ANALYTICAL DIFFERENTIATION OF THE ENERGY CONTRIBUTION DUE TO TRIPLE EXCITATIONS IN FOURTH-ORDER MØLLER–PLESSET PERTURBATION THEORY

Jürgen GAUSS and Dieter CREMER

Lehrstuhl für Theoretische Chemie, Universität Köln, D-5000 Cologne 41, Federal Republic of Germany

Received 15 August 1988

Formulae for the analytical differentiation of the energy contribution due to triple (T) excitations within fourth-order Møller–Plesset (MP4) perturbation theory are derived. Combining these formulae with previously derived formulae for the evaluation of analytical first derivatives of the energy contributions due to single (S), double (D), and quadruple (Q) excitations at MP4 (Chem. Phys. Letters 138 (1987) 131), an algorithm is developed to calculate analytical MP4(SDTQ) energy gradients. Various ways of implementing this algorithm on a computer are discussed and the applicability of the corresponding computer programs is demonstrated by calculating equilibrium geometries, dipole moments, harmonic vibrational frequencies, and infrared intensities for some test molecules at the MP4 (SDTQ) level. The importance of triple excitations for an accurate description of multiple bonds is emphasized.

1. Introduction

Many-body perturbation theory in the form introduced by Møller and Plesset [1] has proven to be a useful method when describing electron correlation effects in atoms and molecules. Møller–Plesset (MP) perturbation theory at second (MP2) and third order (MP3) [2] requires the inclusion of double (D) excitations with respect to the HF reference function into the calculation. At fourth order (MP4), single (S), triple (T), and quadruple (Q) excitations have to be considered in addition to double excitations [3,4]. Since the computation of the correlation correction due to triples is the most time-consuming step, their contributions are often neglected in MP4 calculations [3]. On the other hand, the energy corrections due to triple excitations have been shown to be important for an accurate treatment of electron correlation [5,6]. In some cases, these corrections are even larger than the corrections resulting from single, double, and quadruple excitations. In particular, the description of multiple bonds has been found to be sensitive to the inclusion of triple excitations in MP4 calculations [4,5].

However, a systematic investigation of the effects of triple excitations on molecular properties, e.g. equilibrium geometries, has not been carried out, since this would require a vast amount of computer time aggravated by the fact that, for example, geometry optimizations can only be done with the aid of numerical differentiation procedures. Therefore, it is necessary to develop an analytic differentiation procedure for the MP4(SDTQ) energy which significantly facilitates the calculation of molecular properties such as equilibrium geometry, dipole moment, etc., at the MP4(SDTQ) level.

Analytical techniques for the differentiation of the energy have become powerful tools among modern quantum-chemical methods. They are used to explore potential energy surfaces and to locate energy minima, transition states, and reaction paths [7]. At the HF [8], MC SCF [9], CI [10], quadratic CI [11], and CC [12] levels of theory analytic energy gradients are available and routinely applied. Within MP theory, analytical MP2 [13] and MP3 gradients [14,15] have been developed. For MP4, analytical gradient studies have been

limited so far to MP4(SDQ) [15]. The theoretical background for analytical MP4(SDTQ) energy gradients has been investigated by Bartlett and co-workers [16].

In this paper, formulae for the analytical differentiation of the MP4(T) energy are given and their computational implementation is discussed. In combination with the previously developed MP4(SDQ) gradient program [15], the calculation of analytical MP4(SDTQ) gradients is possible. The applicability of the developed MP4(SDTQ) gradient program is demonstrated by various test calculations.

2. Theory

The energy correction due to triple excitations in fourth-order MP perturbation theory is given by [4]

$$E(\mathrm{MP4(T)}) = \tfrac{1}{36} \sum_{i,j,k} \sum_{a,b,c} w_t(ijk, abc)\, d_t(ijk, abc)\,. \tag{1}$$

