The Møller–Plesset perturbation revisited: origin of high-order divergences

Jean-Paul Malrieu & Celestino Angeli

Taylor & Francis
Taylor & Francis Group

INVITED ARTICLE

The Møller–Plesset perturbation revisited: origin of high-order divergences

Jean-Paul Malrieu[a],[*] and Celestino Angeli[b]

[a]*Laboratoire de Chimie et Physique Quantiques, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex, France;*
[b]*Università di Ferrara, Dipartimento di Scienze Chimiche e Farmaceutiche, Via Borsari 46, I-44121 Ferrara, Italy*

Contrarily to what happens with the Epstein–Nesbet (EN) zeroth-order Hamiltonian, the Møller–Plesset (MP) perturbation operator has diagonal matrix elements, the expression of which is recalled. It is a balance between hole–hole and particle–particle repulsions on one hand and of hole–particle attractions on the other hand. For the double excitations, which dominate the correlation effects, the attractive terms prevail and the second-order MP energy is underestimated, at least for atoms of the first rows of the periodic table. It will be shown that when the perturbation expansion reaches multiple excitations, the diagonal terms of the MP perturbation operator may become larger than the zeroth-order MP excitation energy and creates an oscillating divergence of the series. Several situations of this type will be presented. This divergence is linked to the non-additivity of excitation energies, while this additivity is an underlying assumption for the linked cluster theorem and the coupled cluster method. This analysis may also explain why for heavy atoms the second-order MP energies overshoot the exact correlation energies.

Keywords: Møller–Plesset perturbation theory; Epstein–Nesbet perturbation theory; divergences in perturbation theory; linked cluster theorem; coupled cluster

1. Introduction

Perturbation Theory is the key tool to understand the physics of the quantum many-body problem. It has led to the diagrammatic techniques [1,2] and the fundamental linked cluster theorem. From this comprehension, the coupled-cluster (CC) method [3–6] could be conceived. The theorem rests on a perturbative expansion where the exact Hamiltonian $\hat{\mathcal{H}}$ is partitioned in a zeroth-order $\hat{\mathcal{H}}_0$, chosen to give an independent-particle description of the system, and a perturbation operator $\hat{\mathcal{V}}$,

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \hat{\mathcal{V}}, \tag{1}$$

which allows one to build an order-by-order expansion of the correlated wave function and of the energy. In the most successful partition, $\hat{\mathcal{H}}_0$ is the sum of the Fock operator for the various electrons and it leads to the so-called Møller–Plesset [7] perturbation theory (MP).

Other definitions of the zeroth-order Hamiltonian are conceivable, the most natural of which, called Epstein–Nesbet [8–10], may be defined as the trace of the Hamiltonian in the basis of the single determinants or as the occupation conserving part of the Hamiltonian. In 2000, we have proposed a partition of the Hamiltonian which is in some way intermediate between MP and EN [11]. The advantages and drawbacks of the MP and EN Hamiltonians will be briefly recalled. This work first focuses on the dif-

ference between the low orders of the two series. It recalls why, at least for normal light elements, atoms or molecules, the Møller–Plesset second-order energies are in general underestimated in absolute value, due to an overestimation of the excitation energies. But it points out that this underestimation of the second order corrections is not uniform, some doubly excited configurations may be of much higher energies than estimated by the Møller–Plesset $\hat{\mathcal{H}}_0$. The paper shows a condition for which the MP perturbation series will necessarily exhibit a diverging oscillatory behaviour, at least at high orders of perturbation expansion, as it has been observed [12–14]. Then, we discuss the situations in which this condition may be fulfilled, and we show that they are not exceptional and that in most molecular problems such a high-order divergence may be expected, at least if one works with localised orbitals, as necessary to have linear scaling computational cost. On the contrary, the EN expansion has no reason to exhibit such a behaviour.

2. Choice of the zeroth-order Hamiltonian and nature of the diagonal matrix elements of the Møller–Plesset perturbation operator

The zeroth-order wave function is a single determinant Φ_0, obtained by a preliminary energy minimisation, in general through the Hartree–Fock self-consistent field (SCF) energy minimisation process. It is an antisymmetrised

*Corresponding author. Email: jean-paul.malrieu@irsamc.ups-tlse.fr

product of N ($=2n$ for a closed shell) occupied spin orbitals $1, \bar{1}, \ldots, i, \bar{i}, j, \bar{j}, \ldots, n, \bar{n}$

$$\Phi_0 = \left\| \prod_i^n i\bar{i} \right\|. \qquad (2)$$

In second-quantisation Φ_0 can be equivalently defined as

$$\Phi_0 = \prod_{i=1}^n a_i^\dagger a_{\bar{i}}^\dagger |0\rangle \qquad (3)$$

with $|0\rangle$ indicating the vacuum state. An energy ε_i is attributed to each of these occupied MOs, usually the mean value of the Fock operator for this orbital

$$\varepsilon_i = \langle i|\hat{\mathcal{F}}|i\rangle = \langle \bar{i}|\hat{\mathcal{F}}|\bar{i}\rangle = \langle i|\hat{h}|i\rangle + \sum_{j=1}^n \left[2\langle ij|r_{12}^{-1}|ij\rangle \right.$$
$$\left. - \langle ij|r_{12}^{-1}|ji\rangle \right]. \qquad (4)$$

