ANALYTICAL DIFFERENTIATION OF THE ENERGY CONTRIBUTION DUE TO TRIPLE EXCITATIONS IN QUADRATIC CONFIGURATION INTERACTION THEORY

Jürgen GAUSS and Dieter CREMER
Lehrstuhl für Theoretische Chemie, Universität Köln, D-5000 Cologne 41, Federal Republic of Germany and Theoretical Chemistry, University of Göteborg, S-412 96 Göteborg, Sweden

Received 26 July 1989

Formulæ for the analytical differentiation of the energy contribution due to triple excitations (T) within quadratic configuration interaction (QCI) theory are derived. Combining these formulæ with previously derived formulæ for the evaluation of analytical first derivatives for QCI theory with single (S) and double excitations (D), an algorithm is developed to calculate analytical QCISD(T) energy gradients. The applicability of this algorithm is demonstrated by calculating the equilibrium geometry of CH₂OO at the QCISD(T)/6-31G(d, p) level of theory.

1. Introduction

Pople, Head-Gordon and Raghavachari introduced a new method for the calculation of the electron correlation energy starting from the single determinant approach of Hartree–Fock (HF) theory [1]. The authors called this method quadratic configuration interaction (QCI) since they derived it from normal configuration interaction theory by adding new terms which are quadratic in the configuration coefficients. The quadratic terms ensure size-consistency in the total molecular energy [1]. When only single (S) and double (D) excitations are included in the QCI approach (QCISD), it is closely related to coupled cluster (CC) theory at the CCSD level [2]. Pople and co-workers have stressed the simplicity of QCISD as compared to CCSD and have considered QCISD as an intermediate approach between CC and CI theory. The relationship between QCISD and CCSD has been analyzed by Paldus, Čiček and Jeziorski [3] and the computational requirements of the two methods by Scuseria and Schaefer [4].

In their original paper, Pople and co-workers also introduced a useful approximation for handling triple excitations (T) at the QCI level of theory [1]. They treat the triple excitations as a small perturbation of the solution obtained at the QCISD level. Perturbation theory yields for the energy correction due to the triple the term ΔE(T) (see below) which is simply added to the QCISD energy to yield a QCISD(T) energy. Comparison of QCISD(T) results with those of full CI and other methods have led to promising results in a number of cases [1].

We have presented analytical formulæ for the energy gradient within QCISD theory [5]. We have also developed expressions for the analytical gradient of the energy contribution due to triple excitations within fourth-order Møller–Plesset (MP4) perturbation theory [6]. We have implemented these formulæ in a computer program and have discussed the advantage of using analytical gradients for QCISD and MP4(SDTQ) theory [5,6]. In this work we extend our previous results by presenting the theory for the analytical differentiation of the triple correction ΔE(T) within QCISD(T) theory introduced by Pople and co-workers [1].
2. Theory

The energy correction due to triple excitations in QCISD(T) theory is given by

\[ \Delta E(T) = \frac{1}{3} \sum_{ijk} \sum_{abc} d(ijk, abc) \left[ w(ijk, abc) + 2\tilde{w}(ijk, abc) \right]. \]  

The matrix elements in eq. (1) are defined as

\[ w(ijk, abc) = \sum_d \left[ a_{id}^a \langle bd|dk \rangle + a_{id}^b \langle ca|dk \rangle + a_{id}^c \langle ab|dk \rangle + a_{id}^d \langle ca|dj \rangle + a_{id}^d \langle ca|dj \rangle \right] \]

\[ + \sum_m \left[ a_{im}^{ab} \langle cm|jk \rangle + a_{im}^{bc} \langle am|jk \rangle + a_{im}^{bc} \langle bm|jk \rangle + a_{im}^{bc} \langle cm|ki \rangle + a_{im}^{bc} \langle am|ki \rangle \right] \]

\[ + a_{im}^{ab} \langle bm|ki \rangle + a_{im}^{bc} \langle cm|i j \rangle + a_{im}^{bc} \langle am|i j \rangle + a_{im}^{bc} \langle bm|i j \rangle, \]

