Restricted Hartree–Fock and unrestricted Hartree–Fock as reference states in many-body perturbation theory: a critical comparison of the two approaches

Petr Čársky[1] and Ivan Hubač[2],*

[1] J. Heyrovský Institute of Physical Chemistry and Electrochemistry, Czechoslovak Academy of Sciences, Dolejškova 3, 182 23 Prague 8, Czechoslovakia

[2] Department of Biophysics and Chemical Physics, Faculty of Mathematics and Physics, Comenius University, 842 15 Bratislava, Czechoslovakia

Summary. The paper deals with two topics related to the problem which reference state is better for many-body perturbation theory: restricted Hartree–Fock (RHF) or unrestricted Hartree–Fock (UHF)? The first topic concerns the potential surfaces. Several examples are presented to show shortcomings of the two approaches and a simple way is presented which seems to give a useful potential curve in the whole range of interatomic distances by a composition of RHF and UHF potential curves. The second topic concerns the many-body perturbation theory for open-shell systems in the RHF formalism. The method is critically examined and compared with the ordinary many-body perturbation theory using UHF as the reference. This examination of many-body techniques provides also some insight into the problems inherent of the SCF theory: spin contamination from higher multiplets, localization of orbitals, and self-consistency effects.

Key words: Many-body perturbation theory – Hartree–Fock – Potential surfaces – Open-shell systems

1. Introduction

The aim of this paper is to show some aspects of the choice of the reference state in beyond Hartree–Fock calculations. More specifically we want to show what are the consequences of using restricted Hartree–Fock (RHF) and unrestricted Hartree–Fock (UHF) as reference states in Many-Body Perturbation Theory (MBPT). Theoretically, the choice of the MO basis for highly correlated wave functions is irrelevant. In actual calculations it is however very important and as shown, for example, by Shavitt and collaborators [1], it may considerably improve the convergence of CI expansion and save much computer time. In this paper the choice of the MO basis will be restricted to canonical RHF and UHF orbitals and in the next sections we will discuss the virtues and shortcomings of

* *Present address*: Max-Planck-Institut für Physik und Astrophysik, Institut für Astrophysik, Karl-Schwarzschild-Str. 1, W-8046, Garching b. München, Federal Republic of Germany

RHF-MBPT and UHF-MBPT treatments. The second section will deal with the properties and features of RHF-MBPT and UHF-MBPT potential surfaces. The third section deals with open-shell systems and it presents our attempt to develop [2] a more economic alternative to the commonly used unrestricted Møller– Plesset [3–5] theory (UMP) which would be based on the RHF formalism. Finally, in the last section we present the simulation of the UHF states by means of the perturbed RHF wave functions. We will use it for distinguishing between true correlation effects and self-consistency effects in RHF-MBPT and related double-perturbation approaches.

## 2. RHF-MBPT and UHF-MBPT potential curves

Obtaining potential curves which represent dissociation of a bond in a molecule belongs to the most difficult tasks of applied quantum chemistry. As regards the utility of single-reference RHF and UHF beyond-Hartree-Fock methods, one may find in the older literature clear statements about the superiority of UHF treatments. Figure 1 shows a typical example.

The authors of these calculations [6] argue for the utility of the UHF approach by the following plausible statements. The different space orbitals give enough flexibility to allow the UHF (in the upper part of Fig. 1) to dissociate properly into a hydrogen and fluorine atom, whereas the RHF curve shows the characteristic breakdown at larger interatomic distances. The two lower curves in Fig. 1 were obtained from the configuration interaction (CI) calculations with all single and double replacements out of the RHF and UHF determinants, respectively. The incorrect dissociation still persists in the RHF-CI curve, whereas the UHF-CI method seems to give a useful potential curve with the reasonable values of the equilibrium distances and dissociation energy.

The same arguments were used [7, 8] for emphasizing the utility of UHF-MBPT calculations. At intermediate distances the UMP2 curves show a similar
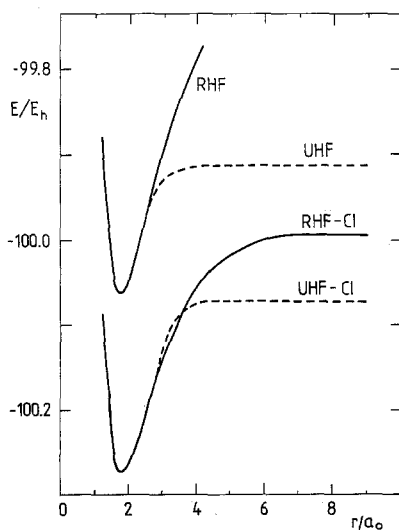

Fig. 1. Potential curves for hydrogen fluoride with the RHF and UHF methods and with the CI-SD expansions based on either an RHF or an UHF reference state [6]. The basis set used was $(11s6p\,1d/5s\,1p)/[5s3p\,1d/3s\,1p]$
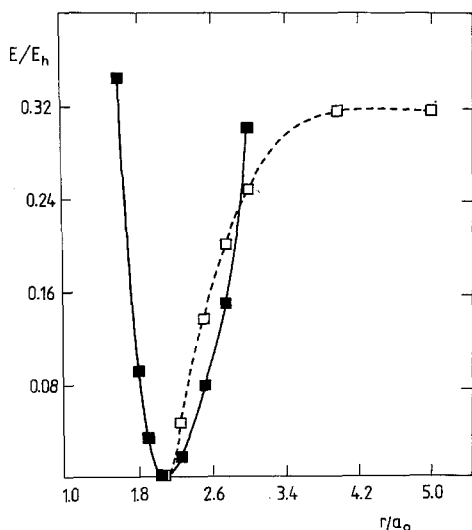
Fig. 2. Potential curves for N_2 given by MBPT for doubles through the sixth order [9, 10]. The *solid line* and *full symbols* represent the RHF-based model, the *dashed line* and *open symbols* the UHF-based model

behavior as the UHF-CI curve in Fig. 1: they usually cross the RMP2 curve and then going to the equilibrium distance the UMP2 energy is higher than RMP2 energy. As seen in Fig. 2 this effect is very distinct in the case of the nitrogen molecule reported by Bartlett and Purvis [9, 10]. Up to about 2.8 A the RHF-MBPT curve follows closely the experimental curve (not shown in Fig. 2), whereas the UHF-based model gives a too steep curve with a wrong curvature. It is assumed [7, 10] that the origin of this incorrect behavior is the spin contamination from higher multiplets. The effect of the spin contamination is even more dramatic on the potential surfaces of some radical reactions. An example in Fig. 3 is taken from paper by Sosa and Schlegel [11], which shows the energy profile for the reaction $OH + C_2H_4 \rightarrow C_2H_4OH$. Both the UHF and UMP4 curves exhibit a distinct activation barrier, which is clearly an artefact of
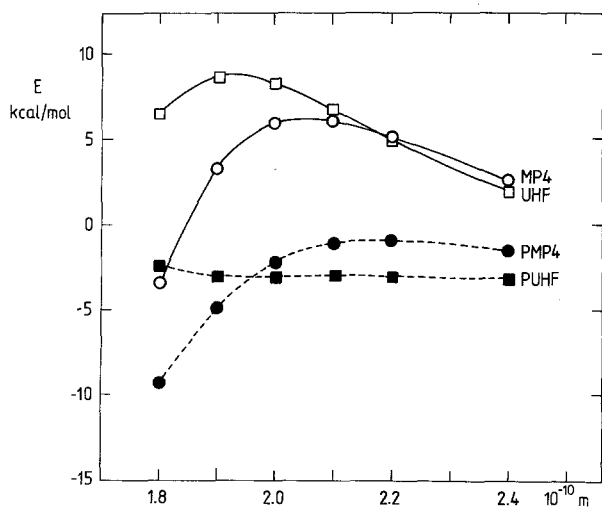

