy}^{ab} \tag{10.1}$$

(b) single excitations accompanied by internal excitations

$$a_{ix}^{by},\, a_{ix}^{yz},\, a_{xy}^{az},\, a_{ixy}^{bzu} \text{ etc.} \tag{10.2}$$

(c) double excitations accompanied by internal excitations

$$a_{ijx}^{aby},\, a_{ijx}^{ayz} \text{ etc.} \tag{10.3}$$

The problem with excitations of type (b), (c) is that their particle rank is unlimited. Fortunately there is some redundancy and it is sufficient to consider only conditional excitations (excitations with spectator line) such as

$$a_{ix}^{bx},\, a_{ix}^{yx},\, a_{ijx}^{abx},\, a_{ixy}^{bxy},\, \ldots \tag{10.4}$$

as well as excitations with spin flip (spin-polarization), where $x\alpha$ is excited to $x\beta$. Instead of considering explicitly excitations of type (10.4) one may also consider separate excitations from the Slater determinants ϕ_μ that constitute the MC-SCF reference function Φ.

Alternatively to (10.4) we may write the basis operator as

$$a_i^b\,|\phi_\mu\rangle\langle\phi_\mu|,\quad a_i^y\,|\phi_\mu\rangle\langle\phi_\mu|,\ldots \tag{10.5}$$

i.e. as products of elements of $\mathcal{L}_c^{(n)}$ with basis operator of type (5.1) which are relevant in MC-SCF theory. Of course the choice (10.5) is equivalent to considering excitations separately from the configurations contained in the MC-SCF functions. In order to avoid that the number of basis operators of type (10.5) becomes too large, one should try not make the MC-SCF-expansion unnecessarily long.

The second iteration becomes rather complicated, because now all single and triple excitations have to be included together with such excitations accompanied by fully internal excitations of arbitrary rank.

We remember that in the single-determinant-reference case we have regarded it as worthwhile to treat the single excitations in a privileged way, because they can be taken care of to infinite order, and because the Hartree–Fock start has privileged them anyway. This led to the construction of the Brueckner orbitals. We can now proceed in the same way. However starting from MC-SCF – more precisely from CAS-SCF – we have not only privileged single excitations but also fully internal excitations, i.e. excitations within the active space. So we should continue to do so, i.e. rather than perform an SCF-calculation in the presence of the most important correlation corrections we should now perform an MC-SCF-calculation. In this way we not only get Brillouin–Brueckner orbitals, both fully occupied and active, but also improved coefficients within the MC-SCF function. If we had the exact Brueckner-MC-SCF this would not only imply that single excitations don't contribute to the exact wave function, but also that the wave function cannot be improved by excitations within the active space.

What one should strive at is to incorporate all excitations with relatively large coefficients in the MC-SCF reference function, such that the norm S of the residual σ is so small that the first iteration cycle is sufficient.

Like in the single-determinant reference case one can obtain a slight improvement over CEPA-0 including explicitly single or triple excitations in the first iteration and/or use partial summations of the σ_D contributions like in the ACPF scheme [31].

11. Stationary many-body perturbation theory for properties. Coupled and uncoupled MP-PT

We now want to study effects of electron correlation on properties, i.e. in principle we want to perform full CI with a Hamiltonian that contains the external perturbation $\mu\Omega$. There are four main possibilities.

(a) we regard the electron interaction λV (or the residual interaction $\lambda\tilde{V}$) and $\mu\Omega$ as two perturbations and apply double perturbation theory.

(b) we treat only $\mu\Omega$ by perturbation theory but take care of the electron interaction in an iterative non-perturbative way.

(c) we use 'finite perturbation theory', i.e. we perform nonperturbative calculations for some selected values of μ and extrapolate to $\mu = 0$.

(d) we treat $\lambda\tilde{V}$ by in powers of λ but $\mu\Omega$ in terms of 'finite perturbation theory', e.g. we perform MP2 calculations for selected values of μ and we extrapolate to $\mu = 0$.

In the double perturbative approaches of type (a) we have again the choice between taking the bare-nuclear Hamiltonian or the Hartree–Fock Hamiltonian as H_0.

Finite perturbation theory of type (c) does not need a special elaboration, the arguments of Sects. 8–10 can directly be used. It is particularly attractive that no new theory is needed. A disadvantage is that finite perturbation theory is hardly recommended for magnetic properties, where the Hamiltonian becomes complex and everything has to be done in complex arithmetics. There are also fundamental problems with finite perturbation theory, since as we shall see (Sect. 15) it does not satisfy a Hellmann–Feynman theorem. This has been known in coupled-cluster theory [32], but is somewhat unexpected for a theory based on stationarity of a well-defined functional.

Double perturbation theory is rather straightforward if one takes the bare nuclear Hamiltonian H_0 as unperturbed and the full electron interaction V, multiplied by λ (which at the end is set equal to 1) as well as $\mu\Omega$ as the two perturbations, i.e. one chooses

$$H(\lambda, \mu) = H_0 + \lambda V + \mu\Omega \tag{11.1}$$

Here one can directly apply the formalism of I. Sect. 5, that has briefly been recapitulated in the introduction, see Eqs. (2.14–16). We just identify V with V_{10}

and Ω with V_{01}. For E_{10} and E_{01} we get in view of (2.16) simple expectation values. For E_{11} we either need Y_{10} or Y_{01}, the construction of E_{20} or E_{02} requires Y_{10} or Y_{01}, respectively.

We are mainly interested in the case that Ω is a one-electron operator, given by (4.9a). We get for the kth order property corresponding to Ω

$$\Omega_k = \sum_l \lambda^l E_{lk} \tag{11.2}$$

E_{0k} is the respective property with complete neglect of the electron interaction, E_{1k} takes care of 1st order effects of the electron interaction etc.

Although this approach is straightforward, it is not recommended in atomic or molecular calculations, because the full electron interaction is too large to be treated as a perturbation, and convergence of the perturbation series is not guaranteed. It is usually better to treat at least the average electron interaction non-perturbatively, i.e. to use Møller–Plesset perturbation theory for the electron correlation (see Sect. 7).

Unfortunately there is no unique way to generalize MP to a double perturbation theory. This has the following reason. In MP theory the unperturbed Hamiltonian H_{HF} is based on a stationary condition for the expectation value $\langle\Phi|H|\Phi\rangle$ with respect to one-praticle transformations. It looks straightforward to base coupled MP-theory on coupled Hartree–Fock [34], where $\langle\Phi|H+\mu\Omega|\Phi\rangle$ is made stationary with respect to one-particle transformations.

This then defines an unperturbed Hamiltonian $H_{HF}(\mu)$, which is μ-dependent (and this not only due to the direct presence of $\mu\Omega$, but also to the μ-dependence of the Coulomb and exchange operators). Consequently we also get an effective electron interaction $\tilde{V}(\mu)$ that is μ-dependent. To define MP theory we multiply $\tilde{V}(\mu)$ by λ (which at the end is set equal to 1), i.e. we deal with the Hamiltonian

$$H(\lambda, \mu) = H_{HF}(\mu) + \lambda\tilde{V}(\mu) = \sum_k \mu^k H_{HF}^{(k)} + \sum_k \lambda\mu^k \tilde{V}^{(k)} \tag{11.3}$$

which has a rather complicated structure, since it also contains terms in μ^2, $\lambda\mu$ etc. Obviously the coupled Møller–Plesset (CMP) approach ('coupled' since it is based on coupled Hartree–Fock) satisfies a Brillouin theorem in the sense

$$\begin{aligned}\lim_{\lambda\to 0} \langle\Psi|[H(\lambda,\mu), X]|\Psi\rangle &= \langle\Psi_{HF}(\mu)|H_{HF}(\mu)|\Psi_{HF}(\mu)\rangle \\ &= \langle\Psi_{HF}(\mu)|H+\mu\Omega|\Psi_{HF}(\mu)\rangle = 0; \quad \forall X \in \mathscr{L}_c^{(1)}\end{aligned} \tag{11.4}$$

for all μ in a neighborhood of $\mu = 0$.

Of course (11.4) implies that singly excited configurations don't contribute to 1st order in λ, and this independently of μ. It further means that the 1st order correlation correction to all properties vanishes.