In eq. (1) the arrays $w_t(ijk, abc)$ and $d_t(ijk, abc)$ are defined as

$$\begin{aligned} w_t(ijk, abc) = & \sum_d [a(ij, ad)\langle bc\|dk\rangle + a(ij, bd)\langle ca\|dk\rangle + a(ij, cd)\langle ab\|dk\rangle + a(ki, ad)\langle bc\|dj\rangle \\ & + a(ki, bd)\langle ca\|dj\rangle + a(ki, cd)\langle ab\|dj\rangle + a(jk, ad)\langle bc\|di\rangle + a(jk, bd)\langle ca\|di\rangle + a(jk, cd)\langle ab\|di\rangle] \\ & + \sum_m [a(im, ab)\langle cm\|jk\rangle + a(im, bc)\langle am\|jk\rangle + a(im, ca)\langle bm\|jk\rangle + a(jm, ab)\langle cm\|ki\rangle \\ & + a(jm, bc)\langle am\|ki\rangle + a(jm, ca)\langle bm\|ki\rangle + a(km, ab)\langle cm\|ij\rangle \\ & + a(km, bc)\langle am\|ij\rangle + a(km, ca)\langle bm\|ij\rangle] \end{aligned} \tag{2}$$

and

$$d_t(ijk, abc) = w_t(ijk, abc)/(\epsilon_i + \epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b - \epsilon_c)\,, \tag{3}$$

where $a(ij, ab)$ denotes the first-order change of the wavefunction in MP perturbation theory,

$$a(ij, ab) = \langle ij\|ab\rangle/(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)\,, \tag{4}$$

$\langle pq\|rs\rangle$ the antisymmetrized two-electron integrals,

$$\langle pq\|rs\rangle = \int \varphi_p(1)\,\varphi_q(2)\,(1/|r_1 - r_2|)[\varphi_r(1)\,\varphi_s(2) - \varphi_r(2)\,\varphi_s(1)]\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\,, \tag{5}$$

and ϵ_p the orbital energies of the occupied and unoccupied spin orbitals φ_p in the HF reference function. Note that in all the formulae labels $i, j, k, \ldots$ refer to occupied and labels $a, b, c, \ldots$ to unoccupied orbitals. If the type of the orbitals is not further specified, indices $p, q, r, \ldots$ are used.

Differentiation of eq. (1) with respect to an external perturbation parameter λ, e.g. the displacement of a nuclear coordinate or the component of a static electric field, yields the following expression for the MP4(T) energy gradient:

$$\begin{aligned} \mathrm{d}E(\mathrm{MP4(T)})/\mathrm{d}\lambda = & \tfrac{1}{2} \sum_{i,j} \sum_{a,b} \langle ij\|ab\rangle^{\lambda} v_t(ij, ab)/(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b) \\ & - \sum_{i,j,k} \sum_{a,b} a(ij, ab)\epsilon^{\lambda}_{ik} v_t(kj, ab)/(\epsilon_k + \epsilon_j - \epsilon_a - \epsilon_b) + \sum_{i,j} \sum_{a,b,c} a(ij, ab)\epsilon^{\lambda}_{ac} v_t(ij, cb)/(\epsilon_i + \epsilon_j - \epsilon_c - \epsilon_b) \\ & + 2 \sum_{i,j,k} \sum_{a} \langle ij\|ka\rangle^{\lambda} r(ijk, a) + 2 \sum_{i} \sum_{a,b,c} \langle ia\|bc\rangle^{\lambda} s(i, abc) - \sum_{i,j} t(i, j)\epsilon^{\lambda}_{ij} + \sum_{a,b} t(a, b)\epsilon^{\lambda}_{ab}\,. \end{aligned} \tag{6}$$

The arrays $v_t(ij, ab)$, $r(ijk, a)$, $s(i, abc)$, and $t(p, q)$ in eq. (6) are independent of the perturbation λ. They are defined by

$$v_t(ij, ab) = \tfrac{1}{2} \sum_k \sum_{c,d} [\langle cd\|bk\rangle d_t(ijk, acd) - \langle cd\|ak\rangle d_t(ijk, bcd)]$$

$$+ \tfrac{1}{2} \sum_{k,l} \sum_c [\langle cj\|kl\rangle d_t(ikl, abc) - \langle ci\|kl\rangle d_t(jkl, abc)] , \quad (7)$$

$$r(ijk, a) = \tfrac{1}{4} \sum_l \sum_{b,c} a(kl, bc)\, d_t(ijl, abc) , \quad (8)$$

$$s(i, abc) = \tfrac{1}{4} \sum_{j,k} \sum_d a(jk, ad)\, d_t(ijk, bcd) \quad (9)$$

and

$$t(p, q) = \tfrac{1}{12} \sum_{k,l} \sum_{a,b,c} d_t(ikl, abc)\, d_t(jkl, abc) , \quad p=i, q=j ,$$

$$= \tfrac{1}{12} \sum_{i,j,k} \sum_{c,d} d_t(ijk, acd)\, d_t(ijk, bcd) , \quad p=a, q=b . \quad (10)$$