The monoelectronic energies might be defined differently and the diagrammatic expansion, as well as the linked cluster theorem, would remain valid. In particular the canonicity of the MOs (i.e. the nullity of the off-diagonal elements of the Fock operator) is not compulsory to establish these two fundamental properties. If the size M of the basis set is larger than n, one gets $M - n$ unoccupied or virtual MOs, orthogonal to the occupied ones, and hereafter labeled r, s, t, $\ldots$. One may define energies of these virtual orbitals, for instance again as mean values of the ground-state Fock operator

$$\varepsilon_r = \langle r|\hat{\mathcal{F}}|r\rangle = \langle \bar{r}|\hat{\mathcal{F}}|\bar{r}\rangle = \langle r|\hat{h}|r\rangle + \sum_{j=1}^n \left[2\langle rj|r_{12}^{-1}|rj\rangle \right.$$
$$\left. - \langle rj|r_{12}^{-1}|jr\rangle \right]. \qquad (5)$$

These virtual MOs are used in the definition of excited determinants, resulting from single, double, multiple substitutions of occupied MOs by virtual ones (hereafter, the indices $i, j, \ldots$ and $r, s, \ldots$ indicate occupied and virtual spinorbitals, respectively)

$$\Phi_{i,r} = a_r^\dagger a_i \Phi_0, \qquad (6)$$

$$\Phi_{ij,rs} = a_s^\dagger a_r^\dagger a_i a_j \Phi_0, \qquad (7)$$

$$\Phi_{ijk,rst} = a_t^\dagger a_s^\dagger a_r^\dagger a_i a_j a_k \Phi_0, \qquad (8)$$

where Φ_0 has been chosen as the vacuum state. Each of these determinants Φ_I is defined from Φ_0 by its holes and its particles. If one introduces a purely monoelectronic zeroth-

order Hamiltonian $\hat{\mathcal{H}}_0$, usually called Møller–Plesset [7],

$$\hat{\mathcal{H}}_0 = \sum_{i=1,2n} \varepsilon_i a_i^\dagger a_i + \sum_{r=2n+1,2M} \varepsilon_r a_r^\dagger a_r + C, \qquad (9)$$

where C is a constant such that

$$\langle \Phi_0|\hat{\mathcal{H}}_0|\Phi_0\rangle = E_0^{(0)} = \langle \Phi_0|\hat{\mathcal{H}}|\Phi_0\rangle, \qquad (10)$$

Φ_0 and all the determinants Φ_I are eigenfunctions of $\hat{\mathcal{H}}_0$, and the associated zeroth-order energy $E_I^{(0)}$ is (except for a constant term C) the sum of the monoelectronic energies of the spin orbitals occupied in this determinant. The zeroth-order energy differences which appear in all the denominators of the terms of the Rayleigh–Schrödinger perturbation expansion are

$$E_0^{(0)} - E_I^{(0)} = \sum_{i=1,2n} \varepsilon_i \langle \Phi_I|a_i a_i^\dagger|\Phi_I\rangle$$
$$- \sum_{r=2n+1,2M} \varepsilon_r \langle \Phi_I|a_r^\dagger a_r|\Phi_I\rangle = \sum_{h(I)} \varepsilon_h - \sum_{p(I)} \varepsilon_p, \quad (11)$$

where h and p are the holes and particles of the Φ_I determinant. An important feature of this choice of $\hat{\mathcal{H}}_0$ is the additivity of the zeroth-order excitation energies, namely the zeroth-order excitation energy to a doubly excited determinant is the sum of the disjoint couples of single excitations that it contains, and similarly the excitation energy to a quadruply excited determinant is the sum of the excitation energies of the disjoint couples of double excitations that it contains, and so on.

With this definition of $\hat{\mathcal{H}}_0$, the perturbation operator $\hat{\mathcal{V}}$ exhibits non-zero diagonal matrix elements

$$\langle \Phi_I|\hat{\mathcal{V}}^{\mathrm{MP}}|\Phi_I\rangle = \sum_{h,h'} \left(J_{hh'} - K_{hh'}\delta_{hh'}^S \right)$$
$$+ \sum_{p,p'} \left(J_{pp'} - K_{pp'}\delta_{pp'}^S \right) - \sum_{h,p} \left(J_{hp} - K_{hp}\delta_{hp}^S \right), \quad (12)$$

where J and K represent, respectively, the Coulomb and exchange integrals and the δ^s symbol is equal to 1 if the spins of the MOs are the same, zero if they are different. In Equation (12), the exchange integrals disappear when $p = p'$ since the two spins are different and K_{pp} (which actually would be a Coulomb integral) is equal to J_{pp} so it never appears in the expression. Hereafter, we shall omit the explicit mention of the exchange integrals. When working with localised MOs, the Coulomb integrals are anyway much larger than the exchange ones.

