Theor Chim Acta (1995) 90: 167–187

Theoretica
Chimica Acta

Multiconfigurational second-order perturbative methods: Overview and comparison of basic properties

Jean-Paul Malrieu, Jean-Louis Heully, Andréi Zaitsevskii
IRSAMC, Laboratoire de Physique Quantique, Université Paul Sabatier,
118, route de Narbonne F-31062 Toulouse cedex, France

Received May 20, 1994 / Accepted August 17, 1994

Summary. The multiconfigurational second-order perturbative treatments of molecular electronic calculations can be classified into four groups: i) quasi-degenerate perturbation theory (QDPT) in the basis of determinants, ii) non-degenerate perturbation theory applied to eigenvectors resulting from a truncated CI, ii) QDPT in a model space of non-interacting multiconfigurational functions, iv) intermediate Hamiltonians theory, and examined according to three criteria: i) risk of numerical instability due to intruder states, ii) ability to treat the effect of the outer-space on the model space component of the wavefunction, especially important for the treatment of weakly avoided crossings, iii) separability for (A* ... B) problems. None of the existing methods satisfies these three criteria, as shown both by model analysis and real ab initio calculations on LiF and CuF.

Key words: Many-electron correlation problem – Perturbation theory – Multiconfigurational approaches

1. Introduction

Perturbation theory has been a very useful tool for the study of ground state electronic energies. From a theoretical point of view it has first enabled the main logical structure of the many-body problem to be established through the linked cluster theorem [1]. As a practical tool it has proved to be quite an efficient method for the evaluation of the correlation energy when starting from the Hartree–Fock (HF) determinant, and Möller–Plesset [2] perturbative evaluations of orders 2 to 4 are routinely used in Quantum Chemistry [3]. However, this method is poorly convergent and becomes unreliable when bonds are broken, i.e. when near-degeneracies occur between the HF determinant and some doubly excited determinants. For excited states one also meets both practical difficulties (lack of convergence) and logical problems since the zeroth-order function cannot be single-determinantal.

Thus, it is clear that some combination of perturbation and variation has to be used. Multireference perturbative methods will be briefly reviewed in this paper and their size-consistency will be analyzed in an elementary intermolecular problem. The first and most natural approach consists in using the quasi-degenerate perturbation theory (QDPT) [4–6], i.e. the perturbative construction of an effective Hamiltonian followed by a diagonalization. The second approach consists in diagonalizing first the restriction of the Hamiltonian to a model space and perturbing the resulting eigenvectors one by one, it may be called contracted MRPT. In a third strategy, one perturbs collectively a subset of the eigenstates of the model space again by using QDPT but in a new and smaller model space. The last method uses the recently proposed intermediate effective Hamiltonians [7]. We show that no method combines satisfactorily the following logical and practical requirements,

(i) to avoid instabilities due to intruder states,
(ii) to revise the content of the eigenfunction in the model space under the effect of the perturbation,
(iii) to be size-consistent.

2. Brief review of combinations of perturbation and diagonalization

2.1 *The three criteria*

The methods will be discussed according to three criteria, namely

(i) the likelihood of convergence for the perturbative step,
(ii) the quality of the components of the wavefunction in the model space,
(iii) the additivity of the energy for non-interacting systems.

The first criterion is quite evident, at least from an elementary viewpoint. Any large perturbative ratio in the expansion of the wave operator should be avoided.

The second one can be visualized with the elementary model problem taken from Ref. [8]. Suppose that the model space is spanned by two diabatic configurations (or determinants) ϕ_1 and ϕ_2 which weakly interact through a small matrix element h (independent on the nuclear coordinate r) and that $\langle\phi_1|H|\phi_1\rangle = 0$ and $\langle\phi_2|H|\phi_2\rangle = ar$, so that at zeroth order the crossing occurs at $r = 0$. Suppose also that the outer-space determinants interact more with ϕ_1 than with ϕ_2, or only with ϕ_1. This situation is schematized in the following matrix, reducing the outer-space to one determinant ϕ_3

$$\begin{array}{c} \phi_1 \\ \phi_2 \\ \phi_3 \end{array} \begin{bmatrix} 0 & h & K \\ h & ar & 0 \\ K & 0 & \Delta E \end{bmatrix}$$

and assuming $\Delta E > ar$ and $0 < K \ll \Delta E$.

It is clear that the actual crossing will take place at $r \simeq -K^2/a\Delta E$ and not at $r = 0$. If one defines the crossing point as the distance r_c at which the coefficients of ϕ_1 and ϕ_2 are equal, it becomes possible to give the following limited development of r_c in powers of ΔE^{-1}, assuming $|K| \gg h$.

$$ar_c = -\frac{K^2}{\Delta E} + \frac{K^2 h}{\Delta E^2} + \frac{4h^3}{\Delta E^2} + \frac{K^4}{\Delta E^3} + \cdots \qquad (1)$$

As will be discussed later, some methods do not manage this phenomenon properly.

The last criterion may also be illustrated through a model problem. Consider an intermolecular A ... B problem, where A and B do not interact. Suppose that one is interested in the spectroscopy of A, and thus that the model space determinants differ only by excitations on A. In a correct method, the inclusion of system B in the intermolecular problem does not affect the spectrum of A. Again, it will be seen that some methods do not satisfy this requirement which we may call the "weak separability condition", by contrast with more severe requirements which would ensure separability for problems where the model space involves excited configurations on both A and B. Notice that this weak separability criterion is much less severe than the cancellation of any unlinked contribution.

2.2 *Quasi-degenerate perturbation theory in the basis of determinants*

In principle, when several determinants become nearly degenerate one should make use of quasi-degenerate perturbation theory [9–14]. One first selects a p-dimensional model space S, and the effect of the outer space takes the form of an order-by-order change (or "dressing") of the matrix elements between the model space determinants. QDPT at order n leads to the construction of an effective Hamiltonian ($H^{\mathrm{eff}(n)}$), and the diagonalization of the effective Hamiltonian provides a nth order evaluation of p eigenenergies and of the components of the corresponding eigenvectors on the model space. At convergence the Bloch effective Hamiltonian [5] will be such that, if P_S is the projector on the model space, and if $H\psi_j = E_j\psi_j$, the p roots of H_{eff} satisfy

$$H^{\mathrm{eff}}|P_S\psi_i> = E_i|P_S\psi_i>, \quad i = 1, \dots, p, \tag{2}$$

provided that there is a good correspondence between the model space and a p-dimensional H-stable subspace (the target space, i.e. a set of p eigenvectors of H), the effective Hamiltonian is entirely defined by Eq. (2) or by

$$H^{\mathrm{eff}} = \sum_i |P_S\psi_i> E_i < P_S\psi_i|^{\perp}. \tag{3}$$

Quasi-degenerate perturbation theory treats all the roots of the effective Hamiltonian in an even-handed manner.

To analyse the possible occurrence of *intruder states*, one should examine the second-order expression of the effective Hamiltonian. If ϕ_I and ϕ_J belong to S,

$$\langle\phi_I|H^{\mathrm{eff}(2)}|\phi_J\rangle = \langle\phi_I|H|\phi_J\rangle + \sum_{\phi\alpha\notin S} \langle\phi_I|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\phi_J\rangle(E_J^0 - E_\alpha^0)^{-1}, \tag{4}$$

where E_J^0 and E_α^0 are the eigenvalues of an yet unspecified zeroth order Hamiltonian H_0

$$H_0\phi_J = E_J^0\phi_J, \quad H_0\phi_\alpha = E_\alpha^0\phi_\alpha.$$

One may easily see that this method will diverge as soon as one of the outer space determinants come close in energy to one of the model space determinants. If the model space is a complete active space (CAS), it involves some multiply-excited determinants of high energy; there is no energy gap between the model space and the outer space and the perturbation will diverge sooner or later.