Fig. 3. Energy profile [11] for the reaction $C_2H_4 + OH \rightarrow C_2H_4OH$. The *solid lines* and *open symbols* denote unprojected UHF and UMP4 calculations, *dashed lines* and *full symbols* the projected UHF and UMP4 calculations. In all calculations the 6-31G* basis set was used

the UHF and UMP4 approaches. The other two curves in Fig. 3 were also obtained from the UHF and UMP4 calculations but the wave functions were corrected by annihilating the largest spin contaminant. Overestimation of activation barriers owing to the spin contamination was also found by Sosa and Schlegel with some other reactions [12, 13]. Their results demonstrated the utility of UHF-MBPT with the spin projection [13, 14], especially at higher orders of MBPT. The method is, however, not free of complications giving discontinuities on potential curves of bond dissociations at the point where the UHF curve coincides with RHF curve [7, 14].

The method for projected unrestricted Møller–Plesset second-order energies was developed by Knowles and Handy [46]. The method is based on annihilating of one component from the wave function and brings enhanced accuracy at low cost. The potential curves by spin-extended Hartree–Fock method were also discussed by Klimo and Tiňo [45]. More relevant for applications is probably the fact that spin-projected methods are not size extensive in a rigorous manner. Recently, the coupled cluster methods based on UHF formalism which annihilate all spin components were developed in Bartlett's group [47–49]. Scuseria [50] formulated an open-shell restricted singles and doubles coupled cluster method including perturbative triple excitation CCSD(T) that does not include any spin contamination for the correlation energies. Spin-adapted restricted linear CCSD method based on splitting of Hamiltonian (see Eq. (9)) was also developed in our group [51].

From Figs. 1 and 2 it is evident that neither RHF nor UHF approaches give realistic potential curves in the whole range of interatomic distances. RHF approaches give curves of correct shape near the equilibrium structure but fails at larger interatomic distances. On the other hand, the UHF approaches may (and mostly do) lead to correct dissociation limits but give unrealistic curves at medium- and short-range distances. A potential curve of a homolytic dissociation cannot be described by a single determinant and thus single reference-based correlation methods like MBPT fail as long as any lower orders are considered. The obvious remedy of this problem is to pass to multireference approaches. It should be realized, however, that for a somewhat larger polyatomic system such an approach becomes prohibitive, so it seems to be still topical in chemical applications to try to exploit single-reference approaches.

Duchovic and collaborators [15] tried to obtain useful potential curves by combining the merits of RHF-MBPT and UHF-MBPT approaches, viz. by linking the short-distance range of the RHF-based curve with the long-distance range of the UHF-based curve. From their data for the $CH_4 \rightarrow CH_3 + H$ dissociation presented in Fig. 4 it is seen that such a linking seemed to be straightforward for the two MP4 curves. Their composite MP4 curve is shown in Fig. 5. Duchovic et al. tested the shape of this curve by comparing it with the Morse curve. As it is seen in Fig. 5, the two curves differ significantly in the intermediate R(C–H) range. Schlegel published later [14] a revised MP4 curve for which the wave function was corrected by annihilating the spin contamination from the triplet state. His curve was in a very good agreement with the potential curve obtained by Brown and Truhlar [16] from the multireference CI-SD calculations. Obviously, the unrealistic sharp bend of the unprojected MP4 curve is due to the spin contamination, though possibly this sharp bend might be eliminated to some extent by an empirical linking of the RMP4 and UMP4 curves noted in the next paragraph. In spite of the high level of sophistication of the calculations, the projected MP4 and multireference CI-SD curves were not in
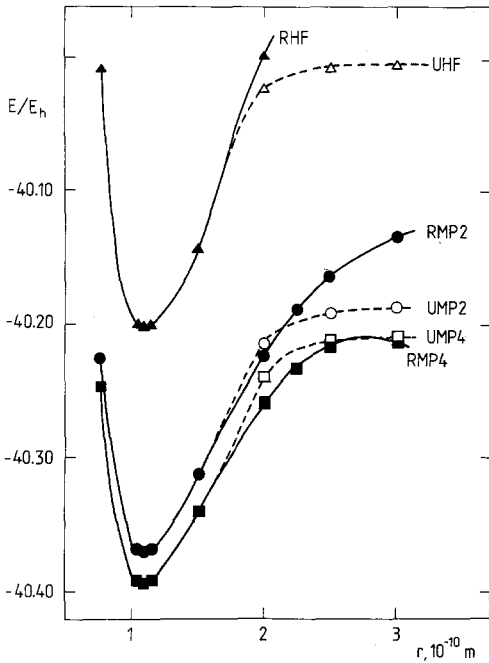
Fig. 4. Potential curves [15] for the dissociation $CH_4 \rightarrow CH_3 + H$. *Open symbols* and *dashed lines* denote UHF calculations, *full symbols* and *solid lines* RHF calculations. The basis set was 6-31G**. Only the critical CH bond length was varied; the other remaining three CH bonds were held constant at 1.986 A and all HCH angles fixed at the tetrahedral values

a satisfactory agreement with the Morse curve. The two curves had still a somewhat sharper bend than the Morse curve. This need not be, however, too much disturbing, since both spectroscopists and theorists believe jointly that the Morse curve is not a reliable standard against which the shape of potential curves for polyatomics should be tested.

We think therefore that it is preferable to judge the shape of potential curves for polyatomics by means of the following function:

$$V = V_0 + f_r y + f_{rr} y^2 + f_{rrr} y^3 + f_{rrrr} y^4 \tag{1}$$

where y is a Morse-like term:

$$y = 1 - e^{-a(r - r_0)} \tag{2}$$
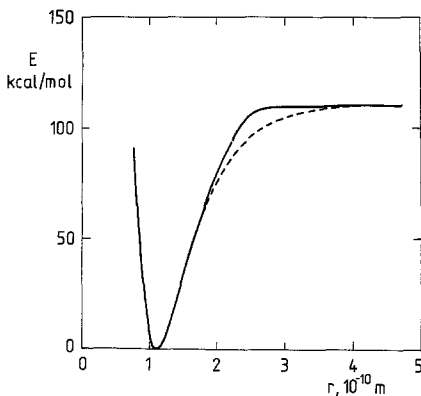
Fig. 5. Potential curves [15] for the dissociation $CH_4 \rightarrow CH_3 + H$. The *solid line* represents a composite MP4 curve which was constructed by linking the MP4 and UMP4 curves. In contrast to Fig. 4, the HCH angles were optimized at each point. The basis set used was 6-31G**. The *dashed line* is the Morse curve
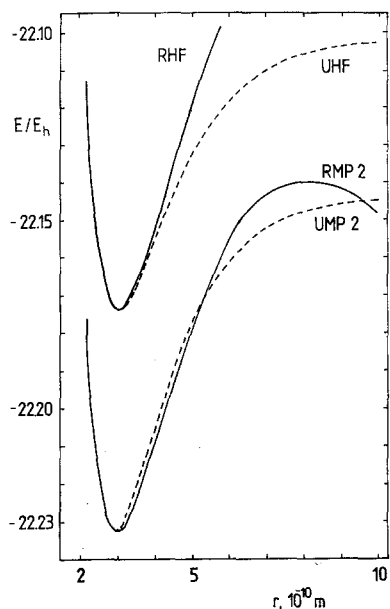
Fig. 6. Potential curves [17] for the symmetrical r(Li-Li) stretch in the D_{3h} structure of Li_3^+. The *dashed lines* correspond to UHF approaches, the *solid lines* to RHF approaches. The basis set used [18] was $(11s3p)/[6s3p]$

