Theor Chim Acta (1986) 69: 357-368

Theoretica Chimica Acta

Orbital-invariant formulation and second-order gradient evaluation in Møller–Plesset perturbation theory

Peter Pulay and Svein Saebø*

Department of Chemistry, The University of Arkansas, Fayetteville, AR 72701, USA

(Received September 28/Accepted October 10, 1985)

Based on the Hylleraas functional form, the second and third orders of Møller-Plesset perturbation theory are reformulated in terms of arbitrary (e.g., localized) internal orbitals, and atomic orbitals in the virtual space. The results are strictly equivalent to the canonical formulation if no further approximations are introduced. The new formalism permits the extension of the local correlation method to Møller-Plesset theory. It also facilitates the treatment of weak pairs at a lower (e.g., second order) level of theory in CI and coupled cluster methods. Based on our formalism, an MP2 gradient algorithm is outlined which does not require the storage of derivative integrals, integrals with three external MO indices, and, using the method of Handy and Schaefer, the repeated solution of the coupled-perturbed SCF equations.

Key words: Perturbation theory — Møller-Plesset theory — Localized orbitals — Pair correlation — Gradients — MP2 — MP2 gradients

1. Introduction

The simplest theoretical methods for the treatment of dynamical electron correlation are the low orders of Møller-Plesset perturbation theory [1, 2]. In particular, the second order of this theory (MP2) provides the simplest possible treatment for dynamical electron correlation. Although MP2 is not particularly accurate by today's standards, it usually recovers 80–120% of the basis set limit correlation

---

* *Present address*: Department of Chemistry, Mississippi State University, Mississippi State, MS 39762, USA.

energy. It is thus well suited to estimate the magnitude of correlation effects. Gradients for the MP2 energy expression have been formulated by Pople et al. [3], using the traditional MP2 energy expression. In the spin-adapted generator state formalism [4], which will be used for closed shells here, this is given by

$$E_c = -\sum_{i \geq j} (1+\delta_{ij})^{-1} \sum_{ab} K_{ij}^{ab}(4K_{ij}^{ab} - 2K_{ij}^{ab})/(e_a + e_b - e_i - e_j). \tag{1}$$

Here and further in the text $i, j, k, l, \ldots$ denote internal (occupied) orbitals, $a, b, c, d, \ldots$ external (virtual) ones, and

$$K_{ij}^{ab} = (ia|jb) = \iint i(1)a(1)r_{12}^{-1}j(2)b(2)\, dr_1\, dr_2. \tag{2}$$

The e's denote Hartree-Fock orbital energies.

The fact that Eq. (1) is necessarily expressed on terms of canonical molecular orbitals is not advantageous in several respects. Pair correlation energies expressed in canonical orbitals cannot be as easily interpreted as those in localized orbitals. The latter provide a natural decomposition into intraorbital, adjacent interorbital and distant interorbital contributions, each with a clear chemical meaning. This decomposition is not easy to obtain using canonical orbitals. For instance, Kellö et al. [5] have recently analyzed correlation energies of prototype organic molecules in terms of bond increments, using a fit to the total Møller-Plesset correlation energies. A similar analysis has been given for empirical correlation energies by George et al. [6]. Our direct evaluation of localized contributions has significant advantages over such approaches. More importantly, it gives the correlation energy in terms of contributions from pairs of bonds rather than simply from bonds, which is physically much more appropriate but yields too many parameters to fit.

A further motivation for the reformulation of the Møller-Plesset theory in terms of non-canonical orbitals was the desire to provide a simplified treatment for distant (or weak) interpair correlation in our local correlation treatment [7, 8]. In previous applications [8], these pairs were either fully included or completely omitted. Neither of these alternatives is completely satisfactory. Although the computational work in the local correlation treatment is strongly reduced compared with all singles and doubles CI, it is still substantial. It is evident that the weak pairs can be satisfactorily treated at the second order level, or a generalization theoreof. The complete neglect of the weak pairs is not justified, as the sum of these small dispersion contributions may be significant, up to 10% of the total correlation energy. In some cases, e.g. in the study of intermolecular interactions, these small terms constitute the effect sought. The present perturbational results form the basis for a simplified but accurate treatment of weak pairs and the dispersion force which will be published separately.

The non-canonical formulation of MP2 theory has another advantage for the local CI method: it provides a good initial estimate for the CI coefficients. In our initial implementation [8] we have observed a somewhat poorer convergence than in canonical-orbital based CI. Part of this was traced to the poorer initial approximation which neglected large pair coupling terms, caused by off-diagonal Fock matrix elements. A straightforward transformation from canonical to local orbitals is impossible in the local CI where each pair has its own correlation basis.