One would also like that a Hellmann–Feyman theorem holds in the sense

$$\frac{\partial}{\partial\mu}\langle\Psi|H(\lambda,\mu)|\Psi\rangle = \langle\Psi|\frac{\partial H(\lambda,\mu)}{\partial\mu}|\Psi\rangle = \langle\Psi|\Omega|\Psi\rangle \tag{11.5}$$

and so to all order in λ. This is not the case, (11.5) only holds if one does not expand in powers of λ. If one does make this expansion, the form (11.3) of $H(\lambda, \mu)$ has to be used, i.e.

$$\frac{\partial}{\partial\mu}\langle\Psi|H|\Psi\rangle = \langle\Psi|H_{HF}^{(1)} + \lambda\tilde{V}^{(1)}|\Psi\rangle \tag{11.6}$$

If one insists on a Hellmann–Feyman theorem in the form (11.5) one must start from a Hamiltonian in which, unlike in (11.3) the μ-dependence is just in a term $\mu\Omega$. This is achieved if we choose $\tilde{H}_{HF}$ and $\tilde{V}$ from ordinary (rather the coupled) Hartree–Fock theory

$$H = H_{HF} + \mu\Omega + \lambda\tilde{V} \tag{11.7}$$

with H_{HF} and $\tilde{V}$ μ-independent, which agree with the μ-dependent counterparts for $\mu = 0$. Obviously for the choice (11.7) the Brillouin theorem (11.4) no longer holds, except at $\mu = 0$.

Double perturbation theory based on (11.7) is related to uncoupled Hartree–Fock theory and should therefore be referred to as uncoupled Møller–Plesset theory (UCMP). It has been in use long before CMP [33].

We shall formulate both variants (Sects. 12 and 13). In this context the careful analysis of Møller–Plesset perturbation theory for properties must be mentioned [34] (see also [35]).

We agree with some of the conclusions of [34], especially as to the superiority of CMP over UCMP; but we have also some reservations as to CMP and we regard the incompatibility of a Brillouin theorem and a Hellmann–Feynman theorem as a serious defect of any kind of MP-based double perturbation theory.

For the double perturbation theory based on the bare nuclear Hamiltonian as H_0, i.e. for the Hamiltonian (11.1) both (11.5) and (11.4) hold, but the latter with $H = H_0 + V$ replaced by H_0. This 'Brillouin-theorem' for the bare nuclear Hamiltonian is, of course, different from the conventional Brillouin theorem. Anyhow, there is no unbalanced treatment between single and multiple excitations and hence no choice between a coupled and an uncoupled variant.

In addition to the dilemma to choose between coupled and uncoupled MP another drawback of double perturbation theory is that the expressions for the E_{kl} are rather lengthy and not easily interpreted, especially for the practically most important quantity E_{22}, the 2nd order correlation correction to a 2nd order property.

A good alternative is perturbation theory for $\mu\Omega$ with a non-perturbative treatment of electron correlation effects (see Sects. 14 and 15).

12. Uncoupled Møller–Plesset (UCMP) theory for properties

We satisfy the Brillouin-condition of Hartree–Fock theory for $X \in \mathcal{L}^{(1)}$ only in the absence of the extra perturbation, i.e. in the limit $\mu \to 0$.

We hence consider double perturbation theory for the decomposition (11.7) of the Hamiltonian, with H_{HF} defined by (4.24) and $\tilde{V}$ by (4.26c).

We get

$$E_0 = \langle\Phi|H_{HF}|\Phi\rangle = \langle\Phi|H|\Phi\rangle = E_{HF} \tag{12.1a}$$

$$E_{10} = \langle\Phi|\tilde{V}|\Phi\rangle = 0 \tag{12.1b}$$

$$E_{20} = \tfrac{1}{2}\langle\Phi|[\tilde{V}, Y_{10}]|\Phi\rangle \tag{12.1c}$$

$$E_{01} = \langle\Phi|\Omega|\Phi\rangle \tag{12.1d}$$

$$E_{02} = \tfrac{1}{2}\langle\Phi|[\Omega, Y_{01}]|\Phi\rangle \tag{12.1e}$$

$$E_{11} = \langle\Phi|[\Omega, Y_{10}]|\Phi\rangle = \langle\Phi|[\tilde{V}, Y_{01}]|\Phi\rangle = 0 \tag{12.1f}$$

$$\begin{aligned} E_{21} &= \tfrac{1}{2}\langle \Phi | 2[[\tilde{V}, Y_{10}], Y_{01}] + [[\Omega, Y_{10}], Y_{10}] + [[[H_{HF}, Y_{10}], Y_{10}], Y_{01}] | \Phi \rangle \\ &= \tfrac{1}{2}\langle \Phi | 2[[\tilde{V}, Y_{01}], Y_{10}] + [[\Omega, Y_{10}], Y_{10}] + [[[H_{HF}, Y_{01}], Y_{10}], Y_{10}] | \Phi \rangle \quad (12.1g) \\ E_{12} &= \tfrac{1}{2}\langle \Phi | 2[[\Omega, Y_{01}], Y_{10}] + [[\tilde{V}, Y_{01}, Y_{01}]] + [[[H_{HF}, Y_{01}], Y_{01}], Y_{10}] | \Phi \rangle \\ &= \tfrac{1}{2}\langle \Phi | [[\tilde{V}, Y_{01}], Y_{01}] | \Phi \rangle \quad (12.1h) \end{aligned}$$

provided that Y_{01} and Y_{10} are determined from

$$C_{10} \equiv \langle \Phi | [\tilde{V}, Z] + [[H_{HF}, Z], Y_{10}] | \Phi \rangle = 0; \quad Z \in \mathscr{L}_c^{(n)} \quad (12.2a)$$

$$C_{01} \equiv \langle \Phi | [\Omega, X] + [[H_{HF}, X], Y_{01}] | \Phi \rangle = 0; \quad X \in \mathscr{L}_c^{(n)} \quad (12.2b)$$

In (12.2b), although we have written $X \in \mathscr{L}_c^{(n)}$, only $X \in \mathscr{L}_c^{(1)}$ gives a non-trivial contribution, and (12.2b) is, in fact, the basic equation of uncoupled Hartree–Fock theory (see Sect. 4). On the other hand due to the Brillouin condition for $\mu = 0$, only double excitations contribute to Y_{10} in (12.2a).

The results for E_0, E_{10}, E_{20}, Y_{10} agree, of course, with the E_0, E_1, E_2, Y_1 of Sect 7 for ordinary MP.

The first-order correlation correction to the first-order property E_{11} (associated with Ω) vanishes since Ω (or Y_{01}) is a pure one-electron operator, Y_{10} (or $\tilde{V}$) a pure two-electron operator. For a similar reason the first and the last terms on the r.h.s. of (12.1h) vanish.

If a Brillouin theorem (11.4) would hold, also E_{12}, i.e. the first-order correlation correction to a second-order property, would vanish. This is, however, not the case (unlike in CMP theory, see Sect. 13).

In UCMP theory a Hellmann–Feynman theorem holds in the form (11.6) or (2.18). In fact the Eqs. (2.16b,c,d) are realized in (12.1d,f,g).

The most interesting quantity is usually E_{22}, i.e. the 2nd-order correlation correction to a 2nd-order property. Unfortunately the expression for this is rather lengthy (see I. 5.13) and its evaluation requires the knowledge of Y_{11} and Y_{20} or Y_{02}. It is probably easier to evaluate Y_{02} than Y_{20}. We hence determine Y_{11} and Y_{02} from

$$\begin{aligned} C_{11} = \langle \Phi | &[[\tilde{V}, X], Y_{01}] + [[\Omega, X], Y_{10}] + [[H_{HF}, X], Y_{11}] \\ &+ \tfrac{1}{2}[[[H_{HF}, X], Y_{01}], Y_{10}] + \tfrac{1}{2}[[[H_{HF}, X], Y_{10}], Y_{01}] | \Phi \rangle = 0 \quad (12.3a) \end{aligned}$$

$$C_{02} = \langle \Phi | [[\Omega, X], Y_{01}] + [[H_{HF}, X], Y_{02}] + \tfrac{1}{2}[[[H_{HF}, X], Y_{01}], Y_{01}] | \Phi \rangle = 0 \quad (12.3b)$$

and get E_{22} as

$$\begin{aligned} E_{22} = \tfrac{1}{2}\langle \Phi | &[[\tilde{V}, Y_{11}], Y_{01}] + 2[[\tilde{V}, Y_{10}], Y_{02}] + [[\Omega, Y_{10}], Y_{11}] \\ &+ [[[\tilde{V}, Y_{10}], Y_{01}], Y_{01}] + \tfrac{1}{2}[[[\Omega, Y_{10}], Y_{01}], Y_{10}] \\ &+ \tfrac{1}{2}[[[\Omega, Y_{10}], Y_{10}], Y_{10}] + \tfrac{1}{2}[[[H_0, Y_{10}], Y_{02}], Y_{10}] \\ &+ \tfrac{1}{2}[[[H_0, Y_{10}], Y_{10}], Y_{02}] + \tfrac{1}{2}[[[H_0, Y_{10}], Y_{11}], Y_{01}] \\ &+ \tfrac{1}{2}[[[H_0, Y_{10}], Y_{01}], Y_{11}] + \tfrac{1}{2}[[[[H_0, Y_{10}], Y_{01}], Y_{10}], Y_{01}] | \Phi \rangle \quad (12.4) \end{aligned}$$

One will probably not want to evaluate expressions that are so complicated and so difficult to interpret.