\[ d(ijk, abc) = \frac{w(ijk, abc)}{\epsilon_e + \epsilon_i - \epsilon_a - \epsilon_b - \epsilon_c}, \]

\[ \tilde{w}(ijk, abc) = a_i^a \langle jk|bc \rangle + a_i^b \langle jk|ca \rangle + a_i^c \langle jk|ab \rangle + a_i^a \langle ki|bc \rangle + a_i^a \langle ki|ca \rangle \]

\[ + a_i^a \langle ki|ab \rangle + a_i^a \langle ij|bc \rangle + a_i^a \langle ij|ca \rangle + a_i^a \langle ij|ab \rangle, \]

\[ \tilde{d}(ijk, abc) = \frac{\tilde{w}(ijk, abc)}{\epsilon_e + \epsilon_i - \epsilon_a - \epsilon_b - \epsilon_c}. \]

In eqs. (2) to (5), the symbols \( a_i^a \) and \( a_i^{ab} \) represent the converged QCISD amplitudes of single and double excitations, respectively. The double-bar integrals \( \langle pq|rs \rangle \) correspond to antisymmetrized two-electron integrals in the spin-orbital representation:

\[ \langle pq|rs \rangle = \int \varphi_p^*(1) \varphi_q^*(2) \varphi_r(1) \varphi_s(2) |r_1 - r_2|^{-1} d\tau_1 d\tau_2, \]

with \( \varphi_i \) being a HF spin orbital and \( \epsilon_i \) being the corresponding orbital energy. The labels \( i, j, k, \ldots \) denote occupied orbitals, labels \( a, b, c, \ldots \) virtual orbitals and labels \( p, q, r, \ldots \) any orbital either occupied or not.

Differentiation of eq. (1) with respect to an external perturbation parameter \( \lambda \), e.g. the displacement of a nuclear coordinate or the component of a static electric field, leads to the following expression for the gradient of the energy correction \( \Delta E(T) \) in QCISD(T) theory:

\[ \frac{d[\Delta E(T)]}{d\lambda} = \frac{1}{3} \sum_{ijk} \sum_{abc} w(ijk, abc) \left[ 2d(ijk, abc) + 2\tilde{d}(ijk, abc) \right] + \frac{1}{3} \sum_{ijk} \sum_{abc} \tilde{w}(ijk, abc) \left[ 2d(ijk, abc) \right] \]

\[ - \frac{1}{12} \sum_{ijkl} \sum_{abc} \epsilon_{ij} d(ijk, abc) \left[ d(ijl, abc) + 2\tilde{d}(ijl, abc) \right] + \frac{1}{12} \sum_{ijkl} \sum_{abc} \epsilon_{ij} d(ijk, abc) \left[ d(ijk, abd) + 2\tilde{d}(ijk, abd) \right], \]

where \( w(ijk, abc) \), \( \epsilon_{ij} \), etc. being the derivative of the elements of \( w, \epsilon, \) etc. with respect to \( \lambda \).

In order to avoid singularities in the evaluation of eq. (7) the derivatives of the Lagrangian multipliers \( \epsilon_{pq} \) rather than those of the canonical orbital energies are calculated [7]. Substitution of eqs. (2) and (4) into eq. (7) leads to

\[ \frac{d[\Delta E(T)]}{d\lambda} = \frac{1}{3} \sum_{ij} \left[ a_{ij}^{ab} \nu_{ij}(i, a) + \frac{1}{2} \sum_a \langle i | a \rangle | b \rangle \langle c | a \rangle \nu_{ij}(i, a) + 2 \sum_i \langle i | a \rangle | b \rangle \langle c | i \rangle s(i, abc) \right] \]

\[ + \sum_i \sum_a \left[ a_{i}^{ab} \nu_{ij}(i, a) + \frac{1}{2} \sum_a \langle i | a \rangle | b \rangle \langle c | a \rangle \nu_{ij}(i, a) + \sum a_{ab} t(a, b) - \sum a_{ij} t(i, j) \right], \]

where \( \left[ a_{ij}^{ab} \right] \) and \( \left[ a_{ij}^{ab} \right] \) denote the derivatives of the amplitudes \( a_{ij}^{ab} \) and \( a_{ij}^{ab} \) with respect to \( \lambda \).