Finally, an important objective of the present work is to simplify gradient evaluation in MP2 theory. The elimination of explicit reference to virtual orbitals obviates some computational steps present in the algorithm given by Pople et al. [3].

Perturbation theory with non-canonical orbitals (in most cases localized orbitals) has been used in previous work. We shall not consider methods in which the zeroth-order wave function is not exactly the Hartree-Fock determinant [9-11]. Even with the Hartree-Fock zeroth-order function, the choice of the zeroth-order Hamiltonian is still arbitrary [12-14], although the sum of the Fock operators (Møller-Plesset theory) is the overwhelmingly used choice [1, 2].

Our goal is to develop a perturbation theory based on localized orbitals which is exactly identical with the customary Møller-Plesset theory. In variational CI as well as in coupled-pair methods localized internal orbitals do not create any principal difficulty and have used occasionally [15-17]. In particular, Kutzelnigg has analyzed the behavior of various correlation schemes under localization [18]. The results obtained do not depend on the unitary mixing of the internal orbitals in variational CI, CPMET, L-CPMET [19] and ACCD [20, 21]. Robb [22] has given a formulation of the many-body perturbation theory in which the zeroth-order Hamiltonian is modified by projection operators to have localized one-electron eigenfunctions. The results obtained are similar to but not identical with the canonical results. Recently Kapuy et al. [23] have derived a diagrammatic formulation of many-body perturbation theory, using localized orbitals. The use of non-canonical orbitals gives rise to new diagrams which are given through fourth order. As their model results show, the localized results in a given order of perturbation theory are still inferior to the canonical ones. For example, in the limit of weak correlation (beta = −10), fourth-order canonical MBPT recovers about 99.5% of the exact correlation energy while the analogous localized form gives only ≃90%. If the localization diagrams are neglected, the result deteriorates further to ≃80%.

We have not attempted a detailed comparison of our method with Kapuy's theory [23]. It seems evident, however, that our method is a generalization of the latter, in which the localization diagrams are summed up to infinite order. The main result of the present work is probably that it is possible to accomplish this with a small amount of extra work, at least to third or partial fourth order; generalization to full fourth order appears to be also possible.

2. Theory

The starting point of our considerations is the Hylleraas functional form of the second-order energy [24]. For real functions this is given by

$$E_2 = 2\langle\Psi_1|\boldsymbol{H} - E_0|\Psi_0\rangle - \langle\Psi_1|\boldsymbol{H}_0 - E_0|\Psi_1\rangle = \min.$$

Here Ψ_0 and Ψ_1 are the zeroth- and first-order wave functions, $\boldsymbol{H}_0$ is the zeroth-order Hamiltonian, and E_0 is the corresponding energy. The functional form of this expression facilitates the evaluation of response properties.

In Møller–Plesset theory with a closed-shell reference function, $\boldsymbol{H}_0$ is the sum of one-electron Fock operators, and, using the efficient generator state matrix formulation [4, 25], the above expression becomes

$$E_2 = \sum_{i \geq j} e_{ij},$$

$$e_{ij} = 2\langle \boldsymbol{K}_{ij}\tilde{\boldsymbol{C}}_{ji}\rangle + \langle \boldsymbol{F}\boldsymbol{C}_{ij}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ji}\rangle + \langle \boldsymbol{S}\boldsymbol{C}_{ij}\boldsymbol{F}\tilde{\boldsymbol{C}}_{ji}\rangle - \sum_k [F_{ik}\langle \boldsymbol{S}\boldsymbol{C}_{kj}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ji}\rangle + F_{kj}\langle \boldsymbol{S}\boldsymbol{C}_{ik}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ji}\rangle]. \quad (3)$$

The matrix notation used here is the same as in [4] which is, in turn, a modification of the SCEP method [26]. Matrix-formulated CI programs have been developed Dykstra [17], Werner, Reinsch and Meyer [27], Ahlrichs [28], and Saunders [29]. In Eq. (3), the brackets $\langle\,\rangle$ denote a matrix trace, the superscript † indicates a matrix transpose, $\boldsymbol{K}$ is defined in Eq. (2), $\boldsymbol{S}$ is the overlap matrix and $\boldsymbol{F}$ is the Fock matrix. Following the SCEP philosophy, the matrix indices may correspond to arbitrary non-orthogonal functions which span the virtual space. They can be expressed directly in AO basis functions, in which case the dimension of the matrices is somewhat larger and the strong orthogonality condition [26] must hold for the CI coefficients $\boldsymbol{C}$. $\boldsymbol{C}^{st}_{ij}$ is the coefficient of the generator state function as defined in [4]:

$$\Psi^{st}_{ij} = \Phi^{st}_{ij} + \Phi^{\bar{s}\bar{t}}_{\bar{i}\bar{j}} + \Phi^{\bar{s}t}_{\bar{i}j} + \Phi^{s\bar{t}}_{i\bar{j}}.$$