The fact that various authors have actually evaluated E_{22} [33, 36] is no argument against this general statement.

13. Coupled Møller–Plesset theory (CMP) for properties

To formulate coupled Møller–Plesset (CMP) theory we must first construct the μ-dependent Fock operator that appears in Eq. (11.3). Remember that $H_{HF}^{(0)} = H_{HF}$ has been defined as the one-electron operator (plus a constant term) which satisfies

$$\langle \Phi | H_{HF}^{(0)} | \Phi \rangle = \langle \Phi | H | \Phi \rangle \tag{13.1a}$$

$$\langle \Phi | [H_{HF}^{(0)}, X] | \Phi \rangle = \langle \Phi | [H, X] | \Phi \rangle; \quad \forall X \in \mathscr{L}_c \tag{13.1b}$$

In analogy we define $H_{HF}(\mu)$ as the one-electron operator which satisfies

$$\langle \Phi | \, e^{-Y(\mu)} H_{HF}(\mu) \, e^{Y(\mu)} | \Phi \rangle = \langle \Phi | \, e^{-Y(\mu)} (H + \mu\Omega) \, e^{Y(\mu)} | \Phi \rangle \tag{13.2a}$$

$$\langle \Phi | \, e^{-Y(\mu)} [H_{HF}(\mu), X] \, e^{Y(\mu)} | \Phi \rangle = \langle \Phi | \, e^{-Y(\mu)} [H + \mu\Omega, X] \, e^{Y(\mu)} | \Phi \rangle \tag{13.2b}$$

Expansion of (13.2) in powers of μ leads to (13.1) and

$$\langle \Phi | H_{HF}^{(1)} + [H_{HF}^{(0)}, Y_1] | \Phi \rangle = \langle \Phi | \Omega + [H, Y_1] | \Phi \rangle \tag{13.3a}$$

$$\langle \Phi | [H_{HF}^{(1)}, X] + [[H_{HF}^{(1)}, X], Y_1] | \Phi \rangle = \langle \Phi | [\Omega, X] + [[H, X], Y_1] | \Phi \rangle \tag{13.3b}$$

$$\begin{aligned} &\langle \Phi | H_{HF}^{(2)} + [H_{HF}^{(1)}, Y_1] + [H_{HF}^{(0)}, Y_2] + \tfrac{1}{2}[H_{HF}^{(0)}, Y_1] | \Phi \rangle \\ &= \langle \Phi | [\Omega, Y_1] + [H, Y_2] + \tfrac{1}{2}[[H, Y_1], Y_1] | \Phi \rangle \end{aligned} \tag{13.4a}$$

$$\begin{aligned} &\langle \Phi | [H_{HF}^{(2)}, X] + [[H_{HF}^{(1)}, Y_1] + [[H_{HF}^{(0)}, X], Y_2] + \tfrac{1}{2}[[[H_{HF}^{(0)}, X], Y_1], Y_1] | \Phi \rangle \\ &= \langle \Phi | [[\Omega, X], Y_1] + [[H, X], Y_2] + \tfrac{1}{2}[[[H, X], Y_1], Y_1] | \Phi \rangle \end{aligned} \tag{13.4b}$$

In view of (13.1b) we can simplify (13.3a) to

$$\langle \Phi | H_{HF}^{(1)} | \Phi \rangle = \langle \Phi | \Omega | \Phi \rangle \tag{13.5a}$$

Noting that $H = H_{HF}^{(0)} + \tilde{V}$ we can rewrite (13.3b) as (having first used that X and Y_1 can be interchanged)

$$\langle \Phi | [H_{HF}^{(1)} - \Omega, X] | \Phi \rangle = \langle \Phi | [[\tilde{V}, Y_1], X] | \Phi \rangle \tag{13.5b}$$

This means, that whenever $H_{HF}^{(1)}$ appears in the form $\langle \Phi | [H_{HF}^{(1)}, X] | \Phi \rangle$ it can be written as

$$H_{HF}^{(1)} = \Omega + [\tilde{V}, Y_1] \tag{13.6a}$$

i.e. that $H_{HF}^{(1)}$ which is a one-electron operator, is equal to the one-electron-part of the r.h.s. of (13.6a)

Since $\tilde{V}$ is a genuine two-electron operator and Y_1 is a one-electron operator, their commutator has no constant part, so also (13.5a) is compatible with (13.6a). Analogously one gets from (13.4) in the same meaning of the equality sign as in (13.6a)

$$H_{HF}^{(2)} = [\Omega, Y_1] + [\tilde{V}, Y_2] + \tfrac{1}{2}[[\tilde{V}, Y_1], Y_1] - [H_{HF}^{(1)}, Y_1] \tag{13.6b}$$

The $\tilde{V}^{(k)}$ defined by (11.3) are then

$$\tilde{V}^{(0)} = \tilde{V} \tag{13.7a}$$

$$\tilde{V}^{(1)} = \Omega - H_{HF}^{(1)} = -[\tilde{V}, Y_1] \tag{13.7b}$$

$$\tilde{V}^{(2)} = -\tilde{H}_{HF}^{(2)} \tag{13.7c}$$

We can now apply double perturbation for the Hamiltonian (11.3).

This is somewhat more complicated than in the UMP case, due to the extra terms $O(\mu^2)$, $O(\lambda\mu)$, $O(\lambda\mu^2)$ in the Hamiltonian (11.3). In addition to the expression (2.16), one has additional contributions

$$E_{11}\colon \langle \Psi_0 | V_{11} | \Psi_0 \rangle = \langle \Phi | \tilde{V}^{(1)} | \Phi \rangle \tag{13.8a}$$

$$E_{12}\colon \langle \Psi_0 | V_{12} + [V_{11}, Y_{01}] + [V_{02}, Y_{10}] | \Psi_0 \rangle = \langle \Phi | \tilde{V}^{(2)} + [\tilde{V}^{(1)}, Y_1] + [H_{HF}^{(2)}, \sigma_{10}] | \Phi \rangle \tag{13.8b}$$

$$E_{21}\colon \langle \Psi_0 | [V_{11}, Y_{10}] | \Psi_0 \rangle = \langle \Phi | [\tilde{V}^{(1)}, \sigma_{10}] | \Phi \rangle \tag{13.8c}$$

Let us first consider the terms for $\lambda = 0$, i.e. for the first label equal to 0

$$\tilde{E}_0 = \langle \Phi | H_{HF}^{(0)} | \Phi \rangle = \langle \Phi | H | \Phi \rangle \tag{13.9a}$$

$$\tilde{E}_{01} = \langle \Phi | H_{HF}^{(1)} | \Phi \rangle = \langle \Phi | \Omega | \Phi \rangle \tag{13.9b}$$

$$\tilde{E}_{02} = \langle \Phi | H_{HF}^{(2)} + \tfrac{1}{2}[H_{HF}^{(1)}, Y_1] | \Phi \rangle = \tfrac{1}{2} \langle \Phi | [\Omega, Y_1] | \Phi \rangle \tag{13.9c}$$

with Y_1 determined from

$$0 = \langle \Phi | [H_{HF}^{(1)}, X] + [[H_{HF}^{(0)}, Y_1], X] | \Phi \rangle = \langle \Phi | [\Omega, X] + [[H, Y_1], X] | \Phi \rangle \tag{13.10}$$

i.e. – as expected – we retrieve CHF theory (see Sect. 4). We have put a tilde on the E_{kl} in (13.9) to mark the difference to the results of Sect. 12. A tilde on Y_1 is not necessary, since the counterpart of Y_1 in Sect. 12 is Y_{01}. Note that Y_{01} and Y_1 are obtained from different stationarity conditions and are hence different.