The overlapping of model space and outer space zero-order spectra can be readily avoided by a simple modification of H_0 via an appropriate shift of unperturbed levels

$$E_J^0 \rightarrow E_J^{0\prime} = E_J^0 + S_J,$$

$$E_\alpha^0 \rightarrow E_\alpha^{0\prime} = E_\alpha^0 + S_\alpha,$$

such that $E_J^{0\prime} \ll E_\alpha^{0\prime}$. In several computational schemes the shift is restricted to the model space ($S_\alpha = 0$). Note that the stabilization of second-order terms in shifted expansions does not indicate suppression of divergence, since the additional diagonal perturbation

$$V' = -\sum_J |\phi_J\rangle S_J \langle\phi_J| - \sum_\alpha |\phi_\alpha\rangle S_\alpha \langle\phi_\alpha| \tag{5}$$

hidden at second order, reappears in the numerators of higher-order terms (see Refs. [15–17] for detailed discussions). Moreover, in many cases the large amplitudes of V' should raise some doubts regarding the reliability of low-order estimates. Nevertheless, shifted second-order schemes have proved their viability in molecular electronic structure calculations [18, 19]. This contradiction might be resolved by attributing the shifted second-order terms to potentially converged expansions for operators differing from conventional effective Hamiltonian, namely intermediate Hamiltonians ([17], see Section E).

Regarding the component of the wavefunction in the model space, one may examine the behaviour of $H^{\mathrm{eff}(2)}$ in the weakly avoided crossing model.

If H_0 is not degenerate in the model space ($E_1^0 = 0$, and $E_2^0 = ar$), the second-order effective Hamiltonian $H^{\mathrm{eff}(2)}$ is given by

$$\begin{array}{c} \phi_1 \\ \phi_2 \end{array} \begin{bmatrix} -\dfrac{K^2}{\Delta E} & h \\ h & ar \end{bmatrix}.$$

The crossing takes place at $ar_c = -K^2/\Delta E$.

If H_0 is forced to be degenerate in the model space with the value $E_1^0 = E_2^0 = ar/2$, the second-order effective Hamiltonian becomes

$$\begin{array}{c} \phi_1 \\ \phi_2 \end{array} \begin{bmatrix} \dfrac{K^2}{ar/2 - \Delta E} & h \\ h & ar \end{bmatrix}$$

and the position of the crossing is given by $K^2 = ar_c(ar_c/2 - \Delta E)$, or, after a limited development in powers of ΔE,

$$ar_c = -\frac{K^2}{\Delta E} + \frac{K^4}{2\Delta E^3}.$$

The term $K^4/\Delta E^3$ of r_c (cf. Eq. (1)) is not obtained from $H^{\mathrm{eff}(2)}$, whatever the choice of H_0, but the leading term is correct.

Concerning the additivity of the energy for the ($A \ldots B$) non-interacting problem, one may easily see from Eq. (4) that if the model space is defined by excitations on A only, the double excitations $D_{k_B}^+$ on B will contribute to all diagonal matrix elements only by the same quantity $-h_{k_B}^2/\Delta E_{k_B}$, where

$$h_{k_B} = \langle\phi_{0_B}|H|D_{k_B}^+\phi_{0_B}\rangle \quad \text{and} \quad \Delta E_{k_B} = \langle D_{k_B}^+\phi_{0_B}|H|D_{k_B}^+\phi_{0_B}\rangle - \langle\phi_{0_B}|H|\phi_{0_B}\rangle.$$

From a much more general point of view, Brandow has demonstrated a generalized linked cluster theorem [9] (i.e. the cancellation of all unlinked contributions) *provided* that the model space is a *CAS model space.* To maintain the size-consistency for the perturbative expansions with shifted zero-order approximations is a rather complicated problem. To ensure the size-extensivity and the weak separability as defined in Section 2.1, purely one-electron shifts are usually employed, i.e.

the shift of zero-order Möller–Plesset energies for many-electron basis function is defined via modification of one-electron levels. Freed ([14, 18] and references therein) proposed a choice of shift parameters such that all the valence orbital energies are lifted to their average value. It is clear, however, that such a scheme cannot guarantee the exact separability of effective Hamiltonians (except for the case of separation into identical fragments) because of the explicit dependence of shifted valence orbital energies of one subsystem on the orbital energies of the others. A similar problem appears when all the valence levels are shifted to the bottom of the valence one-electron spectrum, as prescribed by Hose [20]. Note that such shifts are not restricted to model spaces, as zero-order many-electron energies for all configurations containing at least one valence orbital are modified.

Thus one faces a complete contradiction between the conditions for a good convergence and those for separability. The second-order QDPT approach has been applied with significant success by Freed and coworkers, but it is clear that it cannot converge at high orders since in such cases the target space cannot be defined [16, 21] in a reasonable way.

2.3 Prediagonalization in the model space followed by non-degenerate perturbation treatment

As an alternative approach one may think of a preliminary diagonalization in the nearly degenerate subspace S, followed by a perturbation of the resulting eigenvector ψ_i^S

$$P_S H P_S |\psi_i^S\rangle = E_i |\psi_i^S\rangle, \quad \psi_i^S = \sum_{I \in S} C_{Ii} \phi_I. \tag{6}$$

Since the two vectors ψ_i^S and ψ_j^S do not interact directly, the perturbation is essentially due to the outer space and the energy gap between E_i and the outer space determinant energies will be sufficient to *avoid* unduly *small energy denominators.*

The first solution consisted in perturbing the eigenvector ψ_m from its interaction with single determinants of the outer space. This is the principle of CIPSI [22] and of some recent variants [23, 24] of this method. This algorithm is efficient and provides rather reliable results; it is quite general since the model space does not need to be complete and the size of the prediagonalized space may frequently reach several thousands. For such dimensions of the prediagonalized space, when it is selected to include all the determinants with important coefficients in the desired wave functions, the 2nd-order energy becomes rather small and good accuracy may be reached.

The method is not size-extensive [25], although in practice the deviation from *additivity* in intermolecular problems appears very small [26]. The second-order correction for the state ψ_i is

$$E_i^{(2)} = \sum_{\phi_\alpha \notin S} \frac{|\langle \phi_\alpha | H | \psi_i^S \rangle|^2}{E_i^0 - E_\alpha^0}. \tag{7}$$

Now one must define the zeroth-order energies E_i^0 and E_α^0. The most satisfactory definition is called “barycentric Möller–Plesset H_0” [22]. Using a monoelectronic Möller–Plesset Hamiltonian in terms of monoelectronic functions r and energies E_r

$$H^{\mathrm{MP}} = \sum_r E_r a_r^+ a_r, \tag{8}$$

one may define

$$E_\alpha^0 = \langle\phi_\alpha|H^{\mathrm{MP}}|\phi_\alpha\rangle, \tag{9}$$

$$E_i^0 = \langle\psi_i^S|H^{\mathrm{MP}}|\psi_i^S\rangle = \sum_{I\in S} C_{Ii}^2\langle\phi_I|H^{\mathrm{MP}}|\phi_I\rangle. \tag{10}$$

Note that, in contrast with Eq. (8), the zero-order Hamiltonian defined by Eqs. (9) and (10) is not a sum of one-particle operators.