and the contravariant coefficient matrix $\boldsymbol{C}_{ij}$ is defined as

$$\tilde{\boldsymbol{C}}_{ij} = (1 + \delta_{ij})^{-1}[4\boldsymbol{C}_{ij} - 2\boldsymbol{C}_{ji}].$$

with

$$\boldsymbol{C}_{ji} = \boldsymbol{C}^\dagger_{ij}.$$

Straightforward differentiation of the quadratic second-order energy expression, Eq. (3), leads to the following equation for the MP2 residuals

$$\boldsymbol{T}^2_{ij} = \boldsymbol{K}_{ij} + \boldsymbol{F}\boldsymbol{C}_{ij}\boldsymbol{S} + \boldsymbol{S}\boldsymbol{C}_{ij}\boldsymbol{F} - \sum_k \boldsymbol{S}[F_{ik}\boldsymbol{C}_{kj} + F_{kj}\boldsymbol{C}_{ik}]\boldsymbol{S} = \boldsymbol{0}. \quad (4)$$

This equation can be solved iteratively, in complete analogy to the SCEP method [26]. The resulting correlation energy is identical with the canonical MP2 result. At convergence, the quadratic energy expression, Eq. (3), yields the same pair energies $e_{ij} = \langle(\boldsymbol{K}_{ij} + \boldsymbol{T}^2_{ij})\tilde{\boldsymbol{C}}_{ji}\rangle$ as the more customary linear energy expression does

$$E_2 = \sum_{i \geq j} \langle \boldsymbol{K}_{ij}\tilde{\boldsymbol{C}}_{ji}\rangle. \quad (5)$$

Evaluation of the more accurate quadratic energy expression

$$E_3 = \sum_{i \geq j} \langle(\boldsymbol{K}_{ij} + \boldsymbol{T}_{ij})\tilde{\boldsymbol{C}}_{ji}\rangle \quad (6)$$

yields the Møller-Plesset third order energy. Here $\boldsymbol{T}^2$, the second-order residual, has been replaced by the full doubles residual $\boldsymbol{T}$ (the doubles part of Eq. (32) in [4]):

$$\boldsymbol{T}_{ij} = \boldsymbol{K}_{ij} + \boldsymbol{K}(\boldsymbol{C}_{ij}) + \boldsymbol{F}\boldsymbol{C}_{ij}\boldsymbol{S} + \boldsymbol{S}\boldsymbol{C}_{ij}\boldsymbol{F} + \boldsymbol{Q}_{ij}\boldsymbol{S} + \boldsymbol{S}\boldsymbol{Q}^{\dagger}_{ji} + \boldsymbol{S}(\boldsymbol{G}_{ij} + \boldsymbol{G}^{\dagger}_{ji})\boldsymbol{S}, \tag{7}$$

where

$$\boldsymbol{Q}_{ij} = \sum_k [(\boldsymbol{K}_{ik} - 0.5\boldsymbol{J}_{ik})(2\boldsymbol{C}_{kj} - \boldsymbol{C}_{jk}) - 0.5\boldsymbol{J}_{ik}\boldsymbol{C}_{jk} - \boldsymbol{J}_{jk}\boldsymbol{C}_{ik}], \tag{7a}$$

$$\boldsymbol{G}_{ij} = \sum_{k \geq l} (1 + \delta_{kl})^{-1}[(ik|jl) - \delta_{lj}F_{ik} - \delta_{ki}F_{jl}]\boldsymbol{C}_{kl}. \tag{7b}$$

The off-diagonal Fock matrix elements in Eq. (7b) introduce a significant coupling between pair correlation functions, even at the second-order level. The neglect of these contributions in IEPA (Independent Electron Pair Approximation) [14] is probably responsible for the fact that, contrary to intuitive expectation [18], IEPA deteriorates if localized orbitals are used.

In our local correlation program, the zeroth step is the iterative calculation of the local MP2 energy. In the first iteration step, the residuals, Eq. (7) are evaluated; substituted into the quadratic energy expression, Eq. (6), these yield the Møller-Plesset third-order energy. It is possible to improve the doubles (as well as the singles) CI coefficients using the first-order residuals, and calculate a partial fourth-order energy with little extra work [30]. We omit this step because it inteferes with the efficient conjugate gradient-type convergence accelerator. In the first iteration step, the latter amounts to an overall scaling of the correlation function, and, as can easily be shown, yields the [2/1] Padé approximant.