We now make the ansatz

$$\Psi = e^{\sigma} \Phi; \quad \sigma = \sum_{k,l} \lambda^k \mu^l \sigma_{kl} \tag{13.11}$$

i.e. the Y_{kl} of Sect. 12 are now replaced by the σ_{kl}. For $\mu = 0$, i.e. the 2nd label equal to 0, we get, of course, the same results (12.1b,c) as in Sect. 12, just with Y_{10} replaced by σ_{10} which is solution of (12.2a). The counterpart of (12.2b) would be (13.10) with $X \in \mathscr{L}_c^{(1)}$ replaced by $Z \in \mathscr{L}_c^{(n)}$. However, since both $H_{HF}^{(0)}$ and $H_{HF}^{(1)}$ are one-electron operators, only the $X \in \mathscr{L}_c^{(1)}$ give non-vanishing contributions. Hence $\sigma_{01} \equiv 0$ and Y_1 plays the role of Y_{01}. Let us now look at the mixed terms

$$E_{11} = \langle \Phi | \tilde{V}^{(1)} + [\tilde{V}^{(0)}, Y_1] | \Phi \rangle = \langle \Phi | -[\tilde{V}, Y_1] + [\tilde{V}, Y_1] | \Phi \rangle = 0 \tag{13.12a}$$

$$E_{12} = \tfrac{1}{2} \langle \Phi | 2\tilde{V}^{(2)} + [\tilde{V}^{(1)}, Y_1] + [H_{HF}^{(2)}, \sigma_{10}] + [[\tilde{V}, OY_1], Y_1] | \Phi \rangle = 0 \tag{13.12b}$$

$$\begin{aligned} E_{21} &= \langle \Phi | [\tilde{V}^{(1)}, \sigma_{10}] + [[\tilde{V}^{(0)}, Y_1], \sigma_{10}] + \tfrac{1}{2}[[H_{HF}^{(1)}, \sigma_{10}], \sigma_{10}] \\ &\quad + \tfrac{1}{2}[[[H_{HF}^{(0)}, Y_1], \sigma_{10}], \sigma_{10}] | \Phi \rangle \\ &= \tfrac{1}{2} \langle \Phi | [[\tilde{V}^{(0)}, Y_1], \sigma_{10}] + [[\Omega, \sigma_{10}], \sigma_{10}] + [[[H_{HF}, Y_1], \sigma_{10}], \sigma_{10}] | \Phi \rangle \end{aligned} \tag{13.12c}$$

To see that E_{12} vanishes, one notes that $H_{HF}^{(2)}$ is a one (and zero) particle operator and σ_{10} a two-particle excitation operator, such that the expectation value of the commutator vanishes. The other contributions cancel if one inserts their actual expressions.

The first term is the first expression of E_{21} vanishes because $\tilde{V}^{(1)} = -H_{HF}^{(1)}$ is a one-electron operator and σ_{10} a two-electron operator. The expression for E_{22} is so lengthy that we don't give it here.

It is noteworthy that the 1st order correlation correction vanishes both for a first order property ($E_{11} = 0$) and a second order property ($E_{12} = 0$).

Let us now compare the results of UCMP and CMP. Of course, there is no difference for E_0, E_{10}, E_{20}, which don't involve $\mu\Omega$, there is further agreement for E_{01}. Although E_{02} and $\tilde{E}_{02}$ as given by (12.1e) and (13.9c) look formally the same, they differ because $Y_1 \neq Y_{01}$. The difference between (12.2a) which determines Y_{01} and (13.10) which determines Y_1 is that Ω in (12.2b) is replaced by $H^{(1)}_{HF} = \Omega + [\tilde{V}, Y_1]$ in (13.10). Let us look at the respective functionals that represent the second order energies. These differ by a term

$$\tfrac{1}{2}\langle\Phi|[[\tilde{V}, Y_1], Y_1]|\Phi\rangle \tag{13.13}$$

This term is taken care of in CMP to 0th order in λ and is ignored in UCMP. However, this term appears to 1st order in λ (see (12.1h)) for UCMP, while there is no term to first order in λ for CMP. The difference between UMP and CMP is hence that in CMP the entire functional is made stationary, in UMP only a part (hopefully the dominant one), while the remainder is evaluated with an approximate Y_1. So UMP with inclusion of 1st order correlation corrections is just an approximation to CMP, and corresponds to the first iteration beyond UCHF on the way to CHF.

Sadlej [34] has proposed to refer to contributions like (13.13) that appear in UCMP to 1st (or higher order) in λ, but not in CMP as 'apparent correlation effects' in contrast to 'true correlation effects' as they appear in CMP to 2nd order in λ. For a different derivation of CMP see [37].

What are the final conclusions about CMP and UCMP? For both approaches E_{11}, i.e. the first-order correction to a first-order property vanishes. Both approaches give rather simple results for E_{12}, i.e. the first order correlation correction to a second order property and E_{21} the second order correlation correction for a first order property. In CMP E_{12} vanishes, which means CHF is correct to $O(\lambda)$, to this order CMP is obsolete, while UCMP just corrects approximately the error of UCHF with respect to CHF. For E_{21} there is relatively little difference between CMP and UCMP. For both schemes E_{22} is so complicated that one does not want to evaluate this quantity by double perturbation theory, and rather looks for an alternative approach.

14. Stationary unitary coupled-cluster theory for properties

We now want to study properties on the level of full CI. Let Ψ_0 be the full-CI wave function which makes the expectation value

$$E_0 = \langle\Psi_0|H|\Psi_0\rangle; \quad \Psi_0 = e^{\sigma}\Phi \tag{14.1}$$

stationary, i.e. which satisfies the Brillouin condition

$$\langle\Psi_0|[H, Z]|\Psi_0\rangle = 0; \quad \forall Z \in \mathscr{L}^{(n)}_c \tag{14.2}$$

with H the Hamiltonian, including electron interaction, in the absence of external field $\mu\Omega$. We write the wave function in the presence of the external field as

$$\Psi(\mu) = e^{\tau}\Psi_0; \quad \tau = \sum_{k=1}^{\infty} \mu^k \tau_k \tag{14.3}$$

we then get

$$E = \langle \Psi(\mu) | H + \mu\Omega | \Psi(\mu) \rangle = \sum_{k=1}^{\infty} \mu^k E_k \tag{14.4}$$

$$E_1 = \langle \Psi_0 | \Omega | \Psi_0 \rangle = \langle \Phi | \Omega + [\Omega, \sigma] + \tfrac{1}{2}[[\Omega, \sigma], \sigma] + \cdots | \Phi \rangle \tag{14.5a}$$

$$E_2 = \tfrac{1}{2} \langle \Psi_0 | [\Omega, \tau_1] | \Psi_0 \rangle \tag{14.5b}$$

with τ_1 making the functional

$$F(\tau_1) = \langle \Psi_0 | [\Omega, \tau_1] + \tfrac{1}{2}[[H, \tau_1], \tau_1] | \Psi_0 \rangle \tag{14.6}$$

stationary, i.e. satisfying

$$\langle \Psi_0 | [\Omega, Z] + [[H, Z], \tau_1] | \Psi_0 \rangle = 0; \quad \forall Z \in \mathscr{L}_c^{(n)} \tag{14.7}$$

This looks formally rather simple. The difficulty is that we rarely know Ψ_0, and that even (14.7) is a too high-dimensional linear system of equations.

In the spirit of Sect. 8 we can approximate the expectation value (14.1) by truncating the Hausdorff expansion at CEPA-0 level

$$E_0 = \langle \Phi | e^{-\sigma} H e^{\sigma} | \Phi \rangle \approx E_0^{(2)} = \langle \Phi | H + [H, \sigma] + \tfrac{1}{2}[[H, \sigma], \sigma] | \Phi \rangle \tag{14.8}$$

with the stationary condition

$$\langle \Phi | [H, Z] + \tfrac{1}{2}[[H, Z], \sigma] + \tfrac{1}{2}[[H, \sigma], Z] | \Phi \rangle = 0 \tag{14.9}$$

In the same spirit we truncate the expectation value (14.4)

$$\begin{aligned} E = {} & \langle \Psi_0 | H + \mu\Omega + [H + \mu\Omega, \mu\tau_1] + [H, \mu^2\tau_2] + \tfrac{1}{2}[[H, \mu\tau_1], \mu\tau_1] + O(\mu^3) | \Psi_0 \rangle \\ \approx {} & E^{(2)} = \langle \Phi | H + [H, \sigma] + \tfrac{1}{2}[[H, \sigma], \sigma] + \mu\{\Omega + [H, \tau_1] \\ & + [\Omega + [H, \tau_1], \sigma] + \tfrac{1}{2}[[\Omega + [H, \tau_1], \sigma], \sigma]\} \\ & + \mu^2\{[\Omega, \tau_1] + [H, \tau_2] + \tfrac{1}{2}[[H, \tau_1], \tau_1] \\ & + [[\Omega, \tau_1] + [H, \tau_2] + \tfrac{1}{2}[[H, \tau_1], \tau_1], \sigma] \\ & + \tfrac{1}{2}[[[\Omega, \tau_1] + [H, \tau_2] + \tfrac{1}{2}[[H, \tau_1], \tau_1], \sigma], \sigma]\} + O(\mu^3) | \Phi \rangle \end{aligned} \tag{14.10}$$