It is worth mentioning that the most natural definition of a zeroth-order Hamiltonian is not the Møller–Plesset

[7] one, but the so-called Epstein–Nesbet definition [8–10] which takes the diagonal part of the Hamiltonian in the basis of determinants as the zeroth-order one:

$$\langle \Phi_I | \hat{\mathcal{H}}_0^{EN} | \Phi_I \rangle = \langle \Phi_I | \hat{\mathcal{H}} | \Phi_I \rangle. \tag{13}$$

The corresponding perturbation operator, therefore, has zero diagonal elements

$$\langle \Phi_I | \hat{\mathcal{V}}^{EN} | \Phi_I \rangle = 0. \tag{14}$$

Notice that the off-diagonal elements of $\hat{\mathcal{V}}^{EN}$ and $\hat{\mathcal{V}}^{MP}$ are identical

$$\langle \Phi_I | \hat{\mathcal{V}}^{EN} | \Phi_J \rangle = \langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_J \rangle = \langle \Phi_I | \hat{\mathcal{H}} | \Phi_J \rangle, \, I \neq J. \tag{15}$$

If one defines the norm of the operator as the sum of the absolute values of its matrix elements, the norm of the EN perturbation operator is necessarily smaller than that of the Møller–Plesset one. The EN perturbation series of order 2 incorporates an infinite summation of diagrams of the MP series, as noticed in the 1960s by Kelly [15]. The preference for the choice of the MP perturbation expansion is based on several features:

- The demonstration of the linked cluster theorem [1], which insures the strict separability, and the diagrammatic representation of the order by order corrections to the energy and the wave function require the additivity of the excitation energies, which is not obeyed by the EN operator.
- The first-order EN wave function is not an eigenfunction of the $\hat{S}^2$ spin operator. This defect could easily be repaired by taking a unique zeroth order energy for all the determinants of the same space configuration. This would introduce some small diagonal matrix elements of the new $\hat{\mathcal{V}}$ operator, but these elements would only involve exchange integrals, while the MP one implies Coulomb integrals. The norm of this modified EN perturbation operator would still be significantly smaller than that of $\hat{\mathcal{V}}^{MP}$.
- As will be discussed below the EN energy denominators of the doubly excited determinants in molecules are usually smaller in absolute value than those of the MP expansion. When one considers a bond breaking geometry change the perturbation series diverge, due to a near degeneracy between the closed shell ground state configuration and a doubly excited determinant in which the two electrons of the bond are promoted from the bonding bond MO to the antibonding valence MO. The divergence of the EN energy starts at shorter bond lengths than that of the MP expansion [16]. This sooner divergence may unduly lower

the vibrational frequency associated with the bond lengthening.

3. Comparison between low order contributions in the MP and EN series

Let us consider the effect of the double excitations $ij \rightarrow rs$ on Φ_0. In MP expansion, the first order correction to the wave function is

$$C_{ij,rs}^{(1,MP)} = \frac{\langle rs \, | r_{12}^{-1} | \, ij \rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_r - \varepsilon_s}, \tag{16}$$

while in EN series

$$C_{ij,rs}^{(1,EN)} = \frac{\langle rs \, | r_{12}^{-1} | \, ij \rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_r - \varepsilon_s - \langle \Phi_{ij,rs} | \hat{\mathcal{V}}^{MP} | \Phi_{ij,rs} \rangle}, \tag{17}$$

where $\langle \Phi_{ij,rs} | \hat{\mathcal{V}}^{MP} | \Phi_{ij,rs} \rangle = J_{ij} + J_{rs} - J_{ir} - J_{is} - J_{jr} - J_{js}$. This quantity is negative for the most energy contributing double excitations. Thinking in terms of localised (bond) orbitals, the largest double excitations concern intrabond double excitations from the orbital i of a given bond to the anti-bonding valence MO i^* of the same bond. The two MOs have similar spatial features; therefore, the four Coulomb integrals are of the same order of magnitude and the term $\langle \Phi_{i\bar{i},i^*\bar{i}^*} | \hat{\mathcal{V}}^{MP} | \Phi_{i\bar{i},i^*\bar{i}^*} \rangle = J_{ii} + J_{i^*i^*} - 4J_{ii^*} + 2K_{ii^*}$ is large and negative. Other large contributions come from interbond double excitations $ij \rightarrow i^*j^*$ for which

$$\langle \Phi_{i\bar{j},i^*\bar{j}^*} | \hat{\mathcal{V}}^{MP} | \Phi_{i\bar{j},i^*\bar{j}^*} \rangle = J_{ij} + J_{i^*j^*} - J_{ii^*} - J_{jj^*}$$
$$- J_{ij^*} - J_{ji^*} + K_{ii^*} + K_{jj^*}. \tag{18}$$