We now return to the approximate treatment of weak pairs in correlation theories. It must be emphasized that only localized theories have weak pairs; in canonical-based theories it is usually not possible to identify weak pairs. A practical definition of weak pairs is that the MP2 pair correlation energy does not exceed 3 mE$_\mathrm{h}$. Our starting point is the functional form of the coupled-cluster doubles theory or an approximation to it:

$$\boldsymbol{E}_c = \sum_{i \geq j} \langle(\boldsymbol{K}_{ij} + \boldsymbol{T}_{ij} + \boldsymbol{v}_{ij})\tilde{\boldsymbol{C}}_{ji}\rangle. \tag{8}$$

Here $\boldsymbol{v}_{ij}$ depends on the many-body method used: it is omitted in L-CPMET [19], and it is $e_{ij}\boldsymbol{C}_{ij}$ in CEPA-2 [15]. For CPMET [19] or ACCD [20, 21] $\boldsymbol{v}_{ij}$ is more complex; a concise form in the generator state formalism was given in [4]. Minimization of Eq. (8) with respect to the CI coefficients yields correlation energies which are practically identical with conventional many-body results for CEPA-2 [31] and for ACCD [32], even though these methods cannot be cast exactly in functional form. We would like to retain the functional form in the approximate theory for weak pairs, for the ease of gradient and other response property evaluation.

The simplest approximation for the weak pairs is the replacement of the weak pair energies in Eq. (8) by the second-order expression, Eq. (3). This amounts to the neglect of all couplings between weak pairs as well as those between weak and strong pairs which are not present at the MP2 level. A much better approximation is obtained if the couplings between the weak and strong pairs, as well as certain diagonal type contributions which are phase-matched and therefore important, are retained in the $\boldsymbol{T}$'s. Implementation of these schemes is underway in our laboratory and will be reported separately [33]; preliminary results are very promising. An advantageous feature of these methods is that the functional form of the correlation energy is retained.

It must be pointed out that the calculation of MP2 energies is non-canonical form is in itself not superior to the traditional method; because of the extra computational steps involved, it is likely that it is somewhat less efficient, although the difference is not large, as much of the computation is spent in the integral transformation section. There are three aspects in which the proposed method is useful. First, it can be used in the local basis set approximation [7, 8]. Second, it allows the simplified treatment of weak pairs in accurate correlation calculations. Third, it suggests ways by which the efficiency of the MP2 gradient calculation can be improved.

3. Second order energy gradients and dipole moments

The second-order correlation energy can be written as $\varepsilon(\boldsymbol{I}, \boldsymbol{C}, \boldsymbol{U})$ where $\boldsymbol{I}$ denotes the set of primitive AO integrals, $\boldsymbol{C}$ the correlation coefficients, and $\boldsymbol{U}$ the coefficients of the occupied SCF orbitals. The energy gradient has three contributions, A, B and C, originating respectively from the derivatives of the integrals, correlation coefficients and the SCF coefficients. The first of these can be written as

$$A = \sum_p (\partial\varepsilon/\partial I_p) I_p^x$$

where the superscript x denotes differentiation with respect to one of the nuclear coordinates. It is useful to treat separately the most important term, A1, arising from matrix traces of the form $\langle \boldsymbol{K}_{ij} C_{ji} \rangle$:

$$\begin{aligned} \mathrm{A1} &= 2 \sum_{pqrs} (pq|rs)^x \sum_{i \geq j} U_{pi} U_{rj} C_{ij}^{qs} \\ &= \sum_{p\geq q\, r\geq s\, pq \geq rs} (pq|rs)^x \sum_{ij} [U_{pi} U_{rj} C_{ij}^{qs} + U_{qi} U_{rj} C_{ij}^{ps} + U_{pi} U_{sj} C_{ij}^{qr} + U_{qi} U_{sj} C_{ij}^{pr}] \\ &= \sum (pq|rs)^x [c_{pr}^{qs} + c_{qr}^{ps} + c_{ps}^{qr} + c_{qs}^{pr}]. \end{aligned} \tag{9}$$

The bracketed term in Eq. (9) is the dominant component of the second-order density matrix in AO basis for the MP2 correlation energy. It can be efficiently evaluated by rearranging the coefficients C_{ij}^{pq} to form matrices in the lower (internal) indices. The transformation to the coefficients in the AO basis [denoted by c in Eq. (9)] can then be efficiently performed as a series of matrix multiplications. Using an intermediate array, the computational effort of this step is proportional to $n^2 N^3$ and $n^3 N^2$ where n and N denote the number of occupied and virtual orbitals, respectively.