and a is a parameter which may be optimized together with f_r, f_{rr}, f_{rrr}, and f_{rrrr}. The typical values for a are between 0.5 and 1 and a guess to a may be taken from the Morse functions for diatomics. We have used [17] Eq. (1) for examination of RMP2 and UMP2 potential curves and for finding a justifiable combination of RMP2 and UMP2 curves which would give a realistic potential at any interatomic distance. We have used for that purpose our RMP2 and UMP2 data [17] for the symmetrical Li-Li stretch of the D_{3h} structure of Li_3^+. The respective curves, plotted in Fig. 6, show the same features as curves in Figs. 1 and 2. When Eq. (1) is applied, the RMP2 curve gives a very good fit to all points in the short and intermediate distance range but, as expected, the dissociation energy is very poor. Treating the all-computed RMP2 and UMP2 points as a single data set does not give a good fit. A very good fit is obtained, however, if the data set contains RMP2 points below the crossing point and UMP2 points above the crossing point.

3. RHF-MBPT for open shell systems

Formulation of this version of MBPT originated from our attempt to develop a more economic alternative to commonly used UHF-MBPT which is also called unrestricted Møller–Plesset (UMP) theory. Our strategy [2] was to retain all merits of the MBPT treatments of closed-shell molecules but to base the perturbation expansion on the solution given of the open-shell restricted Hartree–Fock method developed by Roothaan [19]. This method can accommodate several types of open-shell configurations but we shall restrict ourselves to the most common of them, viz., to doublet states in which the unpaired electron occupies a nondegenerate molecular orbital and to triplet states having a

half-closed-shell configuration. For such states the respective Hartree–Fock operator, f_R, becomes [19]:

$$f_R = h + \sum_i (2J_i - K_i) + \sum_m (J_m - \tfrac{1}{2}K_m) + P_T Q + Q P_T - Q \qquad (3)$$

where the operators P_T and Q are defined as:

$$Q = \sum_m K_m \qquad (4)$$

$$P_T = \sum_i |i\rangle\langle i| + \tfrac{1}{2}\sum_m |m\rangle\langle m| \qquad (5)$$

The indices i and m run over the doubly occupied and singly occupied molecular orbitals, respectively. The spinorbitals will be hereafter denoted by the uppercase characters.

Our task is now to modify the usual normal product form of the Hamiltonian [3]:

$$H = \langle \Phi_0 |H| \Phi_0 \rangle + \sum_{AB} \langle A |f| B \rangle N[X_A^+ X_B]$$

$$+ \tfrac{1}{2} \sum_{ABCD} \langle AB |v| CD \rangle N[X_A^+ X_B^+ X_D X_C] \qquad (6)$$

so that it could accommodate the f_R operator instead of the Hartree–Fock operator f. This may be achieved by the following substitution [20]:

$$f = f_R - U \qquad (7)$$

where U is a new one-particle operator defined by the difference between f_R and f:

$$U = \tfrac{1}{2} \sum_m (K_{m\alpha} - K_{m\beta}) + P_T Q + Q P_T - Q \qquad (8)$$

The subscripts α and β are used to denote the spinorbitals $m\alpha$ and $m\beta$. By substituting into Eq. (6) from Eq. (7) we obtain:

$$H = \langle \Phi_0 |H| \Phi_0 \rangle + \sum_A \varepsilon_A N[X_A^+ X_A]$$

$$+ \tfrac{1}{2} \sum_{ABCD} \langle AB |v| CD \rangle N[X_A^+ X_B^+ X_D X_C] - \sum_{AB} \langle A |u| B \rangle N[X_A^+ X_B] \qquad (9)$$

where the orbital energies ε_A and spinorbitals A, B, C, D are the eigenvalues and the eigenfunctions of the f_R operator. Now we are ready to follow closely the MBPT procedure developed for closed-shell molecules [3]. Using Paldus and Čížek's notation [3] we have:

$$K\Psi_0 = k\Psi_0 \qquad (10)$$

$$K = H - \langle \Psi_0 |H| \Psi_0 \rangle \qquad (11)$$

for the perturbed eigenvalue problem:

$$K_0 \Phi_0 = k_0 \Phi_0 \qquad (12)$$

$$K_0 = H_0 - \langle \Phi_0 |H_0| \Phi_0 \rangle \qquad (13)$$

for the unperturbed eigenvalue problem, and:

$$W = \tfrac{1}{2} \sum_{ABCD} \langle AB|v|CD \rangle N[X_A^+ X_B^+ X_D X_C]$$ (14)

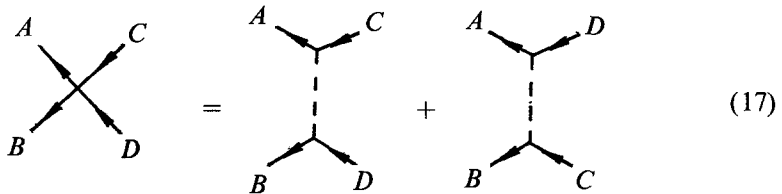for the third term on the right-hand side of Eq. (9). Comparing Eqs. (9), (11), and (14) we obtain:

$$K = K_0 + W - U$$ (15)

Hence the perturbation is represented by two operators. Conventionally we will call W the "true correlation" operator and U the "spin polarization" operator. For $(W - U)$ perturbation the Rayleigh–Schrödinger expansion through third order becomes [3]:

$$k = \langle \Phi_0 | (W - U) Q_0 (W - U) | \Phi_0 \rangle + \langle \Phi_0 | (W - U) Q_0 (W - U) Q_0 (W - U) | \Phi_0 \rangle$$ (16)
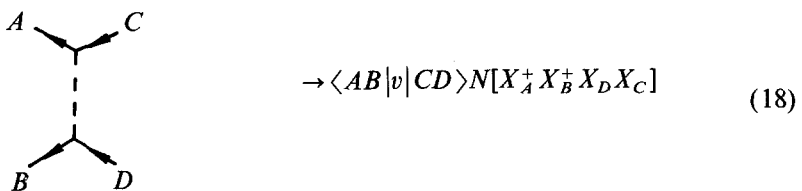
where the first term on the right-hand side of Eq. (16) represents the second-order contribution to the correlation energy $E^{(2)}$, and the second term of Eq. (16) the third-order contribution, $E^{(3)}$.