The matrices \( \nu_{ij}, r, s, t, u \) and \( \nu_{ij} \) are defined in the following way:
\[ v_T(ij, ab) = \frac{1}{2} \sum_k \sum_{cd} \left[ \langle cd| bk \rangle \langle d(ijk, acd) + \hat{d}(ijk, acd) \rangle - \langle cd| ak \rangle \langle d(ijk, bcd) + \hat{d}(ijk, bcd) \rangle \right] \]

\[ + \frac{1}{2} \sum_{kl} \sum_c \left[ \langle ci| kl \rangle \langle d(ikl, abc) + \hat{d}(ikl, abc) \rangle - \langle ci| kj \rangle \langle d(jkl, abc) + \hat{d}(jkl, abc) \rangle \right], \]

\[ r(ijk, a) = \frac{1}{2} \sum_j \sum_{bc} a_{ijc}^{ab} \langle d(ijl, abc) + \hat{d}(ijl, abc) \rangle, \]

\[ s(i, abc) = \frac{1}{2} \sum_k \sum_j a_{ijk}^{ab} \langle d(ijk, bcd) + \hat{d}(ijk, bcd) \rangle, \]

\[ t(p, q) = \frac{1}{12} \sum_{ikl} \sum_{abc} d(ikl, abc) \langle d(jkl, abc) + \hat{d}(jkl, abc) \rangle, \quad \text{for } p=i \text{ and } q=j, \]

\[ = \frac{1}{12} \sum_{ijk} \sum_{cd} d(ijk, acd) \langle d(ijk, bcd) + \hat{d}(ijl, bcd) \rangle, \quad \text{for } p=a \text{ and } q=b, \]

\[ u(ij, ab) = \sum_k \sum_c d(ijk, ab) a_{ikc}, \]

\[ v_T(i, a) = \frac{1}{2} \sum_{jk} \sum_{bc} d(ijk, abc) \langle jk| bc \rangle. \]

The formula for the first derivative of \( \Delta E(T) \) with respect to an external perturbation \( \lambda \) is very similar to the previously derived expression for the analytical gradient of the corresponding energy correction within fourth-order Møller–Plesset (MP4) perturbation theory \[61\]. Differences in the two formulae are due to (a) the different definitions of the amplitudes \( a_{ij}^{ab} \) and \( a_{ij}^{ab} \) in QCISD and MP theory and (b) the additional coupling of single and triple excitations present in QCISD theory but absent in MP4 theory. This coupling leads to terms involving \( v_T(i, a) \) and \( u(ij, ab) \) in eq. (8) as well as to the terms \( d(ijk, abc) \) in eqs. (9) to (12).

As has already been stressed in the evaluation of the QCISD gradients \[5\], explicit calculation of the derivatives of the amplitudes \( a_{ij}^{ab} \) and \( a_{ij}^{ab} \) is very time consuming and should be eliminated utilizing the z-vector method of Handy and Schaefer \[8\]. While the coupled perturbed QC equations that are required to determine the derivatives of \( a_{ij}^{ab} \) and \( a_{ij}^{ab} \), in eq. (8) are identical in QCISD and QCISD(T) theory, the corresponding z-vector equations are not. Here, additional terms resulting from the derivatives of \( a_{ij}^{ab} \) and \( a_{ij}^{ab} \) have to be added to the inhomogeneous terms. Within QCISD(T) theory the z-vector equations are given by

\[ \sum_j \sum_k z^{ij}_{k} C_{jk}^{ab} + \sum_{k' < k} \sum_{b < c} z_{k'}^{ij} C_{k'j}^{ab} = v_T(i, a), \]

\[ \sum_k \sum_{c} z_{ik}^{ij} C_{jk}^{ab} + \sum_{k < l} \sum_{c < d} z_{ik}^{ij} C_{kl}^{ab} = \langle ab| ij \rangle + 2v_T(ij, ab). \]

For a definition of the various \( C \) terms in eq. (15) and a full account of the use of the z-vector method in QC gradient theory compare with ref. \[5\].