When the two bonds i and j are far from each other, the largest integrals are J_{ii^*} and J_{jj^*}, imposing a negative sign. The same type of conclusions can be reached for the excitations from the valence MOs to the non-valence virtual orbitals which are responsible for the dynamical correlation. The non-valence virtual MOs are oscillating orbitals but the $ij \rightarrow rs$ excitations only bring a significant contribution when r and s are close to the occupied MOs i and j. In this case again, the six Coulomb integrals are of the same order of magnitude and the four hole-particle attractions prevail on the hole–hole and particle–particle repulsions. One understands therefore that the second-order MP energy will be underestimated. The third-order correction to the energy will involve a negative large contribution

$$\Delta E_{diag}^{(3,MP)} = \sum_I \frac{\langle \Phi_0 | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle \langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle \langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_0 \rangle}{(E_0^{(0)} - E_I^{(0)})^2}$$
$$\tag{19}$$

as long as the diagonal term is negative and large. The third-order MP correction is usually dominated by this term, and of the same sign as the second-order correction. This has led in the past to the proposal of an empirical correction to the MP second-order energies consisting in dividing them by a factor of about 0.8 [17,18]. One must notice that the ratio

$$\lambda_I = \frac{\langle \Phi_I | \hat{\mathcal{V}}^{\mathrm{MP}} | \Phi_I \rangle}{E_0^{(0)} - E_I^{(0)}} \qquad (20)$$

has no reason to be the same for all double excitations. Let us consider for instance an excitation of two electrons occupying the same orbital i to a virtual orbital r situated very far from i. Then the term

$$\langle \Phi_{i\bar{i},r\bar{r}} | \hat{\mathcal{V}}^{\mathrm{MP}} | \Phi_{i\bar{i},r\bar{r}} \rangle \cong J_{ii} + J_{rr} > 0 \qquad (21)$$

becomes clearly positive. Of course the first order coefficient of such a long range double charge transfer excitation is negligible

$$C_{ii,rr}^{(1,\mathrm{MP})} = \frac{\langle rr \, | r_{12}^{-1} | \, ii \rangle}{2 \, (\varepsilon_i - \varepsilon_r)} = \frac{K_{ir}}{2 \, (\varepsilon_i - \varepsilon_r)}, \qquad (22)$$

since the integral in the numerator is negligible. The third-order diagonal correction relative to this excitation would no longer be negative but positive, but it remains negligible.

The ratio between the MP2 and exact correlation energies is not universal. McCarthy and Thakkar [19,20] have recently observed that the MP2 energies of heavy atoms are too large, compared to more accurate non-perturbative evaluations of the correlation energies. Actually if one considers the excitations from inner shells, which are spatially very compact, their correlation involves excitations to non-valence orbitals. These orbitals must have non-negligible amplitudes in the regions of space where the inner orbitals have their large amplitudes, but they have to be orthogonal to the valence orbitals and may be much more diffuse than the core orbitals. Then in the expression of the term $\langle \Phi_{i\bar{i},r\bar{r}} | \hat{\mathcal{V}}^{\mathrm{MP}} | \Phi_{i\bar{i},r\bar{r}} \rangle$ the hole–hole repulsion J_{ii} may be the largest one, especially for the numerous inner core orbitals, which are spatially very compact, and impose a positive sign. This would explain why the third-order diagonal correction would become positive, and the second-order MP correlation energy appears to be too large in absolute value. In order to validate this explanation, we have calculated, starting from the SCF determinant and using a ANO-RCC $(10s9p8d)$ basis set, the excitation energies relative to double excitations from various occupied ns ($n = 2, 3, 4, 5$) orbitals of the Xe atom. It appears from Table 1 that all double excitations except that from the highest occupied orbital have positive diagonal elements of the Møller–Plesset perturbation operator, which will contribute positively to the

MP3 energy. The ratio $\lambda_I = \frac{\langle \Phi_I | \hat{\mathcal{V}}^{\mathrm{MP}} | \Phi_I \rangle}{E_0^{(0)} - E_I^{(0)}}$ is not large (between 2% and 4%) but it does not decrease when the holes go from the $2s$ to the $4s$ orbitals. Actually the overestimation of the MP2 energies begins to occur for the Xe atom [19,20].

4. Possible divergence of the MP series

A perturbation expansion may diverge for several reasons, but it is possible to identify a case where it will necessarily diverge, and in an oscillating mode. This will occur as soon as for a perturbing vector the ratio λ_I is smaller than -1, i.e. when

$$\langle \Phi_I | \hat{\mathcal{V}}^{\mathrm{MP}} | \Phi_I \rangle > E_I^{(0)} - E_0^{(0)}, \qquad (23)$$

which means that

$$\langle \Phi_I | H | \Phi_I \rangle - \langle \Phi_0 | H | \Phi_0 \rangle > 2 \big(E_I^{(0)} - E_0^{(0)} \big). \qquad (24)$$

Actually, suppose that these inequalities are satisfied for at least one perturber, which appears at order p of perturbation (for the wavefunction), for instance if it is a $2p$-time excited determinant (at order 3 for a sextuply excited determinant). Let call $C_I^{(p)}$ its coefficient at this order.