Let us consider the second-order contributions from the so-called inactive double excitations, i.e. excitations from the orbitals which are doubly occupied in all the determinants from S to those which are always empty within S. To ensure even the weak separability, each of the inactive double excitations D_k^+, possible on all the reference determinants ϕ_I should give a simple energy shift $-h_{D_k^+}^2/\Delta E_k$, independent on the state i. Actually the contribution of all determinants $D_k^+\phi_I$ for a state i will be

$$E^{(2)}_{i,D_k^+} = \sum_{I\in S}\frac{|\langle D_k^+\phi_I|H|\psi_i\rangle|^2}{E_i^0 - E^0_{D_k^+\phi_I}} = h_k^2\sum_{I\in S}\frac{C_I^2}{E_i^0 - E^0_{D_k^+\phi_I}}.$$

Since $E^0_{D_k^+\phi_I} = E_I^0 + \Delta E_k$ with $E_I^0 = \langle\phi_I|H^{\mathrm{MP}}|\phi_I\rangle$,

$$\begin{aligned} E_i^0 - E^0_{D_k^+\phi_I} &= -\Delta E_k + E_i^0 - E_I^0 \\ &= -\Delta E_k\left(1 - \frac{E_i^0 - E_I^0}{\Delta E_k}\right). \end{aligned}$$

If one considers that $\Delta E_k \gg |E_i^0 - E_I^0|$ it becomes possible to write

$$\begin{aligned} E^2_{i,D_k^+} &= -\frac{h_k^2}{\Delta E_k}\left(1 + \frac{E_i^0 - \sum_I C_I^2 E_I^0}{\Delta E_k} + \mathcal{O}(2)\right) \\ &= -\frac{h_k^2}{\Delta E_k}(1 + \mathcal{O}(2)) \quad \forall i. \end{aligned}$$

Thus the barycentric MP definition of H_0 (Eq. (10)) ensures an approximate separability of inactive double excitations. In the practical codes this property has been forced by estimating the effect of inactive double excitation as $-h_{D_k^+}^2/\Delta E_k$, but this is a theoretical weakness of CIPSI. The other choices of H_0, for instance the use of the eigenvalue of $P_S H P_S$

$$E_i^0 = \langle\psi_i^S|H|\psi_i^S\rangle$$

are really poor since increasing the size of S and lowering E_i^0, the absolute value of the denominators in Eq. (7) is increased and the effect of the inactive double excitations is lowered in absolute value. This problems has apparently been ignored in a recently proposed variant [24] of the CIPSI strategy.

Before discussing the behaviour near avoided crossings, let us present briefly the principle of the CASPT2 algorithm of Roos and coworkers [27]. This method, as CIPSI does, perturbs the vector ψ_i^S resulting from the diagonalization of a reference space which is now required to be a Complete Active Space. CASPT2 no longer perturbs ψ_i^S under the effect of single determinants but under the effect of multideterminantal perturbers $D_l^+\psi_i^S$. This change leads to a dramatic increase in speed, despite some complexity due to the non-orthogonality of $D_l^+\psi_i^S$ and $D_n^+\psi_i^S$. The choice of H_0 as a sum of one-electron generalized Fock operators ensures that inactive double excitations have exactly the same effect

$$E^2_{i,D_k^+} = -\frac{h^2_{D_k^+}}{\Delta E_k} \quad \forall i$$

on all eigenstates, as it should be. One should realize that the CASSCF multiconfigurational reference function is not generally an eigenfunction of this purely one-electron H_0 and the perturbation matrix elements

$$\langle \psi_i^S | V | \psi_j^S \rangle \neq \langle \psi_i^S | H | \psi_j^S \rangle = 0, \quad i \neq j$$

should not vanish [28]. The problem has been circumvented *via* projecting H_0 onto the subspaces spanned by ψ_i^S, the rest of eigenvectors of $P_S H P_S$, and single and double replacements from ψ_i^S respectively.

Both CASPT2 and CIPSI are "*contracted*", i.e. the component of the wavefunction in the model space is determined by the first variational step. In our model problem

$$\psi^S_+ = \cos\varphi\,\phi_1 + \sin\varphi\,\phi_2, \quad \psi^S_- = -\sin\varphi\,\phi_1 + \cos\varphi\,\phi_2$$

and $\varphi = \pm\pi/4$ for $r = 0$. At this distance ψ_i^S has equal components on ϕ_1 and ϕ_2, and the perturbation energy will be

$$E^2_{(+)} = -\frac{K^2}{2\Delta E} = E^2_{(-)},$$

i.e. shared between the two eigenstates. As was shown in Ref. [8], if $|K^2/\Delta E| \gg |h|$, the potential energy curves present spurious irregular behaviours with double curve crossing, one near $r = 0$ and the other one near $r = -k^2/a\,\Delta E$. This is due to the contraction of the wavefunctions ψ^S_+ and ψ^S_- which should be mixed under the effect of the outer-space determinant.

CASPT2 has proved to be a very efficient and accurate method, at least when large model spaces are used. It has been applied to the calculation of the vertical spectrum of a large series of conjugated organic molecules [29] and giving a 0.1 eV mean error for the transition energies. However, if one considers one of the most significant errors (0.4 eV), it concerns a ^1B_u state of butadiene [29] which is a strong mixture of a valence configuration of valence bond ionic character and a Rydberg configuration. Since the dynamical correlation of the former configuration is much larger than that of the latter, it should change the valence/Rydberg content of the wavefunction and the 2nd-order results of the contracted procedure are therefore not correct, the spacing between the two ^1B_u states is underestimated and the energy of the lowest is too high.

2.4 Use of quasi degenerate perturbation theory after diagonalization of the model space CI

In view of this difficulty Sheppard et al. [30] and Spiegelmann and Malrieu [8] were led to propose the use of Quasi-degenerate perturbation in the basis of the eigenstates of $P_S H P_S$ (i.e. of the model space CI matrix). Then the interaction between the model space functions and the outer-space determinant redefines off-diagonal effective interactions between the function ψ_i^S and ψ_j^S, which are zero at first order

$$\langle\psi_i^S|H^{\mathrm{eff}(2)}|\psi_j^S\rangle = \sum_{\alpha\notin S}\frac{\langle\psi_i^S|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\psi_j^S\rangle}{E_j^0 - E_\alpha^0}, \tag{11}$$

thus *revising* the *model space component of the eigenvector*. In the weakly avoided crossing model, one sees that

$$\langle\psi_1^S|H^{\mathrm{eff}(2)}|\psi_2^S\rangle = \sin\varphi\cos\varphi\frac{K^2}{\Delta E},$$

which becomes maximum for $r = 0$. At this distance it is clear that

(i) the first-order interaction between ψ_1 and ψ_2 leads to a $\pi/4$ rotation of the vectors ψ_+^S and ψ_-^S with respect to ϕ_1 and ϕ_2,

$$\psi_+ = (\phi_1 + \phi_2)/\sqrt{2}, \quad \psi_- = (\phi_1 - \phi_2)/\sqrt{2},$$

(ii) the second-order interaction through the outer space, when it is larger than the 1st-order one ($|K^2/\Delta E| \gg |h|$) induces a backward rotation; the eigenstates of the effective Hamiltonian are closer to the diabatic configurations ϕ_1 and ϕ_2 as they should be

$$\psi_1 = (\psi_+ + \psi_-)/\sqrt{2} = \phi_1,$$

$$\psi_2 = (\psi_+ + \psi_-)/\sqrt{2} = \phi_2.$$

The prejudice introduced by the 1st-order diagonalization is repaired by the QDPT treatment.

This treatment is a decontracted perturbation, and there is no doubt that a correct treatment of excited states must be decontracted in order to revise the content of the wavefunction in the model space.

At this stage one has only chosen a new basis set of the initial model space. In such a method the *intruder state problems* will be even more severe than when applying QDPT in the basis of single determinants, since the highest eigenvalues of P_SHP_S are pushed above the highest energies of the determinants. The authors of Refs. [8, 30] have immediately suggested a reduction of the model space to the lowest eigenstates of $P_S\mathrm{H}P_S$ of spectroscopic interest. Let us call S' this reduced model space

$$P_{S'} = \sum_{i=1}^{m} |\psi_i\rangle\langle\psi_i|, \quad m < n, \tag{12}$$

$$P_SHP_S|\psi_i\rangle = E_i|\psi_i\rangle.$$

One may call S'' the complementary subspace

$$P_{S''} = P_S - P_{S'} \tag{13}$$

$$P_{S''} = \sum_{j=m+1}^{n} |\psi_j\rangle\langle\psi_j|. \tag{14}$$

Then since $\langle\psi_i|H|\psi_j\rangle = 0$, the effective Hamiltonian built on S' will not have any second-order contribution coming from S''

$$\langle\psi_i|H^{\mathrm{eff}(2)}|\psi_j\rangle = \sum_{\alpha\notin S}\frac{\langle\psi_j|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\psi_j\rangle}{E_j^0 - E_\alpha^0}, \quad i,j\in S'. \tag{15}$$

The energy difference $E_j^0 - E_\alpha^0$ should always be rather large since one is only interested in the lowest states, and the method will avoid any intruder state problem.