Giving a diagrammatic expression to Eq. (16) also follows the traditional scheme [3]. In addition to the expression of the W operator:

(17)
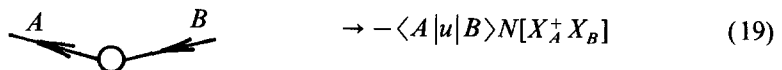
where

$$\rightarrow \langle AB|v|CD \rangle N[X_A^+ X_B^+ X_D X_C]$$ (18)

we also have to introduce the diagram for the U operator:

$$\rightarrow -\langle A|u|B \rangle N[X_A^+ X_B]$$ (19)

Following the diagrammatic rules [3, 21] we can arrive at the diagrammatic representation of Eq. (16). The respective sets of Hugenholtz and Goldstone diagrams are entered in Figs. 7 and 8. Decoding diagrams to mathematical expressions was also done by diagrammatic rules [3] generally valid in MBPT. The explicit formulas for doublet and triplet states may be found in our earlier papers [22, 23].

Shortly after our papers [2], the theory has been independently derived by McDowell [24] and Kvasnička and collaborators [25]. Wilson [26] extended it to the fourth-order by publishing all fourth-order diagrams and Kvasnička and
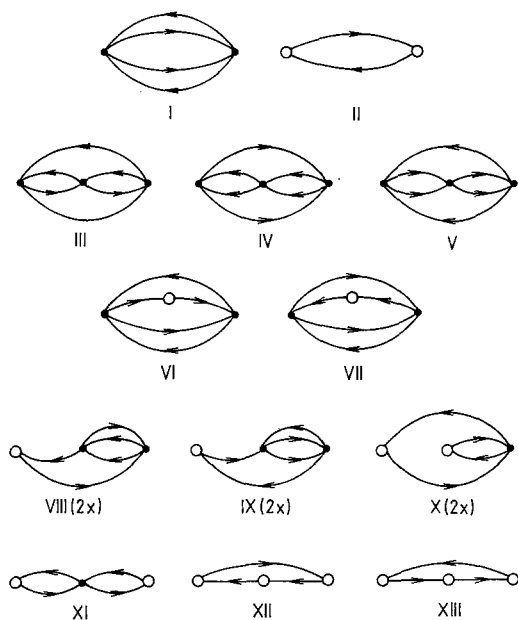
Fig. 7. Hugenholtz diagrams for the second-order (I, II) and the third-order (III–XIII) contributions to the correlation energy of half-closed-shell systems. The diagrams VIII, IX, and X can be obtained in two topologically different ways and their contributions must be therefore counted twice. The diagrams I and III–V are formally the same as those appearing in the closed-shell and unrestricted open-shell treatments [3]. The other diagrams (II, VI–XIII) are due to the additional term in the perturbation ("spin polarization" operator U in Eq. (15); in diagrams it is denoted by *open circles*)
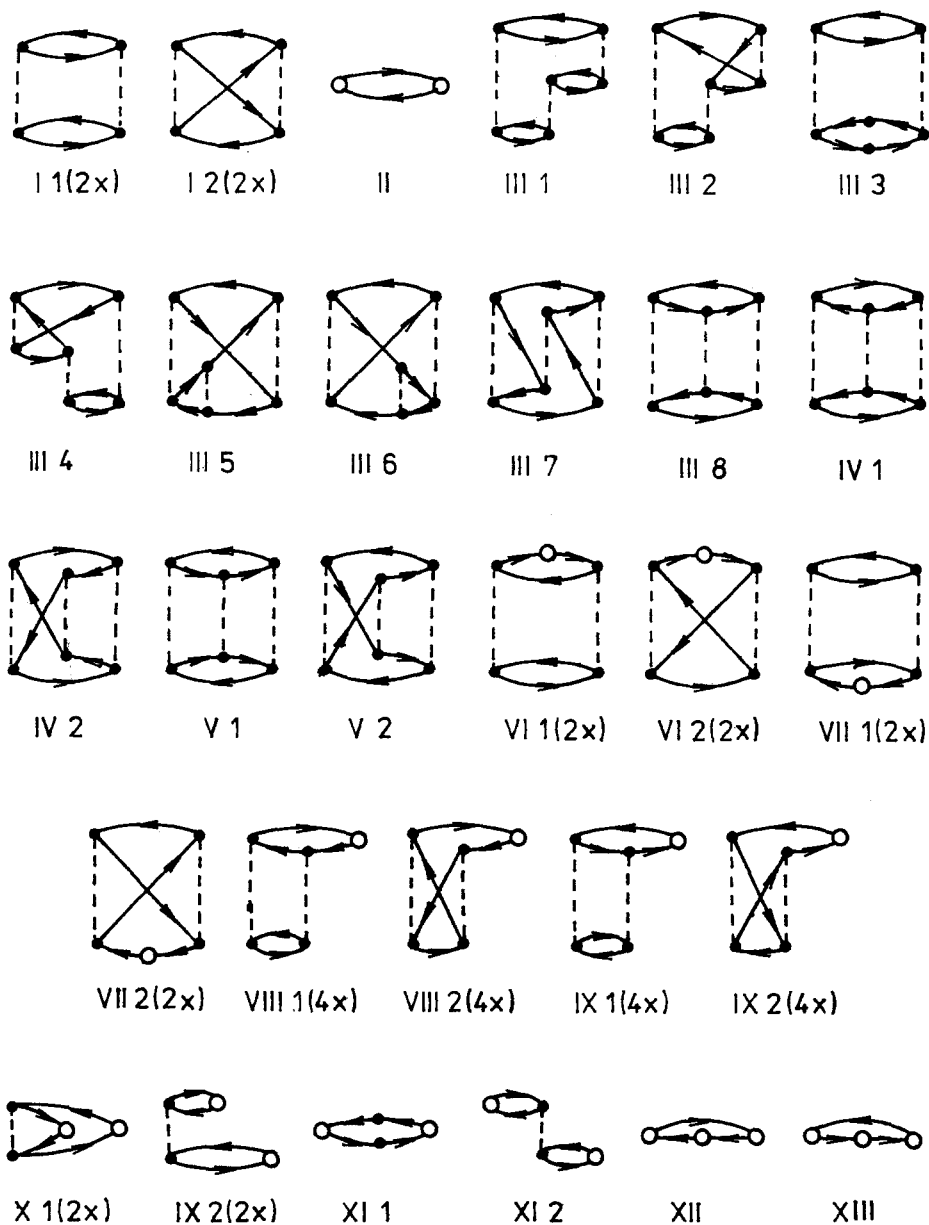
collaborators [25] formulated the theory for the linearized coupled cluster approach.

After the formulation of the theory one could have doubts about its utility for two reasons. First, the theory has been primarily developed for the antici- pated low cost of calculations. Since the number of constructed diagrams (see Figs. 7 and 8) is rather high, one might presume that this expectation need not be met. It turned out however that from the set of diagrams in Fig. 7 only the diagrams III–V are time consuming, whereas the evaluation of the rest is very fast. Practically the cost of the third-order calculation is only slightly higher than that for the ordinary third-order MBPT calculation (MP3) for a closed-shell molecule of the same size and for the same basis set. The second uncertainty concerned the convergence of the perturbation expansion. Since the perturbation now contains two terms, one might think that it would be too large and the convergence therefore rather poor. This of course had to be tested by numerical calculations.