Then, when calculating the wave function at order $p + 1$, a contribution appears which is written

$$\frac{\langle \Phi_I | \hat{\mathcal{V}}^{\mathrm{MP}} | \Phi_I \rangle}{E_I^{(0)} - E_0^{(0)}} C_I^{(p)} = \lambda_I C_I^{(p)}. \qquad (25)$$

At order $p + 2$ one has a contribution equal to $\lambda_I^2 C_I^{(p)}$, and at the order $p + n$ a contribution equal to $\lambda_I^n C_I^{(p)}$. Hence, as soon as $\lambda_I < -1$, whatever the smallness of the coefficient of the determinant when it appears for the first time in the perturbative development, the series will diverge in an oscillatory mode. This unduly large coefficient of a multiply excited determinant will contaminate the coefficients of the less excited determinants, as evident from the order by order expansion of the wave function. The coefficients of the doubly excited determinants will be affected at the order $2p + n$, $p + n$ being the order at which the coefficient of Φ_I becomes meaningless. $C_I^{(p+1)}$ contains the contribution $C_I^{(p)} (1 + \lambda_I)$ and in $C_I^{(p+n)}$ the contribution $C_I^{(p)} \big(1 + \sum_{k=1}^{n} \lambda_I^k \big)$. When the impact of the divergence on the coefficients of the doubly excited determinants becomes manifest, the energy corrections exhibit the same oscillatory behaviour.

When is the critical inequality (23) satisfied? The condition implies that the energy difference between the excited determinant and the reference one is more than two times

Table 1. Analysis of the zeroth-order excitation energies of the Xe atom in the MP and EN partitions for the excitation of two electrons from a ns ($n = 2, 3, 4, 5$) orbital to the lowest empty orbital φ' of s symmetry with orbital energy $\varepsilon_{\varphi'} = 0.0868$ hartrees. The zeroth order description of the ground state, Φ_0, is the HF-SCF determinant, with energy -7441.4758 hartrees. All energies are in hartrees.

Excitation	$\langle \Phi_I | \hat{\mathcal{H}} | \Phi_I \rangle$	$E(\Phi_I) - E(\Phi_0)$	ε_{ns}	$2(\varepsilon_{\varphi'} - \varepsilon_{ns})$	$\langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle$
$(2s)^2 \rightarrow$	-7029.5146	411.9612	-202.2517	404.6769	7.2844
$(3s)^2 \rightarrow$	-7353.0615	88.4143	-42.9735	86.1205	2.2938
$(4s)^2 \rightarrow$	-7423.8466	17.6292	-8.4236	17.0208	0.6085
$(5s)^2 \rightarrow$	-7439.4154	2.0604	-1.0092	2.1919	-0.1315

the MP energy difference. This means that the bielectronic interactions between the holes and the particles are larger than the sum of the monoelectronic energy differences. In view of the fact that the number of hole-particle attractions is larger than the number of repulsions between the holes and between the particles, this seems unlikely at the first glance, but one may perfectly identify the possible occurrence of such a situation. Let us consider first an atomic problem. One must remember the physical meaning of the monoelectronic energies. According to Koopmans' theorem the monoelectronic energy ε_i represents the energy required to extract one electron from the occupied orbital i, i.e. a ionisation potential, while the energy ε_r of a virtual MO r is the energy gained when one electron is added in this MO, i.e. close to an electroaffinity. If one considers the process which sends the six outermost $2p$ electrons of Ne or Cl$^-$ to six virtual very diffuse orbitals of energy near to zero, far from each other and far from the atom, (as can occur in a large basis set), one may calculate the hole–hole repulsion to be

$$\langle \Phi_I | \hat{\mathcal{V}} | \Phi_I \rangle = 15 J_{pp'}, \qquad (26)$$

where $J_{pp'}$ is the Coulomb integral between two $2p$ orbitals of different $\mathcal{L}_z$ values. The factor 15 comes from the numbering of the $(6 \times 5)/2$ bielectronic repulsions between the six electrons of the p shell. This quantity is to be compared with six times the first ionisation potential of the atom. If

$$15 J_{pp'} > 6\varepsilon_p, \qquad (27)$$

the series will diverge. Actually one understands easily that extracting six electrons from the same atom demands much more energy than six times the energy required to extract the first electron since the remaining electrons are subjected to the same nuclear attraction but they are no longer repelled by the already removed electrons. In other words, as the positive charge of the atom increases, the effective nuclear charge affecting the electrons also increases. Looking at Moore's tables [21], one sees that ionising the Ne atom requires 21.6 eV, while ionising it six times, to Ne^{6+}, requires 503 eV, to be compared with six times the first ionisation potential (130 eV). We have calculated

the MP and EN excitation energy for the sextuply excited determinant sending the six $2p$ electrons to the lowest p virtual orbitals in an aug-cc-pVQZ basis set. The MP excitation energy is 6.3422 hartrees, while the EN one is 13.8546 hartrees, i.e. more than twice the MP excitation energy. The ratio λ_I is equal to -1.1845 hartrees. The first-order coefficient of a $(2p \rightarrow 3p)^2$ double excitation is $0.0153/2.114 = 0.0072$. The order of magnitude of the sextuply excited determinant is 0.0072^3 and the sign of divergence of the wave function will be manifest at a very high order.