The QDPT version of CIPSI in the basis of the relevant eigenstates of P_SHP_S is routinely used in our group in the most accurate spectroscopic studies (see for instance Refs. [31, 32]). There is no doubt that an uncontracted modification of the CASPT2 method can also be conceived by calculating off-diagonal 2nd-order coupling between the selected ψ_i vectors, and this would solve the major defects of this efficient method. The use of QDPT in a basis of eigenfunctions of P_SHP_S has been reproposed recently by several authors (Nakano [33], Decleva et al. [34]) and numerical tests illustrate the efficiency of the method.

It is not clear, however, that the method solves all near-degeneracy problems or that it makes possible the computation of reliable potential energy surfaces on a wide domain of geometries. Suppose that one is interested in m roots obtained from the diagonalization of a n-dimensional CI space containing the main physics of the m states. Now if an avoided crossing occurs between the mth root and the $(m + 1)$th root, the defect already discussed for $m = 1$ in the discussion of CASPT2 type algorithms will affect the mth energy. The QDPT can only mix the mth eigenstate with the $(m - 1)$ states of lower energy; it cannot mix it with the $(m + 1)$th eigenstate, and this will result in spurious behaviours, double curve crossings and so on. The response cannot be that one should enlarge the restricted model space S' to dimension $(m + 1)$, since it is likely that for some molecular conformations the $(m + 1)$th eigenstate presents another avoided crossing with the next one, and the problem will reappear somewhere else. So that in practice the diagonalization of P_SHP_S, followed by a reduction of the model space in a basis of multiconfigurational roots, cannot be considered as a general recipe to circumvent both the intruder state problem and the need to revise the model-space component of the eigenvector.

From a qualitative point of view, one may question the ability of this method to treat correctly the effect of the outer space. Consider for instance a valence CAS model space. The outer space should bring what is usually called the external or dynamical correlation. Let us write the valence CAS in an orthogonal valence bond basis set (by defining nearly atomic orthogonal molecular orbitals as the basis of the subspace of active MOs). Then the determinants may be classified as neutral, singly ionic, doubly ionic, etc ... Since they have higher energies, the multiply ionic determinants have lower weights in the CAS wavefunction ψ. But the dynamical correlation (orbital breathing, instantaneous repolarisations effects, Coulomb hole effect) is larger for the (multiply) ionic VB components than for the neutral ones. If one works in the basis of the lowest eigenstates of P_SHP_S, and only takes into account the 2nd-order impact of the outer space in a reduced model space, one will never include these specific effects of dynamical correlation on the highly ionic components, which essentially span the upper eigenstates of P_SHP_S. Dynamical correlation increases the weight of multiply ionic VB structures in the wavefunction with respect to the CAS level, as shown elsewhere [35]. This effect is not taken into account if one builds an effective Hamiltonian on the reduced model space of the lowest eigenstates of P_SHP_S, since these eigenstates are essentially VB-neutral.

Moreover, the method is *not size-consistent*, as numerically demonstrated by Sheppard [36] and it does not satisfy our weak separability criterion. Actually the inactive process D_k^+ possible on all determinants of the model space appears through the interaction with the outer space determinants $D_k^+ \phi_I$. These determinants should not contribute to an effective coupling between ψ_i^S and ψ_j^S since the excitation D_k^+ may concern an infinitely remote electron pair. However, Eq. (15) gives, for the contribution of the excitation D_k^+,

$$\begin{aligned}\langle\psi_i^S|H_{D_k^+}^{\mathrm{eff}(2)}|\psi_j^S\rangle &= \sum_{I\in S}\frac{\langle\psi_i^S|H|D_k^+\phi_I\rangle\langle D_k^+\phi_I|H|\psi_j^S\rangle}{E_j^0 - E_{D_k^+\phi_I}^0}\\ &= \sum_{I\in S}\frac{\langle\phi_i^S|H|D_k^+\phi_I\rangle\langle D_k^+\phi_I|H|\phi_I\rangle C_{Ii}C_{Ij}}{E_j^0 - E_{D_k^+\phi_I}^0}\\ &= h_k^2\sum_{I\in S}\frac{C_{Ii}C_{Ij}}{E_j^0 - E_{D_k^+\phi_I}^0}\end{aligned}$$

If one could write

$$E_j^0 - E_{D_k^+\phi_I}^0 = -\Delta E_k \quad \forall j, I, \tag{16}$$

then one would have

$$\langle\psi_i^S|H_{D_k^+}^{\mathrm{eff}(2)}|\psi_j^S\rangle = \frac{-h_k^2}{\Delta E_k}\sum_{I\in S}(C_{Ii}C_{Ij}) = \frac{-h_k^2}{\Delta E_k}\delta_{ij}, \tag{17}$$

since

$$\sum_{I\in S} C_{Ii}C_{Ij} = \langle\psi_i^S|\psi_j^S\rangle.$$

Then the inactive double excitations D_k^+ would not couple the eigenvectors and would lead to a simply energy shift of the spectrum. However, there is no way to satisfy Eq. (16). In the CIPSI program we have enforced Eq. (17) for inactive excitations, but this practical trick is not satisfactory since one may image an A...B problem in which at short r_{AB} distance some orbitals on B are declared active, while they become practically inactive for the lowest states, the excitation concentrating on A when r_{AB} increases. Then some excitations on B, involving these pseudo-active MOs, will have spurious effects on the spectrum of A. Recent papers supporting this strategy have apparently completely neglected this problem, and have not taken the precaution which has been introduced in CIPSI.

2.5 *Use of intermediate Hamiltonians*

Another strategy consists in using a new perturbation theory and intermediate effective Hamiltonians [7], proposed by our group. This method is based on a partition of the model space into a main model space S_n of dimension m ($m < n$) and an intermediate model space of dimension ($n - m$), i.e. the complement of the main model space in the total model space. Then one searches for an effective Hamiltonian H^{int} built in the total model space, such that its diagonalization gives only m exact roots and the components of the corresponding m eigenvectors in the total model space

$$H^{\mathrm{int}}|P_S\psi_k\rangle = E_k|P_S\psi_k\rangle, \quad k = 1, m. \tag{18}$$

The ($n - m$) other roots are not necessarily exact. Of course the intermediate Hamiltonians are not uniquely defined since one only imposes $m \times n$ conditions, and several variants have been proposed recently. The simplest perturbative approach proposed in Ref. [7] assumed that the zeroth-order Hamiltonian was degenerated in the main model space S_m

$$H_0\phi_I = E_0\phi_I \quad \forall I \in S_m. \tag{19}$$

(This restriction is not compulsory [37, 38].) Then the matrix elements of the 2nd-order intermediate Hamiltonian are given by

$$\langle\phi_I|H^{\mathrm{int}(2)}|\phi_J\rangle = \langle\phi_I|H|\phi_J\rangle - \sum_{\phi_\alpha \notin S} \frac{\langle\phi_I|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\phi_J\rangle}{E^0 - E^0_\alpha} \quad \forall I, J \in S. \tag{20}$$

At this stage $H^{\mathrm{int}(2)}$ appears as an effective Hamiltonian $H^{\mathrm{eff}(2)}$ in which one has forced the zeroth-order energies of the model space determinants to be degenerate. But the higher orders differ, as may be seen in the order-by-order expansion of Ref. [7]. What are the properties of intermediate Hamiltonians?