Calculation of the remaining contributions from integral derivatives is almost trivial as it requires only the evaluation of Fock matrix elements built from integral derivatives instead of the integrals themselves. Denoting these matrices by F^x, these terms are given by

$$\begin{aligned} \mathrm{A2} = \textstyle\sum_{ij} [\langle(\mathbf{F}^x\mathbf{C}_{ij}\mathbf{S}+\mathbf{F}\mathbf{C}_{ij}\mathbf{S}^x)(4\mathbf{C}_{ji}-2\mathbf{C}_{ij})\rangle - \textstyle\sum_k (\mathbf{F}^x_{ik}\langle\mathbf{S}\mathbf{C}_{kj}\mathbf{S}(4\mathbf{C}_{ji}-2\mathbf{C}_{ij})\rangle \\ +\mathbf{F}_{ik}\langle(\mathbf{S}^x\mathbf{C}_{kj}\mathbf{S}+\mathbf{S}\mathbf{C}_{kj}\mathbf{S}^x)(4\mathbf{C}_{ij}-2\mathbf{C}_{ji})\rangle)]. \end{aligned} \tag{10}$$

Note that $\mathbf{F}^x$ is not equal to the true derivative of the Fock matrix, as the latter contains contributions from the derivatives of the orbitals. Note also the unrestricted summations in Eq. (10).

The most important computational feature of this stage of the proposed algorithm is that no storage of the derivative integrals is needed. The latter is a part of the algorithm used in the GAUSSIAN program system [3]. However, as pointed out earlier [34], this is likely to be a bottleneck in larger calculations.

The contributions from the correlation coefficients appear to vanish at first, as the energy expression, Eq. (3), is minimized with respect to the C's [35]. There are, however, contributions arising from the strong orthogonality condition:

$$\mathbf{CSU} = \mathbf{C}^\dagger\mathbf{SU} = \mathbf{0}.$$

In the AO basis, this condition is violated at neighboring geometries if the correlation coefficients **C** are kept constant. Instead of the usual Lagrangian multipler method, we can restore strong orthogonality explicitly by defining suitable coefficient derivatives $\mathbf{C}^x$, in the spirit of the first author's formulation of the SCF gradients [30]. $\mathbf{C}^x$ is not uniquely defined, as an infinitesimal change in the coefficients in the virtual space has no first-order effect on the energy, by virtue of the minimum property. A simple solution is

$$\mathbf{C}^x_{ij} = \mathbf{V}^\dagger\mathbf{C}_{ij} + \mathbf{C}_{ij}\mathbf{V} \tag{11}$$

where

$$\mathbf{V} = -\mathbf{S}\mathbf{U}^x\mathbf{U}^\dagger - \mathbf{S}^x\mathbf{U}\mathbf{U}^\dagger.$$

A similar expression was recently introduced by Jasien and Dykstra [36]. In their formula $\mathbf{V} = \mathbf{SUU}^{x\dagger}$. This does not contain the derivative of the overlap matrix term, as they consider electrical properties where it is usual to assume that the basis set is independent of the perturbation. The sign reversal follows from the interchange of the order of **U** and $\mathbf{U}^x$. More important is the fact that, as our derivation shows, it is sufficient to consider only the *occupied* block of the SCF coefficient matrix **U** in the construction of $\mathbf{C}^x$. Jasien and Dykstra use the whole (occupied plus virtual) transformation matrix; their formula yields identically zero for $\mathbf{C}^x$ if only the occupied block of **U** is used.

The contribution of the second term (CI coefficient derivatives) to the gradient is

$$B = 2\textstyle\sum_{i\geqslant j}\langle\mathbf{T}^2_{ij}\tilde{\mathbf{C}}^x_{ji}\rangle = 2\textstyle\sum_{i\geqslant j}(1+\delta_{ij})^{-l}\langle\mathbf{C}^x_{ji}(4\mathbf{T}^2_{ij}-2\mathbf{T}^2_{ji})\rangle. \tag{12}$$

Note that $\boldsymbol{T}^2$ in the AO basis does not vanish at convergence; only its projection in the virtual space vanishes. The calculation of the whole $\boldsymbol{T}^2$ is only insignificantly more work than the calculation of its virtual components only. Equations (11, 12) show that the second contribution also arises from the change of the internal orbitals with the geometry, and can thus be treated together with the third contributions.

The latter, denoted by D, has terms arising from the derivatives of the internal exchange operator $\boldsymbol{K}$, and the Fock matrix elements. The first of these, D1, is computationally the most significant and, using the singly transformed integrals which are available from the MP2 procedure, can be calculated as

$$\mathrm{D1}=2\sum_{p,i}\left[\sum_{jqs}(pq\,|\,js)(2C^{qs}_{ij}-C^{qs}_{ji})\right]U^{x}_{pi}. \tag{13}$$

This expression shows that the extended set of transformed integrals $(ac\,|\,bj)$ is not required in the derivative evaluation if both the correlation coefficients $\boldsymbol{C}$ and the SCF coefficient derivatives $\boldsymbol{X}^x$ are expressed in AO basis. It can be obtained in a different way by transforming the gradient expression, given by Pople et al. [3] to AO basis.