From the preceding analysis one understands that the hole–hole repulsion plays a key role and that the divergence will preferably affect atoms and molecules presenting a high electron density, as it is the case in Ne and HF, as observed in previous studies [11–13]. The behaviour of the MP expansion for an heavy element would of course be worse, the divergence appearing sooner than for Ne, since the hole–hole repulsions will be very large for inner-shell excitations.

If one considers a heterogeneous system, where the orbitals are atom-, bond- or fragment-centred, or if one works with localised SCF MOs, some of the multiple excitations may send several electrons from the occupied orbitals on a region of space to the virtual orbitals of a remote region of space. In such a case, the hole-particle interactions is very small, and the term $\langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle$ will be positive and large since it reduces to

$$\langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle = \sum_{h, h'} J_{hh'} + \sum_{p, p'} J_{pp'}. \qquad (28)$$

As an example of a similar situation one may imagine the interaction between distant Ne (or Cl$^-$) and Be atoms and consider the sextuple excitations of the six $2p$ electrons from the orbitals of the heavy atom to the six $2p$ orbitals of the Be atom. This determinant appears in the wavefunction at order 3 of perturbation, with a small coefficient, which becomes exponentially small when the interatomic distance increases. For close interatomic distances the quantity $\langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle$ will be negative due to the importance of the hole-particle attractions, but when the distance increases it will become positive and finally will tend to be larger than $6(IP_{Ne}-EA_{Be})$. Since in most molecular

Table 2. Comparison of the zeroth order excitation energies in the MP and EN partitions for the excitation of six or four electrons from the $2p$ orbitals of the Ne atom to the $2p$ empty orbitals of the Be atom as a function of the internuclear distance R. Energies in hartrees, distances in bohrs.

		Six p electrons			Four p electrons		
R	E (HF)	ΔE (MP)	ΔE (EN)	$-\lambda_I$	ΔE (MP)	ΔE (EN)	$-\lambda_I$
2.5	-142.974421	7.1665	14.9244	1.0825	4.6572	7.7036	0.6541
3.0	-143.040230	6.6496	14.8360	1.2311	4.3216	7.6639	0.7734
4.0	-143.091397	5.9298	15.4119	1.5991	3.8878	7.7727	0.9993
6.0	-143.116424	5.5638	17.2489	2.1002	3.6965	8.4057	1.2740
8.0	-143.118335	5.5272	18.5607	2.3580	3.6830	8.9230	1.4227
10.0	-143.118438	5.5241	19.4388	2.5189	3.6826	9.2725	1.5180
15.0	-143.118443	5.5239	20.6367	2.7359	3.6826	9.7683	1.6526

problems such inter-regions multiple excitations belong to the configuration interaction space, one may assess that the divergence of the MP series at high orders is not an exception but a general feature.

This analysis has been numerically tested on the Ne–Be diatomic problem, using a double zeta basis set for simplicity. Table 2 shows the evolution of the MP and EN zeroth-order excitation energies relative to the sextuple excitation $(2p_{Ne})^6 \rightarrow (2p_{Be})^6$, from interatomic distances between 2.5 and 15 bohr. The corresponding λ ratio is always smaller than -1.

Moreover, from Table 2 one sees that the quadruple excitation $(2p_{Ne})^4 \rightarrow (2p_{Be})^4$ also leads to a $\lambda < -1$ ratio when the interatomic distance is larger than 4 bohr.

As a rather funny remark one may add that the EN perturbation series, usually considered as less robust than the MP one, due to its smaller energy denominators (in absolute values) for the double excitations in light atoms and normal element molecules, will not present such diverging behaviours.

5. Conclusion

The present work was motivated by the exotic problem of the observed very high-order oscillatory divergence of the MP series on the Ne atom and the stretched HF molecule [12–14]. The here-studied divergences can be called as due to 'back-door intruders' according to the analysis performed in Ref. [12–14] on the basis of an extension in the complex plane. The divergence is due to a dramatic underestimation of some excitation energies by the Møller–Plesset zeroth order operator. The logical origin of the phenomenon was already correctly understood as due to a ratio between an off-diagonal matrix element of the perturbation operator and the MP energy denominator lower than -1, resulting in the divergent power expansion reported in Equation (14.5.39) at page 778 of Ref. [14]. The sign of the diagonal matrix elements of the MP perturbation operator is usually negative for the double excitations and light atoms, resulting in an underestimation of the correlation

energies at the MP2 level. As discussed here this sign will change when one goes to heavy atoms, which explains the observed underestimation of the MP2 correlation energies (in absolute value) for heavy elements [19,20], the correlation energy of which comes mainly from the numerous inner shell excitations. This sign also changes when one goes to multiple excitations proceeding from occupied orbitals in the same region of space to either diffuse or remote virtual orbitals, due to the prevalence of the hole–hole repulsion (and possibly particle–particle repulsion) over the hole–particle attractions. We have shown that the high-order divergence of the MP series may be expected to be a non-exceptional phenomenon. Of course the most problematic $\langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle$ terms will not affect the reliability of the low-order energies, since they concern either multiply excited determinants or long range double excitations. Indeed, to the second order these double excitations are associated with very small $\langle \Phi_0 | \hat{\mathcal{H}} | \Phi_I \rangle$ matrix elements and do not contribute significantly to the energy.