The contribution of triple excitations to the QCISD(T) gradient comprises two terms, one which is included already in the QCISD energy gradient (after modifying the z-vector equations as described above) and one which is given by additional terms containing derivatives of the two-electron integrals and the orbital energies:

\[ \frac{dE[QCISD(T)]}{d\lambda} = \frac{dE[QCISD]}{d\lambda} + 2 \sum_{ijk} \sum_{a} \langle ij| ka \rangle \langle r(ijk, a) \rangle + 2 \sum_{ij} \sum_{abc} \langle i|ab| bc \rangle \langle s(i, abc) \rangle \]

\[ + \frac{1}{2} \sum_{ij} \sum_{ab} \langle ij|ab \rangle \langle u(ij, ab) \rangle + \sum_{ab} \langle \epsilon_{ij}^{ab} t(a, b) \rangle - \sum_{ij} \langle \epsilon_{ij}^{ab} t(i, j) \rangle. \]

Eq. (16) can be rearranged into a form containing only AO integral derivatives, namely the derivatives of the two-electron integrals \( \langle \mu|\nu| \sigma \rangle \), of the one-electron integrals \( h_{\mu\nu} \), and of the overlap integrals \( S_{\mu\nu} \). In addition, one needs the derivatives \( U_{\mu\nu}^{ab} \) of the spin-orbital coefficients \( \epsilon_{\mu\nu} \) as they are defined in the usual coupled-
perturbed HF (CPHF) theory [9,10]. The solution of the CPHF equations is avoided by again using the \( z \)-vector method [8]. The QCISD(T) energy gradient can be written as
\[
\frac{\text{d}E[\text{QCISD(T)}]}{\text{d}\lambda} = \sum_{\mu \nu \rho} T_{\mu \nu \rho} \langle \mu \nu \rho | \text{op} \rangle + \sum_{\mu \nu} D_{\mu \nu} h^k_{\mu \nu} + \sum_{\mu \nu} C_{\mu \nu} S^k_{\mu \nu},
\]
(17)
where \( T_{\mu \nu \rho}, D_{\mu \nu}, C_{\mu \nu} \) are independent of the perturbation \( \lambda \) but dependent on the solution of the \( z \)-vector equations.

3. Implementation

Computer programs for the analytical evaluation of the energy correction due to triple excitations in QCI theory have been written and combined with the existing programs for the evaluation of the analytical QCISD gradient [5] contained in the program system COLOGNE [11]. Since the evaluation of the various terms of the QCISD(T) gradient is similar to those needed for MP4(T) gradients [6] the implementation follows the same procedure described in our paper on the calculation of the analytical energy gradient for MP4(SDTQ) [6]. In this procedure, storage of the arrays \( d(ijk, abc) \) and \( d(ijk, abc) \) is avoided by using a direct algorithm for the computation of \( v_T(ij, ab), v_T(i, a), r(ijk, a), s(i, abc), t(p, q) \) and \( u(ij, ab) \). This means that the appropriate contributions of the triple amplitudes to these arrays is immediately evaluated when the corresponding triple amplitudes have been determined.

The solution of the \( z \)-vector equations within coupled-perturbed QCI theory as well as the evaluation of the terms for the QCISD gradients has been described in ref. [5]. This procedure has been modified to incorporate the additional inhomogeneous terms in the \( z \)-vector equations. The programs for the analytical evaluation of the QCISD(T) gradients have been carefully checked by comparing results with those obtained by numerical differentiation. Calculation of the analytical QCISD(T) gradient turns out to require 3–4 times the costs of a QCISD(T) single-point calculation. This has also been found in the case of the analytical MP4(SDTQ) gradient [6].