(i) There is no need to fear accidents regarding their *convergence*, provided that the intermediate model space involves all the determinants strongly interacting with those of the main model space. If ϕ_J belongs to the intermediate model space the quantity $(E^0_J - E^0_\alpha)^{-1}$ never appears at any order, and the method accepts near-degeneracies between outer space intruder determinants and intermediate model space determinants. The intermediate model space acts as a variational buffer between the main model space and the outer space. In practice all tests [7, 21, 39, 40] have proved the numerical efficiency of this approach.
(ii) The method is *not contracted*, since the diagonalization occurs after the perturbative correction, as in QDPT. If one considers for instance the weakly avoided crossing model, the main model space must be defined by ϕ_2 at $r < 0$ and ϕ_1 at $r > 0$. The resulting intermediate Hamiltonian, at second order will be

$$\begin{array}{c} \phi_1 \\ \phi_2 \end{array} \begin{bmatrix} \dfrac{K^2}{ar - \Delta E} & h \\ h & ar \end{bmatrix} \quad \text{for } r < 0,$$

$$\begin{array}{c} \phi_1 \\ \phi_2 \end{array} \begin{bmatrix} -\dfrac{K^2}{\Delta E} & h \\ h & ar \end{bmatrix} \quad \text{for } r > 0.$$

Notice that this Hamiltonian is continuous, but has discontinuous derivatives. The position of the crossing ($|C_1| = |C_2|$) satisfies

$$K^2(ar_c - \Delta E) = ar_c$$

and the power expansion of the root gives

$$ar_c = \frac{K^2}{\Delta E} + \frac{K^4}{\Delta E^3} + \cdots$$

which should be compared to Eq. (1). The terms in K^2h and h^3 are missing but the term $K^4/\Delta E^3$ is correct, and if $|K| \gtrsim |h|$ one sees that the intermediate 2nd-order Hamiltonian describes the ground state potential curve more accurately than does the effective Hamiltonian. This is normal, since the intermediate Hamiltonian concentrates here on the lowest root, sacrifying the accuracy of the second one. However, one must mention that going to third-order one introduces an off diagonal matrix element

$$\langle\phi_1|H^{\text{int}(3)}|\phi_2\rangle = -\frac{\langle\phi_1|H|\phi_3\rangle\langle\phi_3|H|\phi_1\rangle\langle\phi_1|H|\phi_2\rangle}{(E_0 - E_3)^2}$$

$$= -\frac{K^2 h}{(ar - \Delta E)^2}$$

for $r > 0$, while for $r < 0$, the main model space is ϕ_2 and $\langle\phi_1|H^{\text{int}(3)}|\phi_2\rangle = 0$. Thus the intermediate Hamiltonian becomes discontinuous at third order in this special case.

However the 2nd-order intermediate Hamiltonian is *not separable*. The inactive double excitations D_k^+ acting on $\phi_I \in S$ will give a determinant $D_k^+ \phi_I$ only interacting with ϕ_I and leading to a diagonal correction

$$\langle\phi_I|H_{D_k^+}^{\text{int}(2)}|\phi_I\rangle = h_k^2/(E^0 - E^0_{D_k^+\phi_I}).$$

The denominator depends on ϕ_I, whatever the choice of E^0, and thus the inactive double excitations, instead of shifting the whole spectrum by a unique amount $-h_k^2/\Delta E_k$ as they should, will have spurious influences on the spectrum and on the component of the wave functions in the model space. This is a major defect of the second-order intermediate Hamiltonian.

The physical origin of this defect is easily understood by considering the following model problem. Imagine an active space made of a ground state HF determinant ϕ_0 and a doubly excited configuration ϕ^*, σ and σ^* being the two active MOs

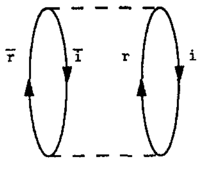
$$\phi_0 = |\text{core}\,\sigma\bar{\sigma}| = |\ldots i\bar{\imath}\ldots\sigma\bar{\sigma}|,$$

$$\phi^* = |\text{core}\,\sigma^*\bar{\sigma}^*| = |\ldots i\bar{\imath}\ldots\sigma^*\bar{\sigma}^*|.$$

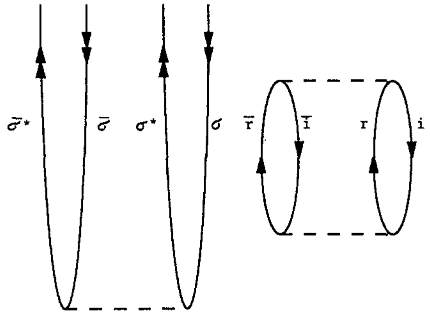
This is a typical CAS for a single bond breaking. Consider now a double excitation from the core MO i to a virtual MO r

$$D_k^+ = a_r^+ a_{\bar{r}}^+ a_{\bar{\imath}} a_i.$$

Then $D_k^+\phi_0 = |\ldots r\bar{r}\ldots\sigma\bar{\sigma}|$ only interacts with ϕ_0, while $D_k^+\phi^* = |\ldots r\bar{r}\ldots\sigma^*\bar{\sigma}^*|$ only interacts with ϕ^*. If we concentrate on the ground state energy, ϕ_0 defines the main model space, and $E_0 = \langle\phi_0|H|\phi_0\rangle$. The diagonal corrections of $H^{\text{int}(2)}$ (diagonal dressings) are $-K_{ir}^2/\Delta(i \to r)^2$ for ϕ_0 and $-K_{ir}^2/(\Delta(\sigma \to \sigma^*)^2 + \Delta(i \to r)^2)$ for ϕ^*. Actually the error comes from the fact that the dressing of ϕ_0 incorporates the 2nd-order diagram



while the dressing of ϕ_0 only incorporates the diagram



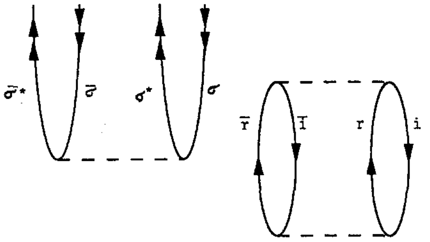
i.e. the 3rd-order correction to the wavefunction in the RS development from ϕ_0

$$\frac{|\phi^*\rangle\langle\phi^*|H|D^+_{(i\to r)^2}\phi^*\rangle\langle D^+_{(i\to r)^2}\phi^*|H|\phi^*\rangle\langle\phi^*|H|\phi_0\rangle}{[\Delta(i\to r)^2+\Delta(\sigma\to\sigma^*)^2]\,\Delta(\sigma\to\sigma^*)^2}$$

and not the other correction of the same order

$$\frac{|\phi^*\rangle\langle\phi^*|H|D^+_{(i\to r)^2}\phi^*\rangle\langle D^+_{(i\to r)^2}\phi^*|H|D^+_{(i\to r)^2}\phi_0\rangle\langle D^+_{(i\to r)^2}\phi_0|H|\phi_0\rangle}{[\Delta(i\to r)^2+\Delta(\sigma\to\sigma^*)^2]\,\Delta(i\to r)^2}$$

This last contribution is pictured by the diagram.