Nevertheless, these remarks are not purely academic, nor teratologic. They call the attention on the limits of validity of monoelectronic zeroth-order operators, which are a reference tool in many-body theories. As soon as the excitation energy to a quadruply excited determinant is not the sum of the complementary double excitations that it contains, the coefficient of this determinant cannot be obtained as sum of products of the coefficients of these double excitations, which is the basic assumption of the coupled cluster of doubles. The quadruple excitation energy to $T_l^+ T_m^+ | \Phi_0 \rangle$ introduces some new Coulomb and exchange integrals between the holes and particles which were not involved in the double excitations $T_l^+ | \Phi_0 \rangle$ and $T_m^+ | \Phi_0 \rangle$. The additivity of excitation energies only concerns non-interacting couples of pairs of electrons, i.e. situations where T_l^+ and T_m^+ imply non-interacting subsystems. But in this case, the EN energy denominators also are additive and the EN series also provide a strict separability, as soon as localised MOs are used. The appearance of non-additive $\langle \Phi_I | \hat{\mathcal{V}}^{MP} | \Phi_I \rangle$ terms can only be managed by four-body operators $T^{(4)}$. The non-additivity of excitation

energies is especially dramatic in low-dimensional spin systems [22–29]. In order to avoid the introduction of many-body operators in the cluster expansion [30–33], one of us (JPM) and co-workers have suggested the following modification [22–29] of the estimate of the coefficients of quadruples

$$C_{T_l^+ T_m^+ \Phi_0} = \sum_{<p,q>, T_p^+ T_q^+ \Phi_0 = T_l^+ T_m^+ \Phi_0} C_p C_q \frac{\Delta E_p + \Delta E_q}{\Delta E_{p+q}}, \tag{29}$$

where

$$C_p = C_{T_p^+ \Phi_0}, \tag{30}$$

$$\Delta E_p = \langle \Phi_0 | \hat{\mathcal{H}} | \Phi_0 \rangle - \langle T_p^+ \Phi_0 | \hat{\mathcal{H}} | T_p^+ \Phi_0 \rangle, \tag{31}$$

$$\Delta E_{p+q} = \langle \Phi_0 | \hat{\mathcal{H}} | \Phi_0 \rangle - \langle T_p^+ T_q^+ \Phi_0 | \hat{\mathcal{H}} | T_p^+ T_q^+ \Phi_0 \rangle. \tag{32}$$

This modified coupled cluster of doubles formulation satisfies the strict separability condition and happens to give very satisfactory results in the spin lattices studies [22–29]. It would be worth testing its efficiency in ab initio quantum chemical problems, in order to see whether it reduces the need for $T^{(4)}$ operators.

The present work focused on the single reference problem. The question of the choice of an appropriate zeroth-order Hamiltonian for multireference problems is much more complex (see for example the discussion in Ref. [34] for the problems related to the extension of the MP and EN partitions to multireference problems). The conditions to establish the generalised linked cluster theorem [2] in the quasi degenerate perturbation theory context are extremely severe: the reference (or model) space must be a complete active space (of dimension N) and the zeroth-order Hamiltonian must again be monoelectronic. Under these conditions not only the QDPT would not converge for most of the chemical problems, but it will be in general impossible to identify N eigenvectors which have large projections on the model space. In practice, the perturbation may face severe intruder state problems. These intruders will in general be 'front-door intruders', having lower energies than some of the model space determinants, whatever the choice of $\hat{\mathcal{H}}_0$, MP or EN. If one alternatively defines a state-specific perturbation from a multi-reference zeroth-order wave function, which is now solution of a bielectronic Hamiltonian, one may either use single determinants or multiconfigurational vectors as perturbers. In the first case, the separability condition can only be satisfied if one uses a complex bielectronic zeroth-order Hamiltonian [35]. In the case of contracted formalisms, the use of the Møller–Plesset Hamiltonian to define the zeroth-order energies of the contracted vectors, which is adopted in the CASPT2 method [36,37], faces severe intruder state problems, which are in practice managed through empirical level

shifts [38,39]. This intruder state problem is avoided in the NEVPT method [40–43], which is also strictly separable. This method uses the bielectronic Dyall Hamiltonian [44] to define the zeroth-order energies of the contracted perturbers, as does, in some sense, the EN Hamiltonian for the single reference problem.