There are two more terms in the gradient expression, originating from the derivatives of the Fock matrix elements in Eq. (3). The first of these, arising from matrix traces like $\langle \boldsymbol{F}\boldsymbol{C}_{ij}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ji}\rangle$, is given by

$$\mathrm{D2}=\sum_{k,r}U^{x}_{rk}\sum_{pq}[G_{pqrk}\sum_{i\geq j}(\boldsymbol{C}_{ij}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ji}+\boldsymbol{C}_{ji}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ij})_{qp}, \tag{14}$$

where

$$G_{pqrk}=4(pq\,|\,rk)-(pr\,|\,qk)-(qr\,|\,pk).$$

The final contribution in the gradient comes from the F_{ik} type terms in Eq. (3). Note first that these terms can be written with unrestricted summations as

$$\sum_{ijk}F_{ik}\langle\boldsymbol{S}\boldsymbol{C}_{kj}\boldsymbol{S}(4\boldsymbol{C}_{ji}-2\boldsymbol{C}_{ij})\rangle=\sum_{ik}F_{ik}N_{ik},$$

where $N_{ik}=\sum_j\langle\boldsymbol{S}\boldsymbol{C}_{kj}\boldsymbol{S}\tilde{\boldsymbol{C}}_{ji}\rangle$.

The contribution of this expression to the MP2 gradients is

$$\mathrm{D3}=\sum_{ik}\left[\sum_{rl}G_{ik,rl}U^{x}_{rl}+\sum_{r}(F_{rk}U^{x}_{ri}+F_{ir}U^{x}_{rk}\right]N_{ik}. \tag{15}$$

The sum of B, D1, D2 and D3 can be represented as

$$\mathrm{BD}=\sum_{ri}U^{x}_{ri}Y_{ri} \tag{16}$$

where Y_{ri} is the sum of the coefficients of U_{ri} in these four terms. In the traditional approach, the orbital derivatives $\boldsymbol{U}^x$ are evaluated for all perturbations ^{x}, and contracted with the corresponding $\boldsymbol{Y}$ matrix elements. The expensive part of this calculation is the repeated solution of the coupled-perturbed Hartree–Fock equations. This can be avoided by using the ingenious method introduced very recently by Handy and Schaefer [37]; for another application of the same idea, see Adamowicz et al. [38]. Handy's device can be used to accelerate the MP2 gradient evaluation. Briefly, the orbital derivatives, written in a vector form as $\boldsymbol{u}^x$, are the solutions of a large linear system of equations in which the left-hand side is the same for all perturbations:

$$\boldsymbol{A}\boldsymbol{u}^x = \boldsymbol{b}^x.$$

In the vector form, Eq. (16) is written as

$$\mathrm{BD} = \boldsymbol{y}^\dagger \boldsymbol{u}^x = (\boldsymbol{y}^\dagger \boldsymbol{A}^{-1})\boldsymbol{b}^x = (\boldsymbol{A}^{\dagger -1}\boldsymbol{y})^\dagger \boldsymbol{b}^x. \tag{17}$$

Therefore a single solution of a linear system of equations, with $\boldsymbol{A}^\dagger$ on the left-hand side, suffices to evaluate the derivatives.

The MP2 gradient algorithm described here has not yet been implemented. It is expected that its implementation will significantly speed up the MP2 gradient algorithm. This is important because the bulk of the correlation correction is recovered at the second-order level.

The functional form of the second-order energy is advantageous in defining other response properties, e.g. dipole moments [39]. The formulas given for the gradient are greatly simplified in this case, if, as usual, the basis set is chosen independent of the perturbation, for instance an external electric field. All two-electron integral derivative contributions vanish then, and the only terms remaining are the contributions from the dipole operator $\boldsymbol{d}$ and the SCF orbital derivatives $\boldsymbol{U}^x$.

Table 1. Localized correlation energy contribution in ethylene[a]

Pair	MP2	MP3	CEPA-2	ACCD
hh	18.768	22.193	23.507	23.216
bb	20.266	22.760	24.811	23.690
Intraorbital	115.604	134.292	143.650	140.244
hh' (gem)	10.130	10.572	10.620	10.372
hh' (cis)	0.985	0.968	0.949	0.919
hh' (trans)	1.124	1.099	1.091	1.055
hb	10.478	10.637	10.637	10.308
bb'	27.716	28.436	29.247	27.820
Interorbital	136.018	138.813	139.659	134.976
Total	251.621	273.105	283.309	275.219