Added to the preceding one it will give the same dressing $-K^2_{i\to r}/\Delta(i\to r)^2$ of the $\langle\phi_i^*|H^{\text{int}}|\phi^*\rangle$, due to the energy denominator additivity but this diagram only appears in the 3rd-order correction to the intermediate Hamiltonian

$$\langle\phi_I|H^{\text{int}(3)}|\phi_J\rangle=\sum_{\alpha\notin S}\sum_{\beta}\frac{\langle\phi_I|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\phi_\beta\rangle\langle\phi_\beta|H|\phi_J\rangle}{E^0-E^0_\alpha\,E^0-E^0_\beta}$$
$$-\sum_{\alpha\notin S}\sum_{\phi_k\notin S_m}\frac{\langle\phi_{=}|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\phi_k\rangle\langle\phi_k|V|\phi_J\rangle}{(E^0-E^0_\alpha)^2}$$

One sees that the contribution missing to $H^{\text{int}(2)}$, needed to obtain separability, belongs to the first summation of the above equation, with

$$\phi_J=\phi_0,\quad \phi_\beta=D^+_{(i\to r)^2}\phi_0,\quad \phi_\alpha=D^+_{(i\to r)]^2}\phi^*.$$

So the intermediate Hamiltonians cannot behave properly regarding separability (and thus size-extensivity) if they are built in the order-by-order expansion proposed in the original development [7]. This promising tool, which avoids the intruder-state nightmare and which behaves correctly in the weakly avoided crossings, must be revised to satisfy the basic separability criterion. This is the subject of a forthcoming work.

Finally, one should mention for completeness the possibility of building an intermediate Hamiltonian in the multiconfigurational basis resulting from a prediagonalization of the model space CI matrix $P_S H P_S$. One obtains, as in Sect. 2.4 a basis of non-interacting vectors ψ, among which the lowest states define a main model space, the other ones being considered as defining the intermediate model space

$$P_S = P_{S'} + P_{S''}, \tag{21}$$

$$P_{S'} = \sum_{k=1,m} |\psi_k\rangle\langle\psi_k|, \tag{22}$$

$$P_{S''} = \sum_{l=m+1,n} |\psi_l\rangle\langle\psi_l|. \tag{23}$$

In Sect. 2.4, one built an effective Hamiltonian on the reduced model space S' associated with P_S, using QDPT. Since $\langle\psi_k|H|\psi_l\rangle = 0$, the intermediate eigenspace vectors did not play any role in $H^{\mathrm{eff}(2)}$. Now one may build an intermediate Hamiltonian in the whole S space, and the intermediate eigenvectors ψ_l will again play a role since in $H^{\mathrm{int}(2)}$ they are coupled with the model space functions through the outer space. Actually, in this approach,

$$\langle\psi_i|H^{\mathrm{int}(2)}|\psi_l\rangle = \sum_{\alpha\notin S} \frac{\langle\psi_i|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\psi_l\rangle}{E_0 - E_\alpha^0} \quad \forall\psi_i\psi_l \in S. \tag{24}$$

It should be mentioned that the latter version of the intermediate Hamiltonian theory is rather costly; it is not clear whether the additional effort required to fulfill the $\{\phi_p\} = \rangle\{\psi_p\}$ transformations is really compensated by improved results in comparison with Eq. (20). A simplified version neglecting the dressing of intermediate states had been proposed in [7] and intensively used in molecular calculations [41, 42]:

$$\langle\psi_p|H_{\mathrm{int}}^{(2)}|\psi_i\rangle = \sum_{\alpha\notin S} \frac{\langle\psi_p|H|\phi_\alpha\rangle\langle\phi_\alpha|H|\psi_i\rangle}{E_i^0 - E_\alpha^0} \quad \forall\psi_p \in S, \quad \psi_i \in S', \tag{25}$$

$$\langle\psi_p|H_{\mathrm{int}}^{(2)}|\psi_k\rangle = 0 \quad \forall\psi_p \in S, \quad \psi_k \in S''. \tag{26}$$

Since $P_S, H_{\mathrm{int}} P_{S''} = 0$, the main eigenvalues of H_{int} coincide with those of the block $P_{S'} H_{\mathrm{int}} P_{S'}$, which is identical with the effective Hamiltonian (Eq. (11)). Indeed, one can easily verify that if $\hat{\psi}$ is a eigenfunction of $P_{S'} H_{\mathrm{int}} P_{S'}$ with the eigenvalue $\hat{E}$, the function

$$\hat{\psi} + \frac{P_{S''}}{\hat{E} - P_{S''} H_{\mathrm{int}} P_{S''}} H_{\mathrm{int}} P_{S'} \hat{\psi}$$

is the eigenfunction of H_{int} (Eqs. (25) and (26)) with the same eigenvalue $\hat{E}$. Moreover, for the systems with a large number of correlated electrons, all the values of $\hat{E} - \langle\psi_k|H_{\mathrm{int}}|\psi_k\rangle$, $\psi_k \notin S'$ (the main model space) should be quite large, so the intermediate-space correction to $\hat{\psi}$ becomes asymptotically insignificant.

Recently Hoffmann [43] has proposed a computational method which, restricted to second order, is based on the hermitized version of Eqs. (25) and (26). Since he does not use an intermediate Hamiltonian, he does not calculate the 2nd-order dressing of the intermediate subblock $P_{S''}H^{\text{int}(2)}P_{S''}$, which appears in our theory. It may be demonstrated that if one omits these corrections, the weakly avoided crossing problem is not solved correctly. In our 3×3 model problem one defines a new basis set

$$\bar{\psi}_1 = \cos\varphi\,\phi_1 + \sin\varphi\,\phi_2,$$
$$\bar{\psi}_2 = \sin\varphi\,\phi_1 + \cos\varphi\,\phi_2,$$

such that $\langle\psi_1|H|\psi_2\rangle = 0$ and one obtains a new expression of the CI matrix

$$\begin{array}{c} \bar{\psi}_1 \\ \bar{\psi}_2 \\ \bar{\phi}_3 \end{array} \begin{bmatrix} H_{11} & 0 & K\cos\varphi \\ 0 & H_{22} & -K\sin\varphi \\ K\cos\varphi & -K\sin\varphi & \Delta E \end{bmatrix}.$$

One shall take ψ_1 as the main model space, ψ_2 as the intermediate model space and the dressed matrix $H^{\text{int}(2)}$ will be

$$\begin{bmatrix} H_{11} + \dfrac{K^2\cos^2\varphi}{H_{11}-\Delta E} & \dfrac{-K^2\sin\varphi\cos\varphi}{H_{11}-\Delta E} \\ \dfrac{-K^2\sin\varphi\cos\varphi}{H_{11}-\Delta E} & H_{22} + \dfrac{K^2\sin^2\varphi}{H_{11}-\Delta E} \end{bmatrix},$$

which is now continuous, with continuous derivatives.

At $r = 0$, $\sin\varphi = \cos\varphi = 1/\sqrt{2}$ if $h < 0$, $H_{11} = h$, $H_{22} = -h$, and the matrix becomes

$$\begin{bmatrix} h + \dfrac{K^2}{2(h-\Delta E)} & \dfrac{-K^2}{2(h-\Delta E)} \\ \dfrac{-K^2}{2(h-\Delta E)} & -h + \dfrac{K^2}{2(h-\Delta E)} \end{bmatrix}.$$

If $K^2/\Delta E \gg |h|$, one sees that the dressing will force a backward rotation, as it should; the lowest eigenstate is essentially concentrated on ϕ_1

$$\psi_1 \simeq (\bar{\psi}_1 - \bar{\psi}_2)/\sqrt{2} = \phi_1$$

and $E_1 \simeq -K^2/\Delta E$.