References

[1] J. Goldstone, Proc. R. Soc. London Ser. A **239**, 269 (1957).
[2] B.H. Brandow, Rev. Mod. Phys. **39**, 771 (1967).
[3] F. Coester, Nucl. Phys. **1**, 421 (1958).
[4] F. Coester and H. Kümmel, Nucl. Phys. **17**, 477 (1960).
[5] J. Cizek, J. Chem. Phys. **45**, 4256 (1966).
[6] R.J. Bartlett, J. Chem. Phys. **93**, 1697 (1989).
[7] C. Møller and M.S. Plesset, Phys. Rev. **46**, 618 (1934).
[8] P.S. Epstein, Phys. Rev. **28**, 695 (1926).
[9] R.K. Nesbet, Proc. R. Soc. London Ser. A **230**, 312 (1955).
[10] P. Claverie, S. Diner, and Jean-Paul Malrieu, Int. J. Quantum Chem. **1**, 751 (1967).
[11] C. Angeli, R. Cimiraglia, and J.P. Malrieu, Chem. Phys. Lett. **317**, 472 (2000).
[12] O. Christiansen, J. Olsen, P. Jørgensen, H. Koch, and P.Å. Malmqvist, Chem. Phys. Lett. **261**, 369 (1996).
[13] J. Olsen, P. Jorgensen, T. Helgaker, and O. Christiansen, J. Chem. Phys. **112**, 9736 (2000).
[14] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).
[15] H.P. Kelly, Adv. Chem. Phys. **14**, 142 (1969).
[16] M.B. Lepetit and J.P. Malrieu, J. Chem. Phys. **87**, 5937 (1987).
[17] P.E.M. Siegbahn, M.R.A. Blomberg, and M. Svensson, Chem. Phys. Lett. **223**, 35 (1994).
[18] P.E.M. Siegbahn, M. Svensson, and P.J.E. Boussard, J. Chem. Phys. **102**, 5377 (1995).
[19] S.P. McCarthy and A.J. Thakkar, J. Chem. Phys. **134**, 044102 (2011).
[20] S.P. McCarthy and A.J. Thakkar, J. Chem. Phys. **136**, 054107 (2012).
[21] C.E. Moore, *Atomic Energy Levels* (National Bureau of Standards, Washington, DC, 1971), Vol. 1, pp. 76–88.
[22] J.P. Malrieu, B. Miguel, and A. Pellegati, Int. J. Quantum Chem. S **29**, 245 (1995).
[23] B. Miguel and J.P. Malrieu, Phys. Rev. B **54**, 1652 (1996).
[24] B. Miguel, M. Couzy, and J.P. Malrieu, Int. J. Quantum Chem. **67**, 115 (1998).
[25] B. Miguel, N. Guihéry, J.P. Malrieu, and P. Wind, Chem. Phys. Lett. **294**, 49 (1998).
[26] M. Al Hajj, N. Guihéry, J.P. Malrieu, and P. Wind, Phys. Rev. B **70**, 099436 (2004).
[27] M. Al Hajj and J.P. Malrieu, Phys. Rev. B **70**, 184881 (2004).
[28] M. Al Hajj and J.P. Malrieu, Phys. Rev. B **72**, 094436 (2005).
[29] J.P. Malrieu, Chem. Phys. **401**, 130 (2012).
[30] R.F. Bishop, J.B. Parkinson, and Y. Xian, Phys. Rev. B **43**, 782 (1992).

[31] R.F. Bishop, R.G. Hale, and Y. Xian, Phys. Rev. Lett. **73**, 3157 (1994).

[32] D.J. Farnell, R.F. Bishop, and K.A. Gernoth, J. Stat. Phys. **108**, 401 (2002).

[33] R.F. Bishop and P.H.Y. Li, Phys. Rev. A **83**, 042111 (2011).

[34] B. Huron, J.P. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1974).

[35] J.L. Heully, J.P. Malrieu, and A. Zaitsevskii, J. Chem. Phys. **105**, 6887 (1996).

[36] K. Andersson, P.Å. Malmquist, B. Roos, A.J. Sadlej, and K. Wolinski, J. Chem. Phys. **94**, 5483 (1990).

[37] K. Andersson, P.Å. Malmquist, and B. Roos, J. Chem. Phys. **96**, 1218 (1992).

[38] B.O. Roos and K. Andersson, Chem. Phys. Lett. **245**, 215 (1995).

[39] N. Forsberg and P.Å. Malmquist, Chem. Phys. Lett. **274**, 196 (1997).

[40] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.P. Malrieu, J. Chem. Phys. **114**, 10252 (2001).

[41] C. Angeli, R. Cimiraglia, and J.P. Malrieu, Chem. Phys. Lett. **350**, 297 (2001).

[42] C. Angeli, R. Cimiraglia, and J.P. Malrieu, J. Chem. Phys. **117**, 9138 (2002).

[43] C. Angeli, M. Pastore, and R. Cimiraglia, Theor. Chem. Acc. **117**, 743 (2007).

[44] K.G. Dyall, J. Chem. Phys. **102**, 4909 (1995).