[a] Nuclear coordinates in Å units: C (±0.672, 0, 0), H (±1.2432530, ±0.9141958), 6-31G* basis set, five-component *d* functions. The C—H localized orbitals are denoted by *h*, the C—C (banana) orbitals by *b*. Correlation energies (with negative sign) in millihartrees (1 $E_h \approx$ 4.359814 aJ)

The first is

$$\langle \boldsymbol{dA} \rangle - \sum_{ijk} d_{ik} \langle \boldsymbol{S}\boldsymbol{C}_{kj}(4\boldsymbol{C}_{ji} - 2\boldsymbol{C}_{ij}) \rangle,$$

and the second is given by Eqs. (11)-(15), the only change being that now $\boldsymbol{S}^x = \boldsymbol{0}$. The orbital derivative terms are often omitted for CI wave functions, although their contribution may be significant [40, 41].

Table 2. Localized correlation energy contributions in puckered oxetane[a]

Pair	MP2	MP3	ACCD
n	15.940	17.292	17.683
n'	16.051	17.492	19.904
CO	20.893	22.700	23.310
CC	20.279	22.849	23.352
h_1	18.689	22.131	23.168
h_2	18.635	22.052	23.087
h_5	18.238	21.505	22.487
h_6	18.308	21.601	22.583
Total intraorbital	225.529	257.354	266.491
n, n'	20.120	19.722	19.439
n, CO	19.727	18.788	18.313
n', CO	19.020	18.145	18.313
CO, CO′	18.567	17.251	17.697
CO, h_1	7.482	7.607	7.405
CO, h_2	7.529	7.645	7.441
h_1, h_2	10.236	10.671	10.436
CO, CC	7.933	8.053	7.836
h_1, CC	9.911	10.259	9.987
h_2, CC	10.113	10.446	10.162
CC, CC′	10.908	11.164	10.845
CC, h_5	10.249	10.530	10.244
CC, h_6	10.058	10.351	10.069
h_5, h_6	10.347	10.699	10.449
Strong interorbital	284.458	283.826	276.602
Weak interorbital	35.400	32.960	31.797
Total	545.387	574.140	574.892

[a] 6-31G* basis, geometry 0.72 rad puckered (see Banhegyi Gy, Pulay P, Fogarasi G (1983) Spectrochim Acta 39A:761. Five-component *d* functions have been used, all energies in millihartrees. *n* and *n*′ denote the upper and lower O lone pairs, CO and CC the corresponding bonding pairs, h_1 and h_2 are the upper and lower alpha C—H bonds, h_5 and h_6 are the upper and lower beta C—H bonds. The local basis set consists of the valence AO's of the atoms on which the orbitals are localized. In the interorbital case, the union of the two local basis sets is used

3. Results

Table 1 shows the localized pair energies for ethylene at the MP2, MP3 [1, 2], CEPA-2 [15] and ACCD [20] levels, using the 6-31G* basis set [42]. The fast Boys localization [43] was used in all calculations presented here. In the future, we plan to compare these results with those obtained using other localization criteria, in particular with the Edmiston-Ruedenberg method [44]. It is interesting to note that while the MP2 intraorbital correlation energies are quite poor (the total intraorbital correlation energy at the MP2 level is 82.4% of the ACCD value), the interorbital ones are much closer to the more accurate ACCD model. The total third-order correlation energy is a good approximation to the ACCD energy. However, this is a result of cancellation: MP3 overestimates the interorbital correlation energies and underestimates the intraorbital ones.

Results for puckered oxetane (trimethylene oxide) using the local correlation basis set approximation [7, 8] are shown in Table 2. We have repeated the MP2 calculations using the full basis set: this yields a total correlation energy of $-556.568\ \mathrm{mE}_h$, i.e. the local basis set recovers 97.99% of the full basis set MP2 energy. As pointed out in our paper [8], most of this difference can be attributed to the elimination of the intramolecular basis set superposition effect, and is thus beneficial. The trends are very similar to the ethylene case. MP2 suffers from a significant underestimation of the intraorbital correlation contributions while MP3 mildly underestimates the intraorbital and overestimates the interorbital terms, giving an overall good agreement in the total correlation energy as compared with the more sophisticated ACCD model. Two conclusions emerge from the above. First, a moderately accurate but inexpensive correlation method can be defined in which the *localized* intraorbital terms are treated at the MP3 level and all other terms at the MP2 one. Second, the total correlation energy may be a very misleading indicator of the convergence properties of Møller-Plesset perturbation theory.

Acknowledgements. Acknowledgement is made to the Donors of The Petroleum Research Fund, administered by the American Chemical Society, for the partial support of this research. This research was also supported by the National Science Foundation under Grant No. CHE-8500487. Professor W. Meyer (Kaiserslautern, Germany) has pointed out to us the role of the off-diagonal Fock matrix elements in CI convergence based on localized orbitals.