We decided [27] therefore to run in parallel RHF-MBPT(3), CI-D, and CEPA calculations for a set of open-shell systems using the same basis set and the molecular geometry. The energies obtained are entered in Table 1. An overall observation from this table is that RHF-MBPT(3) gives somewhat more correla- tion energy than CI-D but somewhat less correlation energy than CEPA. This conforms to the experience found with closed-shell molecules. One further observes that except for CH_3, the ratio between RHF-MBPT(3) and CEPA correlation energies is nearly constant ($98.5 \pm 1\%$); but the same does not hold for the ratio between RHF-MBPT(3) and CI-D. The size inconsistency of the CI approach (if this is limited to doubles only) causes the CI-D correlation energies to become successively smaller if the size of the system is increased (compare F, FHH, HFH versus HFF). We have also computed the MP3, CI-D, and CEPA

I 1(2x) I 2(2x) II III 1 III 2 III 3

III 4 III 5 III 6 III 7 III 8 IV 1

IV 2 V 1 V 2 VI 1(2x) VI 2(2x) VII 1(2x)

VII 2(2x) VIII 1(4x) VIII 2(4x) IX 1(4x) IX 2(4x)

X 1(2x) IX 2(2x) XI 1 XI 2 XII XIII

Fig. 8. Goldstone diagrams derived from the Hugenholtz diagrams in Fig. 7. Some diagrams can be obtained in two or four topologically different ways, so that their contributions to the correlation energy must be multiplied by the indicated factors

energies for related closed-shell molecules and from the energy differences we set up Tables 2 and 3 on energies of reactions and barrier heights for a set of simple chemical reactions. The MBPT(3) data are seen to lie mostly between the CI-D and CEPA results which reflects the trends of Table 1. The RHF-MBPT(3) results obtained were promising and we were inclined to believe that the theory

Table 1a. Energies of open-shell systems. Comparison of RHF-MBPT(3), CI-D, and CEPA results

System	SCF	RHF-MBPT(3)	CI-D	CEPA
F	−99.394521	−0.14698	−0.14346	−0.14751
H_3	−1.589994	−0.05279	−0.05256	−0.05341
$BH_2\,(^2A_1)$	−25.752516	−0.08293	−0.08013	−0.08281
$NH_2\,(^2B_1)$	−55.573224	−0.16546	−0.15882	−0.16737
$NH_2\,(^2A_1)$	−55.523338	−0.16133	−0.15524	−0.16314
HFH	−100.435460	−0.22955	−0.22035	−0.23362
FHH	−100.506711	−0.18767	−0.18260	−0.19071
HFF	−199.202451	−0.40159	−0.37029	−0.41187
CH_3	−39.567926	−0.15715	−0.14452	−0.15225
CH_5	−40.655091	−0.19898	−0.18953	−0.20395

Details of calculations [27]: CI-D and CEPA were actually PNO-CI(D) and CEPA(D); the $(1s)^2$ cores were left uncorrelated; basis set used was $(9s\,5p\,1d/4s\,1p)/[4s\,2p\,1d/2s\,1p]$; all energies are in a.u., for further details see Ref. [27]

Table 1b. Percentage

System	RHF-MBPT(3)/CI-D	RHF-MBPT(3)/CEPA
F	102.4	99.6
H_3	100.4	98.8
$BH_2\,(^2A_1)$	103.5	100.1
$NH_2\,(^2B_1)$	104.2	98.9
$NH_2\,(^2A_1)$	103.9	98.9
HFH	104.2	98.3
FHH	102.8	98.4
HFF	108.5	97.5
CH_3	108.7	103.2
CH_5	105.0	97.6

Table 2. Computed energies of reactions (in kJ/mol)[a]

Reaction	SCF	RHF-MBPT(3)	CI-D	CEPA
$2H \rightarrow H_2$	−347.9	−436.0	−440.9	−440.9
$2F \rightarrow F_2$	+154.5	−84.6	−33.6	−101.9
$H + F \rightarrow HF$	−404.5	−541.1	−533.5	−543.1
$F + H_2 \rightarrow FH + H$	−56.6	−105.1	−92.6	−102.2
$H + F_2 \rightarrow HF + F$	−559.0	−456.5	−500.3	−441.6
$H + CH_4 \rightarrow CH_3 + H_2$	+19.8	−8.7	+3.3	+11.1

[a] Obtained from the entries of Table 1 and additional calculations on closed-shell species. For details see Ref. [27]

would be free of complications in chemical applications. Later, however, we have found a serious failure of the theory for oxygen molecule [28]. The O_2 results together with the results for the methylene molecule are presented in Tables 4 and 5. Whereas the RHF-MBPT data for CH_2 exhibit the same trends as with other open-shell systems listed in Table 1, the situation with the O_2 molecule is

Table 3. Computed barrier heights of some radical reactions (in kJ/mol)[a]

Reaction	SCF	RHF-MBPT(3)	CI-D	CEPA
$H + H_2 \rightarrow H_2 + H$	106.3	52.9	58.3	56.1
$H + FH \rightarrow HF + H$	293.2	213.0	220.0	205.7
$F + H_2 \rightarrow FH + H$	49.6	30.8	39.9	29.2
$H + F_2 \rightarrow HF + F$	71.0	27.5	39.3	19.8
$H + CH_4 \rightarrow CH_3 + H_2$	135.4	82.3	90.7	80.9

[a] Obtained from the entries of Table 1 and additional calculations on closed-shell species. For details see Ref. [27]

Table 4. Comparison of correlation energies given by RHF-MBPT and CI-SD calculations[a] for the molecules CH_2 and O_2

Molecule	SCF	RHF-MBPT(2)	RHF-MBPT(3)	CI-SD
CH_2 ($^3B_1$)	-38.9279	-0.1144	-0.1154	-0.1137
O_2 ($^3\Sigma_g$)	-149.633213	-0.49513	-0.27100	-0.34451

[a] The $1a_1$ orbital in CH_2 was kept doubly occupied; the basis set used was $(9s5p1d/4s1p)/[4s2p1d/2s1p]$, all energies are in a.u., for additional details see Ref. [28]

Table 5. Contributions from diagrams I–XIII to the RHF-MBPT(3) correlation energy[a] of CH_2 and O_2

	Correlation energy (a.u.)	
Diagram	CH_2	O_2
I	-0.11120	-0.47938
II	-0.00319	-0.01575
III	-0.07895	-0.31446
IV	0.02992	0.16427
V	0.02529	0.15886
VI	0.01402	0.07576
VII	0.00827	0.11712
VIII	0.00111	0.01465
IX	0.00029	0.01174
X	-0.00020	-0.00113
XI	-0.00129	-0.00624
XII	0.00022	0.00104
XIII	0.00033	0.00251

[a] See footnote in Table 4

different. Compared to CI-SD, the second-order contribution is too large. The third-order contribution is of opposite sign, but it is also overestimated, so that the resulting total correlation energy at the third order is too low. Obviously, RHF-MBPT converges poorly in this case. Different behavior of CH_2 and O_2 systems may be understood from the inspection of contributions of individual diagrams I–XIII.