The total derivatives $\langle pq\|rs\rangle^\lambda$ of the two-electron integral $\langle pq\|rs\rangle$ are given by

$$\langle pq\|rs\rangle^\lambda = \sum_{\mu,\nu,\sigma,\rho} \langle \mu\nu\|\sigma\rho\rangle^\lambda c_{\mu p} c_{\nu q} c_{\sigma r} c_{\rho s} + \sum_t (U^\lambda_{tp} \langle tq\|rs\rangle + U^\lambda_{tq} \langle pt\|rs\rangle + U^\lambda_{tr} \langle pq\|ts\rangle + U^\lambda_{ts} \langle pq\|rt\rangle) , \quad (11)$$

where $\langle \mu\nu\|\sigma\rho\rangle^\lambda$ denotes the derivative of the antisymmetrized AO integral $\langle \mu\nu\|\sigma\rho\rangle$ and U^λ_{pq} the first-order change of the MO coefficient $c_{\mu q}$ due to the perturbation λ. The coefficients U^λ_{pq} are obtained by solving the coupled-perturbed HF (CPHF) equations [17]. They can be given in terms of the derivatives F^λ_{pq} of the Fock matrix and the derivatives S^λ_{pq} of the overlap matrix. To avoid singularities in the calculation of U^λ_{pq}, the derivatives ϵ^λ_{pq} of the orbital energies are not chosen to be diagonal [18].

In analogy to the formulae given for the MP3- and MP4-(SDQ) energy gradient [15] eq. (6) for the MP4(T) gradient can be rearranged into the following form:

$$dE(\mathrm{MP4(T)})/d\lambda = \sum_{\mu,\nu,\sigma,\rho} T_{\mu\nu\sigma\rho} \langle \mu\nu\|\sigma\rho\rangle^\lambda + \sum_i \sum_a (U^\lambda_{ai} L'_{ai} + U^\lambda_{ia} L''_{ia})$$

$$+ \sum_{i,j} (-\tfrac{1}{2} S^\lambda_{ij} L'_{ij} + \epsilon^\lambda_{ij} K'_{ij}) + \sum_{a,b} (-\tfrac{1}{2} S^\lambda_{ab} L''_{ab} + \epsilon^\lambda_{ab} K''_{ab}) , \quad (12)$$

where the factors $T_{\mu\nu\sigma\rho}$, L'_{pi}, L''_{pa}, K'_{ij}, and K''_{ab} are independent of λ. Explicit formulae for these factors are easily obtained by substituting eq. (11) into eq. (6).

3. Implementation

Programs for the analytical evaluation of MP4(T) energy gradients based on the formulae given in section 2 have been written and implemented in the existing MP4(SDQ) gradient program [15] within the program system COLOGNE [19] [#1].

The critical step of a MP4(T) gradient calculation is the evaluation of the arrays $v_t(ij, ab)$, $r(ijk, a)$, $s(i, abc)$, and $t(p, q)$ which involves the amplitudes $d_t(ijk, abc)$ of the triple excitations (compare eqs. (7) to (10)). In a first version of the program, the triple amplitudes are stored on magnetic disk and retrieved when needed for the evaluation of the various terms. The evaluation of $v_t(ij, ab)$, $r(ijk, a)$, $s(i, abc)$, and $t(p, q)$ is of $\mathcal{O}(N^7)$ where N denotes the number of basis functions, and thus the most expensive step of a

[#1] COLOGNE is a program system developed for a CDC Cyber 176 computer and contains parts of GAUSSIAN 82 [20].