The zeroth order (in μ) is obviously $E_0^{(2)}$ as defined by (14.8), the first order energy $E_1^{(2)}$ can, by means of (14.9) be simplified to

$$E_1^{(2)} = \langle \Phi | \Omega + [\Omega, \sigma] + \tfrac{1}{2}[[\Omega, \sigma], \sigma] + \tfrac{1}{2}[H, [\tau_1, \sigma]] + \tfrac{1}{2}[[[H, \tau_1], \sigma], \sigma] | \Phi \rangle \tag{14.11}$$

This is, unfortunately, not the Hellmann–Feynman result (14.5a) truncated at terms quadratic in σ, but there are additional terms. These only vanish, if rather than (14.9) we satisfy the stationarity condition (14.2) truncated after terms quadratic in σ, i.e.

$$\langle \Phi | [H, Z] + [[H, Z], \sigma] + \tfrac{1}{2}[[[H, Z], \sigma], \sigma] | \Phi \rangle = 0; \quad \forall Z \in \mathscr{L}_c^{(n)} \tag{14.12}$$

To evaluate τ_1 we expand (14.7) in powers of σ and truncate after quadratic terms

$$\begin{aligned} & \langle \Phi | [\Omega, Z] + [[H, Z], \tau_1] + [[\Omega, Z] + [[H, Z], \tau_1], \sigma] \\ & \quad + \tfrac{1}{2}[[[\Omega, Z] + [[H, Z], \tau_1], \sigma] | \Phi \rangle = 0 \end{aligned} \tag{14.13}$$

Using (14.12) and (14.13) we can simplify the second order energy $E_2^{(2)}$ contained in (14.10) to

$$E_2^{(2)} = \tfrac{1}{2} \langle \Phi | [\Omega, \tau_1] + [[\Omega, \tau_1], \sigma] + \tfrac{1}{2}[[\Omega, \tau_1], \sigma], \sigma] | \Phi \rangle \tag{14.14}$$

which one would also get from (14.5b) on expanding the powers of σ and truncating after the quadratic terms.

The important message is that for the calculation of properties on CEPA-0 level one should – for the determination of σ – not use the stationarity condition (14.9) but rather (14.12). To some extent this means that in order to get properties correct to 2nd order in σ we have to make the energy stationary to a higher order in σ. It is also important that Z in (14.12) is not limited to double excitations, since the dominant part of τ_1 will consist of single excitations.

The scheme presented in this section will be called 'perturbative CEPA-0'.

Note that in this section we have never assumed that the single determinant reference function Φ satisfies a Brillouin theorem. In fact this is only useful in perturbation theory, where it guarantees that first-order corrections vanish (see Sect. 9). In a nonperturbative theory as in a CEPA-0-like scheme it only matters that Φ is chosen such that $S = \|\sigma\|$ is sufficiently small, that truncation of the Hausdorff expansion is justified.

It is interesting to note that we have automatically satisfied the Brillouin–Brueckner condition in requiring that the Hellmann–Feynman theorem should be satisfied. In fact, if (14.12) holds for all $Z \in \mathscr{L}_c^{(n)}$, it holds a fortiori for all $X \in \mathscr{L}_c^{(1)}$, for which (14.12) just becomes the Brillouin–Brueckner condition (9.5). If we observe further that τ_1 is dominated by single excitations and realize that (14.12) is needed for the simplification of $E_1^{(2)}$ and $E_2^{(2)}$ with $Z \in \tau_1$, we conclude that (14.12) with $Z \in \mathscr{L}_c^{(n)}$ replaced by $Z \in \mathscr{L}_c^{(1)}$ would 'nearly' have the same effect, i.e. the Hellmann–Feynman theorem would be almost satisfied.

15. Three other variants of a CEPA-0 type approach to properties

Let us now consider what one would get if one simply performs a CEPA-0 calculation in the presence of $\mu\Omega$. There are at first glance two variants, in the first of these the reference function Φ is constructed from Hartree–Fock theory in the absence of the field, and $\mu\Omega$ is only added in the CEPA-0 part. This corresponds somehow to UCMP. In the other variant, related to CMP, one optimizes even Φ in the presence of $\mu\Omega$. Note that in the formalism given in Sect. 14 this distinction was not necessary since the operator τ performs the full transformation from Ψ_0 to $\Psi(\mu)$ and since no stationarity condition was imposed on Φ.

We start with what may be called 'uncoupled CEPA-0'. We hence take the energy functional

$$E(\mu) = \langle \Phi | H + \mu\Omega + [H + \mu\Omega, \sigma], \sigma] | \Phi \rangle \tag{15.1}$$

where σ now depends, of course, on μ

$$\sigma = \sum_k \mu^k \sigma_k \tag{15.2}$$

We expand $E(\mu)$ in powers of μ

$$E(\mu) = \sum_{k=0}^{\infty} \mu^k E_k \tag{15.3}$$

and get

$$E_0 = \langle \Phi | H + [H, \sigma_0] + \tfrac{1}{2}[[H, \sigma_0], \sigma_0] | \Phi \rangle \tag{15.4a}$$

$$E_1 = \langle \Phi | \Omega + [\Omega, \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_0] + [H, \sigma_1] + \tfrac{1}{2}[[H, \sigma_0], \sigma_1] + \tfrac{1}{2}[[H, \sigma_1], \sigma_0] | \Phi \rangle \tag{15.4b}$$

$$\begin{aligned} E_2 = \langle \Phi | &[\Omega, \sigma_1] + \tfrac{1}{2}[[H, \sigma_1], \sigma_1] + \tfrac{1}{2}[[\Omega, \sigma_1], \sigma_0] + [H, \sigma_2] + \tfrac{1}{2}[[H_0, \sigma_0], \sigma_2] \\ &+ \tfrac{1}{2}[[H, \sigma_2], \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_1] | \Phi \rangle \end{aligned} \tag{15.4c}$$

Of course, σ_0 is determined by the stationarity condition for E_0

$$\langle \Phi | [H, Z] + \tfrac{1}{2}[[H, Z], \sigma_0] + \tfrac{1}{2}[[H, \sigma_0], Z] | \Phi \rangle = 0 \tag{15.5}$$

If this is satisfied, E_1 is simplified to the Hellmann–Feynman expression

$$E_1 = \langle \Phi | \Omega + [\Omega, \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_0] | \Phi \rangle \tag{15.6}$$

Stationarity of E_2 leads to the condition for σ_1

$$\langle \Phi | [\Omega, Z] + \tfrac{1}{2}[[\Omega, Z], \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], Z] + \tfrac{1}{2}[[H, Z], \sigma_1] + \tfrac{1}{2}[[H, \sigma_1], Z] | \Phi \rangle = 0 \tag{15.7}$$

If both (15.5) and (15.7) are satisfied, we get for E_2:

$$E_2 = \tfrac{1}{2} \langle \Phi | [\Omega, \sigma_1] + \tfrac{1}{2}[[\Omega, \sigma_1], \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_1] | \Phi \rangle \tag{15.8}$$

While the result for E_1 is the same as before (14.5a), just with σ_0 and σ determined from different stationarity conditions, E_2 is significantly different from (14.14), it only contains correlation corrections to 1st order in σ_0, not to second order in σ_0. This somehow disqualifies 'uncoupled CEPA-0'.