If one omits the dressing of the intermediate space and omits h, the matrix takes the form

$$\begin{bmatrix} -\dfrac{K^2}{2\Delta E} & +\dfrac{K^2}{2\Delta E} \\ +\dfrac{K^2}{2\Delta E} & 0 \end{bmatrix}.$$

The energy is not correct; $E_1 \simeq (-1-\sqrt{5}K^2)/4\Delta E \simeq -0.8K^2/\Delta E$ and the wave function has a spuriously large component on ϕ_2

$$\psi \simeq 0.85\phi_1 - 0.52\phi_2.$$

3. Numerical examples

To illustrate the essential features of some of the second-order methods reviewed above, we present here two numerical examples. First, we apply both contracted and non-contracted second-order schemes to calculate the ground state of the LiF molecule in the region of the weakly avoided ionic–neutral curve crossing. Then, we consider the results of perturbative calculations of the two lowest $^1\Sigma^+$ states of the CuF molecule arising essentially from two strongly coupled configurations of the same Cu^+F^- ionic character, differing in the occupancy of the Cu^+ d-subshell. In both cases a balanced description of non-dynamical correlation is necessary for a qualitatively correct reproduction of the main features of the electronic structure. Our study focuses on the comparison of contracted-versus-decontracted calculation schemes (for detailed numerical analysis of size-extensivity and separability aspects of second-order methods see Refs. [26, 36, 39, 44]).

3.1 The neutral–ionic avoided crossing in LiF molecule

The ground state of this molecule, being purely ionic near the equilibrium, dissociates adiabatically into neutral species passing via an ionic–neutral avoided crossing region at $r_{ac} = 12.2$–12.7 a.u. [45]. Due to the weakness of ionic–neutral interactions and large differences in the correlation energies for ionic and neutral structures, the calculated position of the avoided crossing point is extremely sensitive to the level of correlation treatment.

To make a direct comparison of our results with the full CI data, we used the same contracted Gaussian basis set $(9s4p)/[4s2p]$ Li, $(9s6p1d)/[4s3p1d]$ F, as in Ref. [45]. The one-electron functions were obtained by the CASSCF method [46]. The active space, which also defined the S-space for subsequent second-order calculations, comprised two σ-MO arising from $2p\sigma$ AO of F and $2s$-AO Li. Direct state-specific minimization of the ground state energy fails to produce continuous solutions in the avoided crossing region [45], so we employed the state-average (SA) optimization procedure with equal weights for two lowest $^2\Sigma^+$ states. Since the restricted CASSCF model does not include any dynamical correlation, it substantially underestimates the relative binding energy of the strongly correlated ionic structure, providing an aberrant estimate for the avoided-crossing position ($r_{ac} = 7.5$–8 a.u.).

Three different second-order schemes based on Möller–Plesset-like partitioning of the many-electron Hamiltonians were employed to take external correlations effects into account:

(i) Non-degenerate generalized Möller–Plesset perturbation theory with CASSCF multiconfigurational zero-order functions known as CASPT2 [27, 46].
(ii) Intermediate Hamiltonian perturbation theory with only one (lowest) state considered as the main one (Eq. (20)).

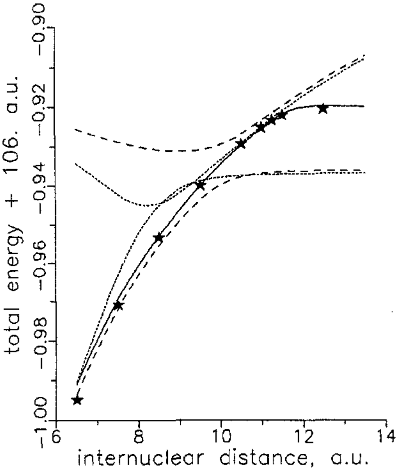


Fig. 1. Calculated potential energy functions for the LiF molecule. Asterisks: full CI [45]. Dotted lines: CASPT2. Solid lines: intermediate Hamiltonian perturbation theory. Dashed lines: effective Hamiltonian QDPT

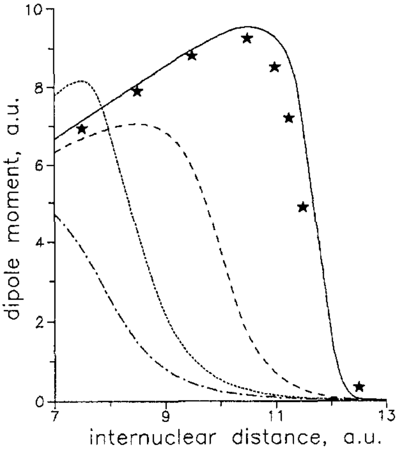


Fig. 2. Ground-state dipole moment function for the LiF molecule. Same designations as in Fig. 1; dashed-dotted line: CASSCF/SA

(iii) Conventional effective Hamiltonian quasi-degenerate perturbation theory (Eq. (11)). In the latter case, the doubly excited configuration was deleted from the model space in order to prevent the appearance of intruder states.

Since the active MO in our calculations were strongly localized, in the two latter cases the zero-order states might be considered as quasi-diabatic, while in the former one we started from an adiabatic description.

The resulting potential curves along with the benchmark full CI energy values from Ref. [45] are plotted in Fig. 1. To visualize the neutral–ionic transfer in the avoided-crossing region, we display in Fig. 2 the ground-state dipole-moment functions computed by the finite-field technique.

The CASPT2 method, which implies the contraction of the reference-space components of the wavefunctions, is obviously unable to improve the description of the avoided crossing. Moreover, the neglect of effective interaction between the perturbed multiconfigurational states leads to violation of the non-crossing rule. This effect had been predicted by Spiegelmann and Malrieu [8] on the basis of a model analysis and has been recently observed in molecular calculations by Nakano [33].

The completely decontracted intermediate Hamiltonian approach reproduces almost quantitatively the full CI results both for potential energy and dipole moment functions, with the error in the avoided-crossing position less than to 0.5 a.u. Certainly this remarkable accuracy of the 2nd-order results can partly be due to cancellation of errors since inclusion of the third-order correction slightly degrades the results (r_{ac} becomes overestimated by 1.0 a.u.) but it does indicate a fairly balanced description of correlation in the two diabatic configurations.

Note that with this particular choice of the zero-order approximation, the resulting curves are smooth. However, in passing from Möller–Plesset to Epstein–Nesbet partitioning of H one can observe a slight change in the slope of the ground state potential function at the point of zero-order level crossing ($r = 7.2$ a.u.) predicted by our model Hamiltonian analysis (Sect. 2.5).

The conventional QDPT provides a qualitatively correct description of the avoided crossing, while r_{ac} is shifted to shorter values due to significant overestimation (by ca. 0.5 eV, or 15%) of the dynamical correlation effects for the neutral structure. This deficiency should be attributed to the inherent tendency of the one-particle Möller–Plesset model to produce too high zero-order energies (and therefore too small energy denominators) for open-shell model states.

3.2 Configuration mixing in the two ${}^1\Sigma^+$ lowest states of CuF

According to the results of large scale CI calculations [47] the ground ($X\,{}^1\Sigma^+$) state of this molecule near the equilibrium is dominated by the closed-shell configuration $(F^-)Cu^+(d^{10})$, which is strongly favored by dynamical correlations, while the excited $B\,{}^1\Sigma^+$ is essentially open-shell ($(F^-)Cu^+d^9s^1$) and significantly less correlated. Within the valence-active-space (2p F, 4s, 3d Cu) state-average CASSCF approximation these correlation effects are neglected and one arrives at a considerable overestimation of configuration mixing (Table 1). As a consequence, in the frame of the contracted CASPT2 approach, one cannot describe correctly the differential correlation energy for the X → B transition, assigning to both states roughly the same amount of dynamical correlation. Moreover, the artificial dependence of the CASSCF mixing coefficients on the internuclear separation leads to aberrations in the shape of the CASPT2 potential curve and underestimation of the ground-state vibrational frequency. In contrast, the shortcomings of the CASSCF/SA zero-order approximation may be substantially corrected by the use of a non-contracted second-order scheme. We employed here a straightforward generalization of the simple intermediate Hamiltonian approach (Eq. (20)) to the case of several non-degenerate main states [37, 39] with Möller–Plesset-type energy denominators. Since we worked in the basis of single Slater determinants, a three-dimensional main model space was chosen to describe properly two singlet states. As one can see from Table 1, this method allowed to rectify the main deficiencies of the CASSCF/SA approximation, reproducing the closed-shell character of the ground-state wavefunction and thus significantly improving the excitation energy estimate and preserving the shape of potential curve. Similar situations with the $d^k s^0 - d^{k-1}s^1 - d^{k-2}s^0$ type configuration mixing controlled by dynamical correlations are perfectly typical for transitional metal compounds (see, for example, the discussion on low-lying states of NiH in [49]); the use of contracted schemes seems to be incompatible with moderate-size MCSCF reference spaces.