As seen in Table 5, the critical point for the poor convergence with O_2 is a large contribution from diagrams VI and VII which contain the U vertex. This suggests that the part of the perturbation which is due to the "spin polarization" operator U is very important for the oxygen molecule. The overall perturbation $W - U$ is likely to be too large, which is a probable reason for the poor convergence of the RHF-MBPT expansion. Our speculation was [28] that such a poor convergence might be expected with any open-shell molecule in which the open shell shares a common space with occupied and/or virtual orbitals, in contrast to radicals with a localized unpaired electron for which small spin polarization and good RHF-MBPT convergence may be expected. Our assumption was supported by a systematic investigation by Balková [29], who showed in a convincing manner that RHF-MBPT(3) works considerably better for open-shell species with localized unpaired electrons than for open-shell species with delocalized unpaired electrons. We present some of her data in Tables 6 and 7 in order to use them for drawing some practical conclusions of this section. It is seen from the two tables that the difference in SCF energies $E_{RHF} - E_{UHF}$ may be taken as a rough measure of the spin polarization and the convergence of the

Table 6. RHF and UHF SCF energies[a]

System	RHF	UHF	$\langle S^2 \rangle$	RHF-UHF
CH	−38.2614	−38.2649	0.756	0.00355
H_2NO	−130.3720	−130.3790	0.762	0.00698
CH_3O	−114.4146	−114.4190	0.758	0.00435
HCO	−113.2421	−113.2464	0.765	0.00433
HO_2	−150.1579	−150.1633	0.763	0.00545
CH_3	−39.5547	−39.5590	0.762	0.00428
CN	−92.1847	−92.2039	1.090	0.01927
O_2	−149.5982	−149.6179	2.030	0.01971
CH_2	−38.9165	−38.9215	2.015	0.00503

[a] Taken from Ref. [29], 6-31G* basis set was used, all energies are in a.u.

Table 7. Comparison of UHF-MBPT and RHF-MBPT expansions for open shell systems[a]

System	MBPT(2)			MBPT(3)		
	RHF	UHF	RHF-UHF	RHF	UHF	RHF-UHF
CH	−0.0870	−0.0737	−0.0133	−0.0976	−0.0901	−0.0075
H_2NO	−0.3588	−0.3228	−0.0360	−0.3352	−0.3344	−0.0008
CH_3O	−0.2905	−0.2649	−0.0256	−0.2966	−0.2870	−0.0096
HCO	−0.3219	−0.2855	−0.0364	−0.2909	−0.2864	−0.0045
HO_2	−0.3625	−0.3338	−0.0387	−0.3434	−0.3419	−0.0015
CH_3	−0.1233	−0.1097	−0.0136	−0.1319	−0.1255	−0.0064
CN	−0.3083	−0.2226	−0.0857	−0.2685	−0.2315	−0.0370
O_2	−0.4557	−0.3234	−0.1323	−0.2772	−0.3219	+0.0447
CH_2	−0.1016	−0.0818	−0.0198	−0.1033	−0.0957	−0.0076

[a] Taken from Ref. [29]; compatible with data of Table 6

$E_{RHF} - E_{UHF}$ difference is extraordinarily high. The situation is of course compli-
cated by a possible contamination of higher multiplets in the UHF solution, so
one should also take in consideration the S^2 value. With the oxygen molecule the
S^2 value is close to that for a pure triplet state and the $E_{RHF} - E_{UHF}$ difference
is likely to be solely due to the spin polarization. On the other hand, with the
cyano radical the S^2 value of 1.09 is very different from the value of 0.75 for a
pure doublet state and, as will be shown in the next section, the spin polarization
is small in this case. We have attempted to formulate the effect of S^2 and
$E_{RHF} - E_{UHF}$ in a compact form in Table 8. As indicated in this table it is
profitable to know which part of the $E_{RHF} - E_{UHF}$ difference is due to the spin
polarization and which part is due to the contamination of higher multiplets. In
the following section we will show how the spin polarization contribution may be
simply estimated by expressing the UHF wave function as a perturbed RHF
wave function.

Finally, some other double-perturbation theories, related to our RHF-
MBPT, should be noted. When studying the system in an external field one can
use two different perturbation methods for the calculation of the external
perturbation effects [30]. The so-called uncoupled Hartree–Fock (UCHF)
[31, 32] perturbation theory has exactly the same structure as our RHF-MBPT
treatment. The operator U in this case corresponds to the presence of external
field. The spinorbital basis on which we define creation and annihilation opera-
tors (Eq. (9)) is determined in the field of the unperturbed Hartree–Fock
potential. The UCHF-based perturbation method has the structure of the
double-perturbation theory and it permits us to distinguish between pure two-
particle (correlation) effects and pure one-particle (self-consistency) effects [30].
On the other hand, in the coupled Hartree–Fock (CHF)-based perturbation
theory [33–36] the spinorbital basis is determined in the presence of the external
field. Formally, therefore, the CHF-based perturbation theory resembles the
UHF-MBPT treatment. There is, however, an important difference. Whereas the

Table 8. Recommendation for using RHF and UHF approches in MBPT(3) treatments of open-shell
systems

$E_{RHF} - E_{UHF}$	$\langle S^2_{UHF} \rangle - \langle S^2_{pure} \rangle^a$	Notes on the use of RHF-MBPT and UHF-MBPT
small	small	Either approach should work; RHF-MBPT is more economic, however.
small	large	This combination is unlikely. If it still occurs, try RHF-MBPT.
large	small	Use preferably UHF-MBPT.
large	large	Avoid using UHF-MBPT. Test the spin polarization (see Sect. 4): if it is small use RHF-MBPT; if it is large user other method than MBPT.

[a] By $\langle S^2_{pure} \rangle$ we mean 0.75 for doublets and 2.0 for triplets

CHF approach is clearly superior to the UCHF approach, use of UHF-MBPT instead of RHF-MBPT seems to be profitable only in exceptional cases of very large self-consistency effects (spin polarization).

Another example of application of double-perturbation theory is the MBPT using noncanonical [24] (e.g., V^{N-1} potential) and localized [37, 38] orbitals. In the noncanonical orbital formulation of the perturbation theory the extra one-particle term corresponds to the change in self-consistency effects. In the localized orbital based perturbation theory the extra one-particle perturbation is the consequence of unitary transformation performed on the occupied and virtual single-particle functions. In this case the nonzero off-diagonal elements of the Hartree–Fock operator enter the perturbation expansion. Since the unitary transformation does not change the total energy and the wave function of the system, the extra one-particle perturbation does not correspond to self-consistency effects. Therefore the number of double-perturbation diagrams is substantially reduced. This is due to the fact that (using our terminology) all diagrams containing the one-particle extra term U are vanishing, unless the entering and leaving lines of the U vertex are two hole lines or two particle lines. Thus the diagrams II and VIII–XIII (see Fig. 7) having the U vertex linked with one hole and particle line give zero contributions in this approach and the only nonvanishing "localized" diagrams through third order are the diagrams [38] VI and VII.

As regards the applications the external field perturbation theories have already been developed to a state of routine calculations [9, 41] whereas the localized orbital-based perturbation treatments have been the subject of intensive studies [39, 40, 53–55]. An important contribution in this field was done by Saebo and Pulay [53–55].

An interesting method based on Møller–Plesset partitioning of Hamiltonian for open-shell and multiconfigurational SCF reference states was developed by Wolinski, Sellers, and Pulay [52]. They followed the idea which they used in formulating the many-body perturbation theory for localized orbitals [53–55]. The basic idea by which they differ from our approach [2] is that they use nondiagonal zeroth-order Hamiltonian. The reference function accounts for nondynamical electron correlation and is an eigenfunction of S^2 and S_z. The poor convergence of unrestricted Møller–Plesset perturbation theory was studied by Gill, Pople, and Radom [56]. They showed that the spin contamination and poor UMP convergence are closely related. Our conclusion is that the current problem is to formulate the spin-adapted ROHF CC method where the important steps were already done in the work of Bartlett [47–49, 57], Scuseria [50] and also in our group [51].