Let us now consider the 'coupled' variant. It becomes more lengthy. We first make

$$\langle \Phi | e^{-Y}(H + \mu\Omega)\, e^{Y} | \Phi \rangle \tag{15.9}$$

stationary to 0th and 2nd order in μ with respect to one-particle transformations. This leads to

$$\langle \Phi | [H, X] | \Phi \rangle = 0 \tag{15.10a}$$

$$\langle \Phi | [[H, Y_1], X] + [\Omega, X] | \Phi \rangle = 0 \tag{15.10b}$$

as in CHF theory. We then take care of electron correlation on CEPA-0 level, i.e. we take the functional

$$E(\sigma) = \langle \Phi | e^{-Y}\{H + \mu\Omega + [H + \mu\Omega, \sigma] + \tfrac{1}{2}[[H + \mu\Omega, \sigma], \sigma]\}\, e^{Y} | \Phi \rangle \tag{15.11}$$

and make this stationary with respect to variations of σ. We first expand σ and then $E(\sigma)$ in powers of μ and get

$$E_0 = \langle \Phi | H + [H, \sigma_0] + \tfrac{1}{2}[H, \sigma_0], \sigma_0] | \Phi \rangle \tag{15.12a}$$

$$\begin{aligned} E_1 = \langle \Phi | &[H, Y_1] + [[H, \sigma_0], Y_1] + \tfrac{1}{2}[[[H, \sigma_0], \sigma_0], Y_1] + \Omega + [\Omega, \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_0] \\ &+ [H, \sigma_1] + \tfrac{1}{2}[[H, \sigma_1], \sigma_0] + \tfrac{1}{2}[[H, \sigma_0], \sigma_1] | \Phi \rangle \end{aligned} \tag{15.12b}$$

$$\begin{aligned} E_2 = \langle \Phi | &[H, Y_2] + [[H, \sigma_0], Y_2] + \tfrac{1}{2}[[[H, \sigma_0], \sigma_0], Y_2] \\ &+ \tfrac{1}{2}[[H, Y_1], Y_1] + \tfrac{1}{2}[[[H, \sigma_0], Y_1], Y_1] + \tfrac{1}{4}[[[[H, \sigma_0], \sigma_0], Y_1], Y_1] \\ &+ [\Omega, Y_1] + [[\Omega, \sigma_0], Y_1] + \tfrac{1}{2}[[[\Omega, \sigma_0], \sigma_0], Y_1] \\ &+ [H, \sigma_2] + \tfrac{1}{2}[[H, \sigma_2], \sigma_0] + \tfrac{1}{2}[[H, \sigma_0], \sigma_2] + \tfrac{1}{2}[[H, \sigma_1], \sigma_1] \\ &+ [\Omega, \sigma_1] + \tfrac{1}{2}[[\Omega, \sigma_1], \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_1] \\ &+ [[H, \sigma_1], Y_1] + \tfrac{1}{2}[[[H, \sigma_1], \sigma_0], Y_1] + \tfrac{1}{2}[[[H, \sigma_0], \sigma_1], Y_1] | \Phi \rangle \end{aligned} \tag{15.12c}$$

The stationarity conditions are

$$\langle \Phi | [H, Z] + \tfrac{1}{2}[[H, Z], \sigma_0] + \tfrac{1}{2}[[H, \sigma_0], Z] | \Phi \rangle = 0 \qquad (15.13a)$$

$$\begin{aligned}&\langle \Phi | \tfrac{1}{2}[[H, \sigma_1], Z] + \tfrac{1}{2}[[H, Z], \sigma_1] + [\Omega, Z] + \tfrac{1}{2}[[\Omega, Z], \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], Z] \\ &\quad + [[H, Z], Y_1] + \tfrac{1}{2}[[[H, Z], \sigma_0], Y_1] + \tfrac{1}{2}[[H, \sigma_0], Z], Y_1 | \Phi \rangle = 0 \qquad (15.13b)\end{aligned}$$

Using (15.13a) and (15.10a) we can simplify (15.12b) to

$$E_1 = \langle \Phi | [[H, \sigma_0], Y_1] + \tfrac{1}{2}[[[H, \sigma_0], \sigma_0], Y_1] + \Omega + [\Omega, \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0] , \sigma_0] | \Phi \rangle \qquad (15.14a)$$

This is not the desired Hellmann–Feynman expression like (14.5a)

When one simplifies (15.12c) by means of (15.13a,b) the expression is still very lengthy and differs even more from a Hellmann–Feynman expression than does (15.14a). That 'coupled CEPA-0' does not satisfy a Hellmann–Feynman theorem is a strong argument against this approach.

What can we do after having discarded both coupled and uncoupled CEPA-0? Let us wonder what we have to do in order that a Hellmann-Feynman theorem holds. This leads us to the 'coupled Brillouin–Brueckner CEPA-0' approach for properties. We easily see that we get a Hellmann–Feynman expression for E_1, if we replace the Brillouin condition (15.10a) by the Brillouin–Brueckner condition

$$\langle \Phi | [H, X] + [[H, \sigma_0], X] + \tfrac{1}{2}[[[H, \sigma_0], \sigma_0], X] | \Phi \rangle = 0 \qquad (15.15a)$$

Then we get, in fact

$$E_1 = \langle \Phi | \Omega + [\Omega, \sigma_0] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_0] | \Phi \rangle \qquad (15.14b)$$

On first glance (15.15a) looks different from the Brillouin–Brueckner condition formulated in Sect. 9. However, by means of the Jacobi identity one can reformulate (15.15a) to

$$\begin{aligned}&\langle \Phi | [H, X] + [[H, X], \sigma_0] + \tfrac{1}{2}[[[H, X], \sigma_0], \sigma_0] | \Phi \rangle \\ &\quad + \langle \Phi | [H, [\sigma_0, X]] + \tfrac{1}{2}[[H, [\sigma_0, X]], \sigma_0] + \tfrac{1}{2}[[H, \sigma_0], [\sigma_0, X] | \Phi \rangle \qquad (15.15b)\end{aligned}$$

We recognize the first part of (15.15b) as (9.4), while the second part vanishes due to the CEPA-0 stationarity condition (15.10b) for $Z = [\sigma_0, X]$. Similarly the CHF condition (15.10b) does not lead to a sufficient simplification of (15.12c). We rather should satisfy the 'coupled-Brueckner–Hartree–Fock condition'

$$\begin{aligned}&\langle \Phi | [[H, Y_1], X] + [[H, \sigma_0], Y_1], X] \\ &\quad + [[\Omega, \sigma_0], X] + \tfrac{1}{2}[[[[H, \sigma_0], \sigma_0], \sigma_0], Y_1], X] + \tfrac{1}{2}[[[\Omega, \sigma_0], \sigma_0], X] | \Phi \rangle = 0 \qquad (15.16)\end{aligned}$$

Both (15.15a) and (15.16) imply that we make the energy in the presence of correlation effects stationary with respect to one-particle transformations, i.e. that we minimize the functional (15.11) with respect to variations of Y_1 (and, of course, also with respect to variations of σ). To minimize (15.9) instead is, so to say, not consistent. This stresses again that a Brillouin condition for the reference function, which is meaningful if the electron correlation is treated by perturbation theory, is not of much use in a nonperturbative context.

If (15.13a,b) and (15.15a, 16) hold, the expression (15.12c) for E_2 can be simplified to

$$\begin{aligned}E_2 = \tfrac{1}{2}\langle \Phi | &[\Omega, Y_1] + [[\Omega, \sigma_0], Y_1] + \tfrac{1}{2}[[[\Omega, \sigma_0], \sigma_0], Y_1] \\ &+ [\Omega, \sigma_1] + \tfrac{1}{2}[[\Omega, \sigma_0], \sigma_1] | \Phi \rangle \qquad (15.17)\end{aligned}$$

We see that the response to the perturbation $\mu\Omega$ is shared by two operators Y_1 and σ_1. The difference between Y_1 and σ_1 is that Y_1 is – by construction – a one-electron operator, while σ_1 is – in view of the Brillouin–Brueckner condition – a two-electron operator. Note also that Φ is the Brueckner determinant rather than the Hartree–Fock determinant.

We realize that the 'coupled Brillouin–Brueckner CEPA-0' is not very different from the 'perturbed CEPA-0' of Sect. 14. The operator τ of Sect. 14 is now split into two parts σ_1 and Y_1. Only the dominant part Y_1 is treated 'exactly' as compared to 'perturbed CEPA-0', while the 'correlation correction' σ_1 is treated on a lower level of approximation.

The message is that the 'perturbed CEPA-0 approach' of Sect. 14 is probably the best choice, but the 'coupled Brueckner CEPA-0' may be a decent approximation to 'perturbed CEPA-0', while both 'uncoupled' and 'coupled' CEPA-0 have to be discarded. From this section we can also learn something about the validity of finite perturbation theory in a CEPA-0 context. In fact we have simulated this in the present section by 'uncoupled' and 'coupled' CEPA-0.

We conclude that a finite perturbation theory in which one performs the starting Hartree–Fock calculation for $\mu = 0$ and where one includes $\mu\Omega$ only in the CEPA-0 part, is inaccurate for second-order properties, as is seen from (15.8) as compared to (14.4). It is acceptable for first-order properties, but these are, anyhow, more easily evaluated as expectation values.

If one includes $\mu\Omega$ already in the Hartree–Fock calculations and continues then with CEPA-0 one gets results that don't satisfy the Hellmann–Feynman theorem, i.e. the result for a first-order property will be different depending on whether it is obtained by finite perturbation theory or as an expectation value. The error of the property will be linear rather than quadratic in the error of the wave function. Similar problems arise for second-order properties.