Table 1. Calculated characteristics of the two lowest ${}^1\Sigma^+$ states of CuF. Contracted Gaussian basis sets and pseudopotentials from Ref. [47, 48]

	Active-space wavefunction composition $s^0d^{10}:s^1d^9$		$X^1\Sigma^+ - B^1\Sigma^+$ vertical excitation energy (cm^{-1})	Ground state vibrational frequency (cm^{-1})
	$X^1\Sigma^+$	$B^1\Sigma^+$		
CASSCF/SA	59:36	40:55	10347	620
CASPT2	"	"	10435	560
$H^{\mathrm{int}(2)}$	84:14	16:82	15182	617
Exptl.	—	—	19300	623

Table 2. Summary of the existing 2nd-order multireference algorithms

	Convergence	Effect of the outer space on $P_S\psi$	Separability of inactive exc.
QDPT-(2)	Bad (intr. states)	Correct	Exact if $S \equiv$ CAS
Diagon. of $P_S\mathrm{H}P_S$ + PT2			
CIPSI	Good	None	Forced
CASPT2	Good	None	Exact
Diagon. of $P_S\mathrm{H}P_S$ + QDPT2			
CIPSI	Good	Correct in some well-behaved problems only	Forced[a]
Intermediate Hamiltonian $H^{\mathrm{int}(2)}$	Good	Correct	Not satisfied

[a] The inactive double excitations are assumed to cause an uniform energy shift of all the states considered

4. Conclusion

The discussions of the present paper are summarized in Table 2. One sees that, for the time being, there is no satisfactory multiconfigurational second-order perturbative scheme. The QDPT-2 scheme is formally the best when the model space is a CAS, but then it suffers dramatic intruder state problems which can only be masked at low order by level shifts. The techniques consisting in perturbing a single vector resulting from the diagonalization of the model space CI, such as CIPSI and CASPT2, have the major drawback of being contracted, which necessarily results in spurious behaviour near weakly avoided crossings. For the time being, the QDPT-2 version of CIPSI, working in a physically meaningful model space spanned by a few eigenstates of $P_S\mathrm{H}P_S$, seems the most reliable tool, but the separability of inactive double excitations has to be enforced. One may expect that the efficient CASPT2 algorithm will be rewritten promptly in an effective Hamiltonian formalism along the same lines and will become the most satisfactory method. The intermediate Hamiltonian theory in terms of single determinants seemed a promising tool, since it is uncontracted and avoids the intruder state problem, but in its order-by-order development it does not satisfy the elementary separability requirement. Our next paper will give a solution to this difficulty.

Acknowledgement. AZ is grateful for the financial support provided by the French Foreign Affairs Ministry.

References

1. Goldstone J (1957) Proc Roy Soc London Ser A 239:267
2. Möller C, Plesset MS (1934) Phys Rev 46:618
3. Hehre WJ, Pople JA, Radom L, Schleyer PvR (1986) Ab initio molecular orbital theory. Wiley, New York
4. Van Vleck JH (1929) Phys Rev 33:467
5. Bloch C, Horowitz J (1958) Nucl Phys 8:91
6. des Cloizeaux J (1960) Nucl Phys 20:321
7. Malrieu JP, Durand Ph, Daudey JP (1985) J Phys A: Math Gen 18:809
8. Spiegelmann F, Malrieu JP (1984) J Phys B: At Mol Phys 17:1235; 17:1259
9. Brandow BH (1967) Rev Mod Phys 39:771
10. Sandars PGH (1969) Adv Chem Phys 14:365
11. Lindgren I (1974) J Phys B: At Mol Phys 7:2449
12. Kvasnicka V (1977) Adv Chem Phys 36:345
13. Hose G, Kaldor U (1979) J Phys B: At Mol Phys 12:3827
14. Freed KF, Sheppard M (1982) J Phys Chem 86:2130
15. Kaldor U (1985) Intern J Quantum Chem 28:103
16. Durand Ph, Malrieu JP (1987) In: Lawley KP (ed) Ab initio methods of quantum chemistry, Vol 1. Wiley, p 321
17. Zarrabian S, Paldus J (1990) Intern J Quantum Chem 38:761
18. Freed KF (1989) In: Kaldor U (ed) Many-body methods in quantum chemistry. Springer, Berlin, p 1
19. Rawlings DC, Davidson ER (1983) Chem Phys Lett 98:424
20. Hose G (1989) In: Kaldor U (ed) Many-body methods in quantum chemistry. Springer, Berlin, p 43
21. Evangelisti S, Daudey JP, Malrieu JP (1987) Phys Rev A: Gen Phys 35:4930
22. Huron B, Malrieu JP, Rancurel R (1973) J Chem Phys 58:5745
23. Wolinski K, Pulay P (1989) J Chem Phys 90:3647
24. Harrison RJ (1991) J Chem Phys 94:5021
25. Malrieu JP (1982) Theoret Chim Acta 62:163
26. Rubio J, Povill A, Illas F, Malrieu JP (1992) Chem Phys Lett 200:559
27. Anderson K, Malmqvist P, Roos BO, Sadlej AJ, Wolinski K (1990) J Phys Chem 94:5483; Anderson K, Malmqvist P, Roos BO (1992) J Chem Phys 96:1218
28. Morrison JC, Froese Fischer C (1987) Phys Rev A 35:2429
29. Serrano-Andres L, Merchan M, Nebot-Gil I, Lindh R, Roos BO (1993) J Chem Phys 98:3151
30. Sheppard MG, Schneider BI, Martin RL (1983) J Chem Phys 79:1364
31. Spiegelmann F, Gadéa FX, Castex MC (1990) Chem Phys 145:173
32. Audouard E, Spiegelmann F (1991) J Chem Phys 94:6102
33. Nakano H (1993) J Chem Phys 99:7983
34. Lisini A, Decleva P (1992) Chem Phys 168:1
35. Clotet A, Daudey JP, Malrieu JP, Rubio J, Spiegelmann F (1990) Chem Phys 147:293
36. Sheppard MG (1984) J Chem Phys 80:1225
37. Heully JL, Evangelisti S, Durand Ph (1994) Int J Quantum Chem, to be published
38. Zaitsevskii AV (1993) J Phys II France 3:435
39. Zaitsevskii AV (1993) J Phys II France 3:1593
40. Heully JL, Daudey JP (1988) J Chem Phys 88:1046
41. Zaitsevskii AV, Dementev AI (1990) Chem Phys Lett 168:589
42. Zaitsevskii AV, Dementev AI (1991) Optics Commun 86:461
43. Hoffmann MR (1993) Chem Phys Lett 210:193
44. Kozlowski PM, Davidson ER (1994) J Chem Phys 100:3672

45. Bauschlicher CW, Langhoff SR (1988) J Chem Phys 89:4246
46. Andersson K, Blomberg MRA, Fülscher MP, Kellö V, Lindh R, Malmquist P-Aa, Noga J, Olsen J, Ross BO, Sadlej AJ, Sieghban PEM, Urban M, Widmark P-O (1991) MOLCAS, version 2, University of Lund, Sweden
47. Ramirez-Solis A, Daudey JP (1989) Chem Phys 134:111
48. Nodeless pseudopotentials have been included in MOLCAS2 by M Pélissier, Université de Poitiers, France
49. Bauschlicher CW, Langhoff SR, Komornicki A (1990) Theoret Chim Acta 77:263