4. Application

To illustrate the applicability and usefulness of our program for the evaluation of analytical QCISD(T) energy gradients we have calculated the equilibrium geometry of carbonyl oxide, \( \text{CH}_2\text{O} \), employing the 6-31G(d, p) basis set [12]. \( \text{CH}_2\text{O} \) is a 1,3-dipolar species that is formed as an important intermediate during the oxidation of unsaturated hydrocarbons by ozone [13,14]. Its electronic properties are best described by the resonance structures 1–6 shown in scheme 1 [15]. Resonance structure 1 corresponds to a 1,3 \( \pi,\pi \)-biradical while resonance structures 2 and 3 describe a 1,3 zwitterion. Structures 4, 5 and 6 are probably of less importance for the wavefunction of \( \text{CH}_2\text{O} \) [15].

Restricted HF calculations exaggerate the zwitterionic character of \( \text{CH}_2\text{O} \) as is clearly revealed by the HF/6-31G(d, p) geometry shown in table 1. The CO bond length is rather short comparable to that found for formaldehyde, \( \text{CH}_2\text{O} \), while the OO bond length is close to that found for OO bonds in peroxides [13,14]. Hence, the HF geometry of \( \text{CH}_2\text{O} \) is best described by resonance structure 2.

Inclusion of correlation effects leads to a lengthening of the CO bond and a shortening of the OO bond. The resulting geometry is typical of 1,3-dipolar species with significant biradical character, i.e. correlation increases the importance of resonance structure 1. Table 1 lists geometries and energies of \( \text{CH}_2\text{O} \) computed with analytical gradients at four different levels of theory, namely MP2 [10], MP4(SDQ) [16], MP4(SDTQ) [6], and QCISD(T). The calculated geometries differ considerably from the HF geometry. Furthermore, they reveal that the heavy atom bond lengths are very sensitive to the method used. Therefore, the predictions of the
At the QCISD(T) level, the CO bond length decreases to 1.287 Å while the OO bond length increases to 1.356 Å. The CO bond lengthening at the QCISD(T) level is consistent with the suggestion that the biradical character (resonance structure 1) dominates the electronic structure of the molecule. This is not unexpected since MP2 is known to overestimate the biradical character of 1,3-dipolar species \([14,17]\). At MP4(SDQ), the CO bond is calculated to be 0.02 Å shorter and the OO bond length 0.03 Å longer than the corresponding MP2 values, reflecting the increased zwitterionic character and a stronger contribution of resonance structure 2 to the wavefunction at this level of theory.

When triple excitations are included at the MP4 level, the geometry of CH₂OO changes in an unexpected way. The alternation in the heavy-atom bond lengths is again reduced due to a lengthening of the CO bond by 0.04 Å and a shortening of the OO bond by 0.02 Å relative to the MP4(SDQ) values. In this way, the CO bond becomes slightly longer than the OO bond, indicative of an admixture of resonance structure 4 to resonance structure 1. This result contradicts the known properties of carbonyl oxides, which suggest an electronic structure of CH₂OO close to either 2 or 3.

An answer to the question as to whether the MP4(SDTQ) results are chemically relevant is given by the QCISD(T) calculations. The QCISD(T)/6-31G(d, p) energy of CH₂OO is 1.64 mhartree below that obtained at the MP4(SDTQ)/6-31G(d, p) level of theory (table 1). The QCISD(T) correlation energy is -0.56151 hartree, which has to be compared with a MP4(SDTQ) correlation energy of -0.56846 hartree. At the QCISD(T) level, the CO bond length decreases to 1.287 Å while the OO bond length increases to 1.356 Å.
(table 1). Hence, the ratio of the heavy-atom bond lengths clearly predicts the importance of resonance structures 2 and 3 for the electronic structure of CH$_2$OO in line with the experimental observations. We conclude that the QCISD(T) description of CH$_2$OO is more reliable than the MP4(SDTQ) description.

Acknowledgement

Valuable assistance by E. Kraka and F. Reichel is acknowledged. This work has been supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. Calculations have been carried out on the CDC CYBER 76M of the Rechenzentrum der Universität Köln and a CRAY XMP 48 of the National Supercomputer Center (NSC), University of Linköping, Sweden. DC thanks the NSC for a generous allotment of computer time.

References