Even in finite perturbation theory, in order to get consistent results, one should at least satisfy the Brillouin–Brueckner condition for all selected values of μ.

16. Coupled MC-CEPA-0 for properties

In coupled MC-CEPA-0 theory for properties we proceed as in Sect. 14, just with the reference function Φ not the single Slater determinant of Hartree–Fock theory but an MC-SCF function. Since in Sect. 14 we have never used explicitly that Φ is a single Slater determinant, almost everything can be taken over. Of course, the operator basis is a little more complicated, as discussed in Sect. 9 and it is less obvious than in the closed-shell case which basis operators should be included in the first iteration cycle, i.e. on CEPA-0 level.

17. Conclusions

Stationary perturbation theory based on the Lie algebraic formulation of the variation principle – is the appropriate frame for the derivation of Hartree–Fock theory and MC-SCF theory as well as of coupled Hartree–Fock and coupled MC-SCF in a compact and elegant way. Also many-body perturbation theory as well as Møller–Plesset theory can be so formulated in a simple and transparent manner. The main interest in the application of the Lie algebraic formulation of the variation principle is, however, the quadratically convergent iterative approach to the construction of the full-CI wave function, where the first iteration cycle is the well-known CEPA-0 approach. Protagonists of coupled-cluster (CC) theory might object at this point that a straightforward CC approach is – at its lowest level, i.e. CC-D or CC-SD – not much more complicated than CEPA-0 and gives usually more satisfactory results. This is true, but conventional coupled-cluster theory does not satisfy a stationarity condition of the type that we need for a unitary formulation. The author is convinced that on the long run the advantages of the hierarchy of approximations to full CI (including higher than the first iteration and the option for a multiconfiguration reference, but also corrections like CPF) will take over as compared to conventional coupled-cluster theory (especially if one is interested in properties, for which a different philosophy than for energies is required).

The treatment of properties has, in the theory of the coupled cluster approach, so far only played a marginal role [28, 32]. There is no doubt that the stationary context is much more important for properties than for energies, and that the unitary formulation wins back much of what it may have lost for energies. It is almost imperative to use such methods for the calculation of properties including correlation effects, which do obey a Hellmann–Feynman theorem, such that there is no ambiguity as to how the property should be evaluated.

Double perturbation theory (DPT) does not appear to be a good choice for the calculation of correlation effects on a property. DPT based on the bare nuclear Hamiltonians as describing the unperturbed problem is formally straightforward, but has the drawback that the full electron iteraction is too large in order to be treated by perturbation theory. DPT based on the Møller–Plesset pertition of the Hamiltonian faces the dilemma that two variants, coupled Møller–Plesset (CMP) and uncoupled Møller–Plesset (UCMP) can be defined, none of which is entirely satisfactory. CMP satisfies a Brillouin theorem which guarantees that the first-order correlation corrections, not only to the energy, but also to 1st order and 2nd order properties vanish, such that the dominant correlation corrections are in the second-order terms. On the other hand CMP does not satisfy a Hellmann–Feynman theorem which would guarantee that the error in the property is quadratic in the error of the wave function. UCMP does satisfy a Hellmann–Feynman theorem, but no Brillouin theorem. Hence there are correlation effects on properties formally to 1st order in correlation, but these are no 'true' correlation effects since they are taken care of to 0th order in coupled Hartree–Fock (on which CMP is based). The most serious drawback of either variant of MP double perturbation theory is that the most interesting quantity, namely E_{22} the 2nd order correlation correction to a 2nd order property is so complicated that one would not hesitate to prefer an alternative non-perturbative way.

A theory of properties on CEPA-0 level, i.e. with a non-perturbative treatment of correlation effects, can be formulated in a rather transparent way such that both a Brillouin and a Hellmann–Feynman-theorem hold. However, there is a price to pay. The Brillouin-condition that one has to satisfy is rather a Brillouin–Brueckner condition, i.e. it implies stationarity of the energy with respect to variations of the reference function in the presence of the leading terms that describe correlation. Moreover, even the counterpart of coupled Hartree–Fock theory must be performed on the same level, i.e. in the presence of correlation.

We have, in this paper, never specified the type of the external perturbation described by $\mu\Omega$. In fact one will mostly be concerned with magnetic interactions, because for electric perturbations finite perturbation theory is rather simple, while finite perturbation theory for magnetic perturbations would require full complex arithmetics, which is certainly inconvenient. Of course, the Hamiltonian of a molecule in a magnetic field is not linear in the field strength, but contains also a quadratic term. The necessary generalization of the formalism is, however, almost trivial. The same is true for two magnetic perturbations as in the theory of chemical shifts. More serious, at any level of theory is the gauge problem. A transformation to individual gauge origins for different orbitals in the spirit of the IGLO approach is possible, but complicates the formalism.

Note that even if finite perturbation theory is possible, it need not be the best choice, since it does not necessarily satisfy a Hellmann–Feynman theorem.

We have, at various instances, assumed that the perturbation $\mu\Omega$ is a one-electron operator. Some results are based on this assumption. If one wants to apply the present theory to, say, an Ω which is a two-electron operator, some modifications in the final formulas may be necessary.

One should mention that we have always argued in terms of a spin-orbital basis, while in practice one will mostly deal with spinfree systems and Lie algebras with spinfree excitation operators are advantageous. However, the necessary modifications are rather straightforward.

The reader may want to know which method is recommended for the evaluation of correlation corrections to properties. This is the perturbed CEPA-0 approach of Sect. 14 which satisfies a Hellmann–Feynman theorem and a Brillouin–Brueckner condition. Finite perturbation theory based on CEPA-0 can be applied, provided that for all selected values of the external perturbation a Brillouin–Brueckner condition is obeyed.

Finally one should not forget that effects of dynamic correlation on properties are often small, while non-dynamical correlation effects can be very pronounced. This means that a most important step towards a treatment of correlation effects for properties is often coupled MC-SCF as discussed in Sect. 4.

Acknowledgement. The author thanks Ch. van Wüllen and J. Noga for fruitful discussions on this subject.

Appendix

Commutators in the particle-hole picture

The excitation operators a^p_q, a^{pq}_{rs} etc. satisfy the following commutation relations

$$[a^p_q, a^r_s] = \delta^r_q a^p_s - \delta^p_s a^r_q \tag{A.1a}$$

$$[a^{pq}_{rs}, a^t_u] = \delta^t_r a^{pq}_{us} + \delta^t_s a^{pq}_{ru} - \delta^p_u a^{tq}_{rs} - \delta^q_u a^{pt}_{rs}, \text{ etc.} \tag{A.1b}$$

contractions from upper right to lower left appear with a + sign, from upper left to lower right with a − sign.

Whenever one deals with the case that a particular single Slater determinant Φ plays a privileged role, it is convenient to define Φ as a 'physical vacuum' and to introduce a particle-hole formalism with respect to this vacuum. One then defines new operators

$$\tilde{a}^p_q = a^p_q - \delta^p_q n_q \tag{A.2a}$$

$$\tilde{a}^{pq}_{rs} = a^{pq}_{rs} + (-\delta^p_r a^q_s + \delta^p_s a^q_r)n_p + (-\delta^q_s a^p_s + \delta^q_r a^p_s)n_q + (-\delta^r_p \delta^s_q + \delta^r_q \delta^s_p)n_p n_q \tag{A.2b}$$

The $\tilde{a}$ operators have the interesting property that

$$\tilde{a}^p_q|\Phi\rangle = 0; \qquad \tilde{a}^{pq}_{rs}|\Phi\rangle = 0 \tag{A.3a}$$

which implies a fortiori that

$$\langle\Phi|\tilde{a}^p_q|\Phi\rangle = 0; \qquad \langle\Phi|\tilde{a}^{pq}_{rs}|\Phi\rangle = 0 \quad \text{etc.} \tag{A.3b}$$

The commutators of the $\tilde{a}$ operator differ from those of the a operators in the presence of cross contractions. Consider e.g.

$$[\tilde{a}^p_q, \tilde{a}^r_s] = [a^p_q, a^r_s] = \delta^r_q a^p_s - \delta^p_s a^r_q = \delta^r_q \tilde{a}^p_s - \delta^p_s \tilde{a}^r_q + \delta^r_q \delta^p_s (n_p - n_q) \tag{A.4a}$$

The last term of the final expression in (A.4a) is such a cross contraction. To the expectation value of a commutator only full contractions contribute, e.g.