4. Simulation of the UHF states by means of the perturbed RHF wave functions

In order to obtain deeper insight into the problems of RHF-MBPT, we decided to undertake a study [42] of spin polarization effects, which may be called more generally self-consistency effects, separately from correlation effects. The idea was not new and since Musher's paper [43] the problem of self-consistency effects was treated in the literature several times. Probably the most elaborate treatment was given by Rossky and Karplus [44]. Whereas the essence of their treatment is a perturbation expansion of UHF spinorbitals through RHF spinorbitals, we preferred a more transparent approach based on the perturba-

tion expansion for the UHF Slater determinant. We disregarded the fact that the UHF wave function is not the eigenfunction of the S^2 operator. Using this simplification, one may assume for the UHF wave function the following first-order expansion:

$$\Phi_0^{\text{``UHF''}} = \Phi_0^{\text{RHF}} + Q_0 U \Phi_0^{\text{RHF}} \qquad (20)$$

where Q_0 and U have the same meaning as in Eq. (16). Since Eq. (20) implies the use of the intermediate overlap:

$$\langle \Phi_0^{\text{RHF}} | \Phi_0^{\text{``UHF''}} \rangle = 1 \qquad (21)$$

the simulated UHF energy is given by:

$$E_{\text{``UHF''}} = \frac{\langle \Phi_0^{\text{``UHF''}} | H | \Phi_0^{\text{``UHF''}} \rangle}{\langle \Phi_0^{\text{``UHF''}} | \Phi_0^{\text{``UHF''}} \rangle} \qquad (22)$$

The self-consistency effects are then represented by the difference $E_{\text{``UHF''}} - E_{\text{RHF}}$, where $E_{\text{``UHF''}}$ is given by Eq. (22) and E_{RHF} is simply $\langle \Phi_0^{\text{RHF}} | H | \Phi_0^{\text{RHF}} \rangle$. Substitution of Eq. (20) into the numerator of Eq. (22) gives us:

$$\langle \Phi_0^{\text{``UHF''}} | H | \Phi_0^{\text{``UHF''}} \rangle = \langle \Phi_0^{\text{RHF}} + \Phi_0^{\text{RHF}} U Q_0 | H | \Phi_0^{\text{RHF}} + Q_0 U \Phi_0^{\text{RHF}} \rangle \qquad (23)$$

Equation (23) may be rewritten as:

$$\langle \Phi_0^{\text{``UHF''}} | H | \Phi_0^{\text{``UHF''}} \rangle = \langle \Phi_0^{\text{RHF}} | H | \Phi_0^{\text{RHF}} \rangle + 2 \langle \Phi_0^{\text{RHF}} U Q_0 | H | \Phi_0^{\text{RHF}} \rangle$$
$$+ \langle \Phi_0^{\text{RHF}} U Q_0 | H | Q_0 U \Phi_0^{\text{RHF}} \rangle \qquad (24)$$

and by substituting H from Eq. (9), it becomes ready for the diagrammatical treatment [3]. After some manipulation [39] we get it in the form:

$$\langle \Phi_0^{\text{``UHF''}} | H | \Phi_0^{\text{``UHF''}} \rangle = \langle \Phi_0^{\text{RHF}} | H | \Phi_0^{\text{RHF}} \rangle S^{(1)} + II + XI + XII + XIII \qquad (25)$$

where $S^{(1)}$ is the overlap of the first-order wave function:

$$\langle \Phi_0^{\text{``UHF''}} | H | \Phi_0^{\text{``UHF''}} \rangle = S^{(1)} \qquad (26)$$

The self-consistency effects may therefore be expressed from Eqs. (22), (25), and (26) as:

$$\frac{\langle \Phi_0^{\text{``UHF''}} | H | \Phi_0^{\text{``UHF''}} \rangle}{\langle \Phi_0^{\text{``UHF''}} | \Phi_0^{\text{``UHF''}} \rangle} - \langle \Phi_0^{\text{RHF}} | H | \Phi_0^{\text{RHF}} \rangle = \frac{1}{S^{(1)}} (II + XI + XII + XIII) \qquad (27)$$

Balková [29] extended our treatment for the second-order contributions by assuming:

$$\Phi_0^{\text{``UHF''}} = \Phi_0^{\text{RHF}} + Q_0 U \Phi_0^{\text{RHF}} + Q_0 (U - \Delta e) Q_0 U \Phi_0^{\text{RHF}} \qquad (28)$$
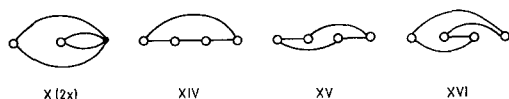
Since

$$\Delta e = \langle \Phi_0^{\text{RHF}} | U | \Phi_0^{\text{RHF}} \rangle = 0 \qquad (29)$$

Eq. (28) becomes:

$$\Phi_0^{\text{``UHF''}} = \Phi_0^{\text{RHF}} + Q_0 U \Phi_0^{\text{RHF}} + Q_0 U Q_0 U \Phi_0^{\text{RHF}} \qquad (30)$$

Fig. 9. Most important second-order "self-consistency" diagrams resulting from the perturbation expansion of the UHF wave function (Eqs. (30) and (31)). Next 18 diagrams obtained [29] and not shown in this figure are assumed to give very small contributions. As in Fig. 7, diagram X must be counted twice. The meaning of vertices is the same as in Fig. 7

By substituting for $\Phi_0^{\text{"UHF"}}$ in Eq. (22) she obtained:

$$\frac{\langle \Phi_0^{\text{"UHF"}} | H | \Phi_0^{\text{"UHF"}} \rangle}{\langle \Phi_0^{\text{"UHF"}} | \Phi_0^{\text{"UHF"}} \rangle} - \langle \Phi_0^{\text{RHF}} | H | \Phi_0^{\text{RHF}} \rangle$$

$$= \frac{1}{S^{(2)}} (II + XI + XII + XIII + X + XIV$$

$$+ XV + XVI + 18 \text{ disregarded diagrams}) \tag{31}$$

The disregarded diagrams are expected to give very small contributions. They contain five U vertices, four U vertices and one W vertex, or three U vertices and two W vertices and they would appear in the fourth order (together with diagrams XIV–XVI, see Fig. 9) and the fifth order of the RHF-MBPT theory presented in Sect. 3. The overlap integrals $S^{(1)}$ and $S^{(2)}$ in Eqs. (27) and (31) are given by [29, 42]:

$$S^{(1)} = 1 + II^* \tag{32}$$

$$S^{(2)} = 1 + II^* + 2(XII^* + XIII^*) + XIV^* + XV^* + XVI^* \tag{33}$$

where the asterisks mean that the denominators of the diagrams II and XII–XVI in Figs. 6 and 8 are squared.