MP4(SDTQ) gradient calculation. An analytic MP4(SDTQ) gradient calculation requires about 2–3 times the costs of a single MP4(SDTQ) energy calculation and, therefore, it is preferable compared to numerical differentiation procedures.

If stored, the number of $d_t(ijk, abc)$ elements limits the applicability of the program. For a modern super computer with several Gbyte of disk space, MP4(SDTQ) gradient calculation with up to 60 basis functions, e.g. DZ+P calculations on molecules with three heavy atoms, are feasible with such a program. The amount of disk space needed will be 10 to 50 Mword in dependence of the ratio of occupied to virtual orbitals. On the other hand, TZ+2P calculations on three- and four-heavy-atom systems with more than 70 basis functions require several hundred Mword of disk space and, therefore, will not be possible on a routine basis.

The second version of our MP4(T) gradient program is based on a direct algorithm for the evaluation of $v_t(ij, ab)$, $r(ijk, a)$, $s(i, abc)$, and $t(p, q)$, which does not require storage of the triple amplitudes. The arrays $v_t(ij, ab)$, $r(ijk, a)$, and $s(i, abc)$ depend linearly on $d_t(ijk, abc)$ and, thus, are easily computed within a direct algorithm: After a batch of $d_t(ijk, abc)$ elements has been calculated, the corresponding $v_t(ij, ab)$, $r(ijk, a)$, and $s(i, abc)$ elements are obtained by multiplying $d_t(ijk, abc)$ with the appropriate antisymmetrized integrals and amplitudes $a(ij, ab)$, respectively.

The t matrices are obtained by summation over products of two $d_t(ijk, abc)$ elements differing in one index. A direct evaluation of these matrices is troublesome. In the original MP4(SDTQ) program of Pople and coworkers [4,20], all triple amplitudes for fixed lables a, b, c are calculated and processed together. Accordingly, a direct evaluaton of $v_t(ij, ab)$, $r(ijk, a)$, $s(i, abc)$, and $t(i, j)$ is possible within this computation scheme. In order to overcome the problem of calculating for the construction of $t(a, b)$ simultaneously two triple amplitudes which differ in at least one virtual orbital index, we recompute the $d_t(ijk, abc)$ array. In the second computation of $d_t(ijk, abc)$, all elements for fixed i, j, k are obtained together, and, accordingly, $t(a, b)$ can easily be formed. In this way, the storage of the triple amplitudes is avoided at the cost of additional $\mathcal{O}(N^7)$ operations. MP4(SDTQ) gradient calculation will now require about 3–4 times the computer time of a single energy calculation, which still compares favorable with a numerical determination of MP4(SDTQ) gradients. In contrast to our first MP4(SDTQ) gradient program, our second program is not limited by the magnetic disk space available. Therefore, it is advantageously applied in large scale calculation with more than 50 basis functions.

4. Applications

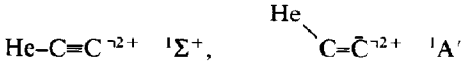
To demonstrate the applicability of the MP4(SDTQ) gradient program, we calculated for NH_3 the equilibrium geometry, the dipole moment, harmonic vibrational frequencies, and infrared intensities at the MP4(SDTQ) level using analytically evaluated energy gradients and the 6-31G(d) basis [21]. The dipole moment is determined as the first derivative of the energy with respect to a static electric field. Force constants and dipole moment derivatives were computed by numerical differentiation of the analytically determined gradients. Table 1 summarizes the MP4(SDTQ) results together with the corresponding data obtained at the HF, MP2, MP3, and MP4(SDQ) level. The inclusion of triple excitations in the MP4 calculation has only small effects on the molecular parameters of NH_3, which is in line with previous results on molecules with single bonds only [22].

A different observation is made for molecules with multiple bonds. As an example, we have calculated the equilibrium geometry of a helium–carbon dication with a triple bond. Helium chemistry has recently become an attractive research field for computational studies [23,24]. A considerable number of helium containing mono- and di-cations have been predicted with the aid of ab initio calculations as stable or metastable. In principle, these cations and dications should be experimentally observable. The dication $HeCC^{2+}$, for which two metastable isomers (**1** and **2**) have been found on the singlet potential energy surface [24], is one example in this area.