$$\langle\Phi|[\tilde{a}^p_q, \tilde{a}^r_s]|\Phi\rangle = \delta^r_q \delta^p_s (n_p - n_q) \tag{A.5a}$$

$$\langle\Phi|[a^j_b, a^a_i|\Phi\rangle = \delta^a_b \delta^j_i \tag{A.5b}$$

$$\langle\Phi|[a^b_j, a^i_a]|\Phi\rangle = -\delta^i_j b^b_a \tag{A.5c}$$

The commutator between a one-particle and a two-particle cannot have a full contraction, hence

$$\langle\Phi|[\tilde{a}^p_q, \tilde{a}^{rs}_{tu}]|\Phi\rangle = 0 \tag{A.6}$$

The only nonvanishing expectation value of commutators between two-particle-operators are

$$\langle\Phi|[\tilde{a}^{kl}_{cd}, \tilde{a}^{ab}_{ij}]|\Phi\rangle = (\delta^k_i \delta^l_j - \delta^k_j \delta^l_i)(\delta^a_c \delta^b_d - \delta^a_d \delta^b_c) \tag{A.7a}$$

$$\langle\Phi|[\tilde{a}^{cd}_{kl}, \tilde{a}^{ij}_{ab}]|\Phi\rangle = -(\delta^i_k \delta^j_l - \delta^i_l \delta^j_k)(\delta^c_a \delta^d_b - \delta^c_b \delta^d_a) \tag{A.7b}$$

for the expectation value of double commutators one gets e.g.

$$\langle\Phi|[[\tilde{a}^p_q, \tilde{a}^r_s], \tilde{a}^t_u]|\Phi\rangle = (\delta^r_q \delta^t_s \delta^p_u - \delta^p_s \delta^t_q \delta^r_u)(n_u - n_t) \tag{A.8}$$

and the only nonvanishing matrix elements of type (A.8a) are

$$\langle\Phi|[[\tilde{a}^i_r, \tilde{a}^r_a], \tilde{a}^a_i]|\Phi\rangle = 1 \tag{A.9a}$$

$$\langle\Phi|[[\tilde{a}^r_a, \tilde{a}^i_r], \tilde{a}^a_i]|\Phi\rangle = -1 \tag{A.9b}$$

$$\langle\Phi|[[\tilde{a}^a_r, \tilde{a}^r_i], \tilde{a}^i_a]|\Phi\rangle = -1 \tag{A.9c}$$

$$\langle\Phi|[[\tilde{a}^r_i, \tilde{a}^a_r], \tilde{a}^i_a]|\Phi\rangle = 1 \tag{A.9d}$$

Such expectation values are conveniently represented by diagrams. 'Vacuum expectation value' diagrams must be completely contracted, i.e. have no external lines. They must further be connected, such that there is at least one contraction between the terms of the inner commutator, and so between the terms of the outer commutator.

For the application of the particle-hole formalism the Hamiltonian must be written in this formalism. One gets

$$H = h_q^p a_p^q + \tfrac{1}{2} g_{rs}^{pq} a_{pq}^{rs} = f_q^p \tilde{a}_p^q + \tfrac{1}{2} g_{rs}^{pq} \tilde{a}_{pq}^{rs} \tag{A.10a}$$

$$f_q^p = h_q^p + (g_{qr}^{pr} - g_{rq}^{pr}) n_p \tag{A.10b}$$

The one-electron Brillouin condition (4.3) becomes:

$$\langle \Psi_0 | [H_0, a_u^t] | \Psi_0 \rangle = \langle \Psi_0 | [H_0, \tilde{a}_u^t] | \Psi_0 \rangle = f_u^t (n_u - n_t) = 0 \tag{A.11a}$$

$$f_a^i = f_i^a = 0 \tag{A.11b}$$

The construction of the Hessean (4.12a) is now very simple, if one chooses a basis in which:

$$f_q^p = \varepsilon_p \delta_q^p \tag{A.11c}$$

$$H_{aj}^{ib} = \langle \Phi | [\tilde{a}_a^i, [H, \tilde{a}_j^b]] | \Phi \rangle = (\varepsilon_b - \varepsilon_j) \delta_a^b \delta_j^i + \tilde{g}_{aj}^{ib} = (H_{ib}^{aj})^* \tag{A.12a}$$

$$H_{ab}^{ij} = \tilde{g}_{ab}^{ij} = (H_{ij}^{ab})^* \tag{A.12b}$$

References

1. Kutzelnigg, W (1992) Theoret Chim Acta 83:263, paper I of this series
2. Löwdin PO (1955) Phys Rev 97:1474,1490,1509
3. Kutzelnigg W (1989) J Mol Struct (Theochem) 202:11
4. Dalgarno A, Stewart AL (1956) Proc Roy Soc A 238:269, 276
5. Kutzelnigg W (1982) J Chem Phys 77:3081
6. Kutzelnigg W (1989) in: Mukherjee D (ed) Aspects of many-body effects in molecules and extended systems, Lecture Notes in Chemistry, Vol 50, p 35. Springer, Berlin
7. Kutzelnigg W (1979) Chem Phys Lett 64:383
8. Kutzelnigg W (1980) Int J Quantum Chem 18:3
9. Dalgarno A (1959) Proc Roy Soc A 251:282
10. Hinze J, Roothaan CCJ (1967) Progr Theor Phys 40:37; Hinze J (1973) J Chem Phys 59:6424
11. Levy B, Berthier G (1968) Int J Quantum Chem 2:307; Levy B (1970) Chem Phys Lett 4:17; (1970) Int J Quant Chem 4:2970
12. Werner HJ (1987) Adv Chem Phys 69:1
13. Shepard R (1987) Adv Chem Phys 69:63
14. Jaszunski M, Sadlej AJ (1975) Theoret Chim Acta 40:157; Daborn GT, Handy NC (1983) Mol Phys 49:1277
15. Kutzelnigg W (1977) in: Schaefer III HF (ed) Modern theoretical chemistry, Vol 3a, p 129. Plenum, NY
16. Reitz H, Kutzelnigg W (1979) Chem Phys Letts 66:111
17. Kümmel H, Lührmann D (1972) Nucl Phys A 191:525; 194:225
18. Kutzelnigg W (1984) in: Kümmel H, Ristig L (eds) Recent progress in many-body theories, Lecture notes in physics, Vol 198. Springer, Berlin
19. Kutzelnigg W (1992) in: Mukherjee D (ed) Applied many-body methods in spectroscopy and electronic structure. Plenum, p 4
20. Kutzelnigg W (1984) J Chem Phys 80:822
21. Kutzelnigg W (1991) Theoret Chim Acta 80:349
22. Cizek J (1966) J Chem Phys 45: 4256; (1969) Adv Chem Phys 14:35
23. Bartlett RJ (1989) J Phys Chem 93:1697 and references therein
24. Meyer W (1973) J Chem Phys 58:1017
25. Pople JA, Head-Gordon M, Raghavachari K (1987) J Chem Phys 87:5968; (1989) 90:4653
26. Arponen J (1983) Ann Phys (NY) 151:311
27. Ahlrichs R, Scharf P, Ehrhard C (1985) J Chem Phys 82:890

28. Bartlett RJ, Kucharski SK, Noga J (1989) Chem Phys Lett 155:133
29. Kutzelnigg W, Smith VH (1964) J Chem Phys 41:869; Smith VH, Kutzelnigg W (1968) Ark Fys 38:309
30. Kümmel H (1979) Nucl Phys A 317:199
31. Gdanitz RJ, Ahlrichs R (1988) Chem Phys Lett 143:413
32. Monkhorst H (1977) Int J Quant Chem 11:421
33. Kelly HP (1969) Adv Chem Phys 14:129; Musher JI (1969) J Chem Phys 46:369; Doran MB (1974) J Phys B 7:558; Itajagiki T, Saika A (1979) J Chem Phys 70:2378
34. Sadlej AJ (1981) J Chem Phys 73:320
35. Caves TC, Karplus M (1969) J Chem Phys 50:3649; Nerbrandt PO, Roos B, Sadlej AJ (1979) Int J Quant Chem 15:135; Epstein ST, Sadlej AJ (1979) Int J Quantum Chem 15:147; Sadlej AJ (1983) Int J Quant Chem 23:147; Diercksen GHF, Roos B, Sadlej AJ (1981) J Chem Phys 73:1253
36. Oddershedde J, Jørgensen P, Yeager DK (1984) Comput Phys Rep 2:33; Geertsen J, Oddershedde J (1986) J Chem Phys 85:2112; Gallasso V, Fronzoni G (1986) J Chem Phys 84:3215; Fronzoni G, Galasso V (1986) Chem Phys 103:29; Bouman TD, Hansen AE (1990) Chem Phys Lett 175:292
37. Gauss J (1992) Chem Phys Lett 191:614