Table 9. Simulation of the UHF energy by means of the perturbation expansion of the UHF wave function. All entries are in a.u. and were obtained [29] with the 6-31G* basis set

System	$E_{\text{RHF}} - E_{\text{UHF}}$	ΔS^2 [a]	self-consistency effect		
			1. order Eq. (27)	1. + 2. order Eq. (31)	perc.[b]
CH	0.00355	0.006	−0.00317	−0.00329	92.6
H_2NO	0.00698	0.012	−0.00593	−0.00632	90.6
CH_3O	0.00435	0.008	−0.00374	−0.00390	89.6
HCO	0.00433	0.015	−0.00361	−0.00382	88.4
HO_2	0.00545	0.013	−0.00452	−0.00474	86.9
CH_3	0.00428	0.012	−0.00339	−0.00356	83.3
CN	0.01927	0.340	−0.00399	−0.00456	23.7
O_2	0.01971	0.030	−0.01683	−0.01909	96.9
CH_2	0.00503	0.015	−0.00387	−0.00411	81.8

[a] ΔS^2 is difference $\langle S^2 \rangle - \langle S_{\text{pure}}^2 \rangle$, by $\langle S_{\text{pure}}^2 \rangle$ we mean 0.75 for doubles, 2.0 for triplets
[b] Absolute values and percentage with respect to the SCF result $E_{\text{RHF}} - E_{\text{UHF}}$

Some numerical data obtained [29] with Eqs. (27) and (31) are summarized in Table 9. We ordered the systems treated according to the decreasing percentage of the recovered energy difference $E_{RHF} - E_{UHF}$.

With the doublet states it is clearly seen in Table 9 that if the spin contamination of higher multiplets is small (S^2 is close to 0.75), Eq. (31) gives about 90% of the energy difference $E_{RHF} - E_{UHF}$. As the spin contamination becomes larger, the self-consistency effects are getting less and less important and in the extreme case of the cyano radical, Eq. (31) gives only 24% of the $E_{RHF} - E_{UHF}$ difference. With the triplet states (presented at the bottom of Table 9) we do not see such a clear-cut trend because of a small variation in S^2.

Equation (31) also permits us to make a classification of diagrams appearing in the RHF-MBPT(3) theory. The diagrams II and V–XIII may be called "spin polarization" or "self-consistency" diagrams and the rest of Fig. 7 may be referred to as the "correlation" diagrams.

References

1. Shavitt I, Rosenberg BJ, Palalikit S (1976) Int J Quant Chem Symp 10:33
2. Hubač I, Čársky P (1980) Phys Rev 22A:2392
3. Paldus J, Čížek J (1975) Adv Quantum Chem 9:105
4. Bartlett RJ, Silver DM (1976) in: Calais JL, Goscinski O, Linderberg J, Öhrn Y (eds) Quantum science, Plenum Press, New York, p 393
5. Pople JA, Binkley JS, Seeger R (1976) Int J Quant Chem Symp 10:1
6. Roos BO, Siegbahn PEM (1977) in: Schaefer III HF (ed) Modern theoretical chemistry Vol 3, Plenum Press, New York, p 277
7. Klimo V, Tiňo J (1980) Mol Phys 41:483
8. Klimo V, Tiňo J (1980) Mol Phys 46:541
9. Bartlett RJ, Purvis GD (1980) Phys Scripta 21:255
10. Bartlett RJ (1981) Ann Rev Phys Chem 32:359
11. Sosa C, Schlegel HB (1987) J Am Chem Soc 109:7007
12. Sosa C, Schlegel HB (1986) Int J Quant Chem 29:1001
13. Sosa C, Schlegel HB (1987) Int J Quant Chem 21:267
14. Schlegel HB (1986) J Chem Phys 84:4530
15. Duchovic RJ, Hase WL, Schlegel HB, Frisch MJ, Raghavachari K (1982) Chem Phys Lett 89:120
16. Brown FB, Truhlar DG (1985) Chem Phys Lett 113:441
17. Čársky P, Špirko V (1986) unpublished results
18. Gerber WH, Schumacher E (1978) J Chem Phys 69:1692
19. Roothaan CCJ (1960) Rev Mod Phys 32:179
20. Cederbaum LS, Schirmer J (1974) Z Phys 271:221
21. Hubač I, Čársky P (1978) Top Curr Chem 75:97
22. Čársky P, Zahradník R, Hubač I, Urban V, Kellö V (1980) Theor Chim Acta 56:315
23. Čársky P, Hubač I (1981) Coll Czech Chem Commun 46:1324
24. McDowell K (1981) Int J Quantum Chem 19:271
25. Kvasnička V, Biskupič S, Laurinc V (1981) Mol Phys 42:1345
26. Wilson S (1982) Theor Chim Acta 61:343
27. Čársky P, Hubač I, Staemmler V (1982) Theor Chim Acta 60:445
28. Čársky P, Svrček M, Hubač I, Brown FB, Shavitt I (1982) Chem Phys Lett 85:17
29. Balková A (1985) Thesis, Slovak Academy of Sciences and Comenius University, Bratislava
30. Hubač I, Sadlej AJ (1982) Czech J Phys B32:1116
31. Kelly HP (1969) Adv Chem Phys 14:129
32. Tuan DFT, Epstein ST, Hirschfelder JO (1966) J Chem Phys 44:431
33. Adamowicz L, Sadlej AJ (1978) Chem Phys Lett 53:337

34. Sadlej AJ (1979) J Phys Chem 83:1653
35. Sadlej AJ (1980) Acta Phys Polon A57:879
36. Bartlett RJ, Purvis III GD (1979) Phys Rev 20A:1313
37. Diner S, Malrieu JP, Claverie P (1969) Theor Chim Acta 13;1, 18
38. Kapuy E, Csépes Z, Kozmutza C (1983) Int J Quant Chem 23:981
39. Čížek J, Förner W, Ladik J (1983) Theor Chim Acta 64:107
40. Laidig WD, Purvis GD, Bartlett RJ (1985) J Phys Chem 89:2161
41. Diercksen GHF, Sadlej JA (1981) J Chem Phys 75:1253
42. Hubač I, Čársky P (1983) Int J Quant Chem 24:141
43. Musher JI (1967) J Chem Phys 46:369
44. Rossky PJ, Karplus M (1980) J Chem Phys 72:6085
45. Klimo V, Tiňo J (1984) Int J Quant Chem 25:733
46. Knowles PJ, Handy NC (1988) J Chem Phys 88:6991
47. Rittby M, Bartlett RJ (1988) J Phys Chem 92:3033
48. Watts J, Bartlett RJ (1990) J Chem Phys 93:6104
49. Sekino H, Bartlett RJ (1986) J Chem Physe 85:3945
50. Scuseria GE (1991) Chem Phys Lett 176:28
51. Neogrády P, Urban M, Hubač I (to be published)
52. Wolinski K, Sellers HL, Pulay P (1987) Chem Phys Lett 140:225
53. Saebo S, Pulay P (1987) J Chem Phys 86:914
54. Pulay P, Saebo S (1986) in: Jorgensen P, Simons J (eds) Geometrical derivatives of energy surfaces and molecular properties. Reidel, Dordrecht, p 95
55. Pulay P, Saebo S (1986) Theoret Chim Acta 69:357
56. Gill PMW, Pople JA, Radom L (1988) J Chem Phys 89:7307
57. Purvis III GD, Sekino H, Bartlett RJ (1988) Coll Czech Chem Com 53:2203