Table 1
Theoretical energies, geometries, dipole moments, harmonic vibrational frequencies, and infrared intensities of NH_3 [a]

	HF/6-31G(d)	MP2/6-31G(d)	MP3/6-31G(d)	MP4(SDQ)6-31G(d)	MP4(SDTQ)/6-31G(d)
r_{NH}	1.003	1.017	1.017	1.019	1.021
α_{HNH}	107.2	106.4	106.3	106.1	105.9
E	−56.18436	−56.35738	−56.36894	−56.37192	−56.37429
μ	1.920	1.965	1.955	1.958	1.957
ν_1	3822	3661	3648	3611	3591
ν_2	3689	3504	3510	3470	3446
ν_3	1850	1756	1758	1753	1749
ν_4	1209	1160	1172	1178	1179
I_1	0.9	1.4	0.0	0.3	0.9
I_2	0.3	0.1	0.3	0.7	1.1
I_3	42.7	39.9	39.3	36.2	34.3
I_4	218.3	188.3	183.8	174.2	167.8

[a] Energies in hartree, distances in Å, angles in deg, dipole moments μ in D, vibrational wave numbers ν in cm^{-1}, and intensities I in km/mol.



1 2

Previous calculations at the HF and MP2 level have shown [24] that electron correlation plays an important role in the determination of relative stabilities and geometries of the two isomers of $HeCC^{2+}$. Thus, we have reoptimized the equilibrium geometry of **1** and **2** at the MP4(SDQ) and MP4(SDTQ) level of theory using the 6-31G(d, p) basis [21] and analytically evaluated forces. Table 2 summarizes the MP4 results together with the previously published MP2 data [24].

At all correlation-corrected levels, **2** is predicted to be more stable than **1**. The energy difference increases from 39.0 kcal/mol at MP2/6-31G(d, p) to 50.7 at MP4(SDQ)/6-31G(d, p) and 56.6 kcal/mol at MP4(SDTQ)/6-31G(d, p), thus showing the importance of higher-order corrections and in particular of triple excitations in fourth-order MP perturbation theory. The geometries of **1** and **2** vary significantly with the the-

Table 2
Theoretical energies and geometries of the two isomers **1** and **2** of $HeCC^{2+}$ [a]

	MP2/6-31G(d, p) [b]	MP4(SDQ)/6-31G(d, p)	MP4(SDTQ)/6-31G(d, p)
$HeCC^{2+}(^1\Sigma^+(4\pi))$, **1**			
E	−77.24208	−77.26255	−77.27452
r_{CC}	1.199	1.193	1.204
r_{CHe}	1.082	1.090	1.091
$HeCC^{2+}(^1A'(2\pi))$, **2**			
E	−77.30438	−77.34342	−77.36470
r_{CC}	1.423	1.447	1.447
r_{CHe}	1.409	1.420	1.415
α_{HeCC}	93.6	99.0	97.2
$\Delta E = E(1) - E(2)$	39.1	50.7	56.6

[a] Energies in hartree, distances in Å, angles in deg, and energy differences in kcal/mol.
[b] See ref. [24].

oretical level employed in the calculation. Going from MP2 to MP4 the CHe bonds in **1** and **2** are lengthened by about 0.01 Å. However, the main effect is due to MP4(SDQ) and triple excitations have relatively small effects on the CHe bond distances. Similarly, the CC bond in **2** is increased by 0.025 Å compared to the MP2 result when the MP4(SDQ) method is used. The triples do not lead to a further change in the bond length.

The importance of triple excitations is revealed by the CC distances obtained for **1** at the various levels. Compared to the MP2 result, MP4(SDQ) shortens the CC triple bond in **1** by 0.006 Å. On the other hand, inclusion of triple excitations lengthens the CC bond in **1** by 0.011 Å (!). This result suggests that in order to get reasonable r_e values for multiple bonds triple excitations have to be considered. However, it cannot be excluded that higher-order corrections (MP5) reduce the MP4(T) effect slightly, similar to the changes found for MP2 and MP3.

The MP4(SDTQ) gradient calculations for the two helium compounds took, on a CDC Cyber 176, about 2.3 times the computer time of the preceding MP4 (SDTQ) energy calculation, which underlines the usefulness of the analytical MP4(SDTQ) energy gradients.

Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. All calculations have been carried out on the CDC Cyber 176 of the Rechenzentrum der Universität Köln.

References

[1] C. Møller and M.S. Plesset, Phys. Rev. 46 (1934) 618.
[2] J.A. Pople, J.S. Binkley and R. Seeger, Intern. J. Quantum Chem. Symp. 10 (1976) 1.
[3] R. Krishnan and J.A. Pople, Intern. J. Quantum Chem. 14 (1978) 91.
[4] R. Krishnan, M.J. Frisch and J.A. Pople, J. Chem. Phys. 72 (1980) 4244.
[5] M.J. Frisch, R. Krishnan and J.A. Pople, Chem. Phys. Letters 75 (1980) 66.
[6] R.J. Bartlett, H. Sekino and G.D. Purvis, Chem. Phys. Letters 98 (1983) 66.
[7] G. Fitzgerald, R.J. Harrison and R.J. Bartlett, J. Chem. Phys. 85 (1986) 5143.
[8] P. Pulay, Mol. Phys. 17 (1969) 197.
[9] J.D. Goddard, N.C. Handy and H.F. Schaefer III, J. Chem. Phys. 71 (1979) 1525;
S. Kato and K. Morokuma, Chem. Phys. Letters 65 (1979) 19.
[10] R. Krishnan, H.B. Schlegel and J.A. Pople, J. Chem. Phys. 72 (1980) 4654;
B.R. Brooks, W.D. Laidig, P. Saxe, J.D. Goddard, Y. Yamaguchi and H.F. Schaefer III, J. Chem. Phys. 72 (1980) 4652.
[11] J. Gauss and D. Cremer, Chem. Phys. Letters 150 (1988) 280.
[12] G. Fitzgerald, R.J. Harrison, W.D. Laidig and R.J. Bartlett, Chem. Phys. Letters 117 (1985) 433;
A.C. Scheiner, G.E. Scuseria, J.E. Rice, T.J. Lee and H.F. Schaefer III, J. Chem. Phys. 87 (1987) 5361;
G.E. Scuseria and H.F. Schaefer III, Chem. Phys. Letters 146 (1988) 23.
[13] J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, Intern. J. Quantum Chem. Symp. 13 (1979) 225.
[14] G. Fitzgerald, R. Harrison, W.D. Laidig and R.J. Bartlett, J. Chem. Phys. 82 (1985) 4375;
E.A. Salter, G.W. Trucks, G. Fitzgerald and R.J. Bartlett, Chem. Phys. Letters 141 (1987) 61.
[15] J. Gauss and D. Cremer, Chem. Phys. Letters 138 (1987) 131.
[16] G. Fitzgerald, R.J. Harrison and R.J. Bartlett, J. Chem. Phys. 85 (1986) 5143.
[17] J. Gerratt and I.M. Mills, J. Chem. Phys. 49 (1968) 1719.
[18] N.C. Handy, R.D. Amos, J.F. Gaw, J.E. Rice and E.D. Simandiras, Chem. Phys. Letters 100 (1985) 151.
[19] J. Gauss and D. Cremer, COLOGNE, Universität Köln (1988).
[20] J.S. Binkley, M.J. Frisch, D.J. DeFrees, K. Raghavachari, R.A. Whiteside, H.B. Schlegel, E.M. Fluder and J.A. Pople, GAUSSIAN 82, Carnegie-Mellon University, Pittsburgh (1985).
[21] P.C. Hariharan and J.A. Pople, Theoret. Chim. Acta 28 (1973) 213.
[22] W.J. Hehre, L. Radom, P. von R. Schleyer and J.A. Pople, *Ab initio molecular orbital theory* (Wiley, New York, 1986) ch. 6.
[23] W. Koch and G. Frenking, J. Chem. Soc. Chem. Commun. (1986) 1095; Intern. J. Mass Spectrom. Ion Processes 74 (1986) 133;
W. Koch, J.R. Collins and G. Frenking, Chem. Phys. Letters 132 (1986) 330.
[24] W. Koch, G. Frenking, J. Gauss, D. Cremer and J.R. Collins, J. Am. Chem. Soc. 109 (1987) 5917.