References

1. Bartlett RJ, Silver MD (1975) J Chem Phys 62:325; Bartlett RJ (1981) Ann Rev Phys Chem 32:359
2. Pople JA, Binkley JS, Seeger R (1976) Int J Quantum Chem Symp 10:1
3. Pople JA, Krishnan R, Schlegel HB, Binkley JS (1979) Int J Quantum Chem Symp 13:225
4. Pulay P, Saebo S, Meyer W (1980) J Chem Phys 81:1901
5. Kellö V, Urban M, Noga J, Diercksen GHF (1984) J Am Chem Soc 106:5864
6. George P, Trachtman M, Brett AM, Bock CW (1977) Int J Quantum Chem 12:61
7. Pulay P (1983) Chem Phys Lett 100:151
8. Saebo S, Pulay, P (1985) Chem Phys Lett 113:13
9. Diner S, Malrieu JP, Claverie P (1969) Theor Chim Acta (Berl.) 13:1, 18
10. Laidig W, Purvis III GD, Bartlett RJ (1982) Int J Quantum Chem Symp 16:561

11. Cullen JM, Zerner MC (1982) Int J Quantum Chem 22:497
12. Amos AT, Musher JI (1971) J Chem Phys 54:2380
13. Davidson ER (1973) J Chem Phys 57:1999
14. Kutzelnigg W (1977) In: Schaefer III HF (ed) Methods of electronic structure theory, p 129. Plenum Press, New York
15. Meyer W (1973) J Chem Phys 58:1017
16. Ahlrichs R, Lischka H, Staemmler V, Kutzelnigg W (1975) J Chem Phys 62:1225
17. Dykstra CE, Chiles RA, Garrett MD (1981) J Comp Chem 2:266
18. Kutzelnigg W (1975) In: Chalvet O, Daudel R, Diner S, Malrieu JP (eds) Localization and delocalization in quantum chemistry, p. 143. Reidel, Boston
19. Cizek J (1966) J Chem Phys 45:4256
20. Chiles RA, Dykstra CE (1981) Chem Phys Lett 80:69
21. Janowski K, Paldus J (1980) Int J Quantum Chem 18:1243
22. Robb MA (1975) In: Diercksen GHF, Sutcliffe BT, Veillard A (eds) Computational techniques in quantum chemistry and molecular physics, p 435. Reidel, Boston
23. Kapuy E, Csepes Z, Kozmutza C (1983) Int J Quantum Chem 23:981
24. Hylleraas EA (1930) Z Phys 65:209
25. Matsen FA (1981) Int J Quantum Chem Symp 15:163
26. Meyer W (1976) J Chem Phys 64:2901; Dykstra CE, Schaefer III HF, Meyer W (1976) J Chem Phys 65:2740
27. Meyer W, private communication; Werner H-J, Reinsch EA (1982) J Chem Phys 76:3144
28. Ahlrichs R (1979) Computer Phys Commun 17:31
29. Saunders VR, Lenthe JH van (1983) Mol Phys 48:923
30. Bartlett RJ, Shavitt I (1977) Chem Phys Lett 50:190
31. Pulay P (1983) Int J Quantum Chem Symp 17:257; Pulay P, Saebo S (1985) Chem Phys Lett 117:37
32. Saebø S, Meyer W, Pulay P: to be published
33. Saebø S, Hamilton T, Pulay P: to be submitted
34. Pulay P (1983) J Chem Phys 78:6043
35. Pulay P (1969) Mol Phys 17:197
36. Jasien PG, Dykstra CE (1985) Int J Quantum Chem 28:411
37. Handy NC, Schaefer III HF (1984) J Chem Phys 81:5031
38. Adamowicz L, Laidig WD, Bartlett RJ (1984) Int J Quantum Chem Symp 18:245
39. Bartlett RJ, Purvis GD (1979) Phys Rev A 20:1313
40. Nerbant PO, Roos B, Sadlej AJ (1979) Int J Quantum Chem 15:135
41. Raghavachari K, Pople JA (1981) Int J Quantum Chem 20:1067
42. Hariharan PC, Pople JA (1973) Theor. Chim Acta (Berl) 28:213
43. Boys SF (1968) In: Löwdin PO (ed) Quantum theory of atoms, molecules and the solid state, p. 253. Academic Press, New York 1968
44. Edmiston C, Ruedenberg K (1963) Rev Mod Phys 35:457

Note added in proof

After the submission of this paper, the paper of Handy et al (Handy NC, Amos RD, Gaw JF, Rice JE, Simandiras ED (1985) Chem Phys Lett 120:151) appeared. This paper describes a new algorithm for the evaluation of the first and second derivatives of MP2 energy. Their gradient algorithm is quite similar to the one described in our paper.