Theor Chim Acta (1996) 93: 87-99

Theoretica Chimica Acta

Analytical first derivatives of the energy in the MNDO half-electron open-shell treatment

S. Patchkovskii, W. Thiel

Organisch-chemisches Institut, Universität Zürich, Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland

Received July 27, 1995 / Final revision received September 22, 1995 / Accepted September 22, 1995

Summary. Using the Z-vector formalism the analytical gradient of the energy in the half-electron open-shell treatment is derived and implemented for semiempirical MNDO-type methods. The computation time is shown to scale as $O(N^3)$ with the size of the system, with the memory requirements growing as $O(N^2)$. The evaluation of the analytical gradient is significantly faster than the half-electron SCF calculation, so that routine full geometry optimizations become possible for large open-shell systems. The approach can easily be extended to the treatment of the small CI expansions typically encountered in semiempirical computations.

Key words: MNDO – Half-electron method – Analytical derivatives – Z-vector

1 Introduction

MNDO [1], AM1 [2], and PM3 [3] are widely used in theoretical studies of molecular structure and reactivity, particular in organic chemistry [4–6]. The recently proposed MNDO/d method [7–10] promises an improved semiempirical description of inorganic compounds including hypervalent molecules and transition metal complexes. Applications of these MNDO-type methods usually involve explorations of multidimensional potential surfaces which are greatly facilitated if the gradient of the energy with respect to the nuclear coordinates can be evaluated efficiently.

For closed-shell restricted Hartree-Fock (RHF) wavefunctions, analytic gradients have long been available in MNDO [11], but a simple finite-difference procedure with a constant density matrix and recalculated two-center integrals is also efficient. In both cases, the computational effort for the gradient scales as $O(N^2)$ with the number of basis functions (N), compared with a formal scaling

---

Correspondence to: W. Thiel

of $O(N^3)$ for the RHF treatment. Analogous remarks apply to the unrestricted Hartree-Fock (UHF) method which has therefore been quite popular in semiempirical studies of open-shell systems. However, despite its variational nature and its efficiency, the UHF approach cannot be considered entirely satisfactory because the UHF wavefunction is generally not an eigenfunction of the S^2 operator. Moreover, in certain cases, semiempirical UHF calculations may exaggerate correlation effects and thus yield artificially low energies for open-shell species [12] which causes an imbalanced description of closed-shell and open-shell regions on a given potential surface.

The exact RHF treatment of open-shell systems [13] is more complicated than the usual closed-shell RHF procedure. The approximate half-electron RHF method [14], on the other hand, is conceptually simple [15] and easily implemented, but analytical gradient evaluation is more difficult than in the case of the exact variational RHF treatment [13]. The first derivatives of the half-electron energy are closely related to the energy gradient in configuration interaction (CI) and multiconfigurational self-consistent-field (MCSCF) approaches which have received considerable attention at the *ab initio* level (see Refs. [16–20] for reviews). Analytical first derivatives of the half-electron energy have been derived in a recent semiempirical study [21] as the limiting case of small CI expansions. The proposed algorithm [21] exhibits an $O(N^4)$ execution time, however, so that the gradient evaluation remains the time-determining step for large molecules, in comparison with the energy evaluation that scales as $O(N^3)$.

The present paper introduces a more efficient formulation of the analytical energy gradient in semiempirical half-electron RHF calculations. Section 2 contains the theoretical derivations, section 3 describes the implementation of the new formalism for MNDO-type methods (including MNDO/d), and section 4 reports on the performance obtained with this implementation.

2 Theory

In the exact open-shell RHF treatment, the electronic energy is given by [22]:

$$E_{\mathrm{ROHF}} = \sum_i n_i H_{ii} + \sum_{ij} \left(a_{ij} J_{ij} + b_{ij} K_{ij} \right) \ , \tag{1}$$

where n_i is the average occupation number of the i-th molecular orbital (MO), a_{ij} and b_{ij} are constants characteristic of the structure of the open shell, and H_{ii}, J_{ij} and K_{ij} are one-electron, Coulomb and exchange integrals over the corresponding MOs, respectively. The open-shell MOs are determined by a variational minimization of E_{ROHF}. In the half-electron approach, this is replaced by a variational minimization of a simpler quasi closed-shell expression:

$$E_{\mathrm{R}} = \sum_i n_i H_{ii} + \frac{1}{4} \sum_{ij} n_i n_j \left(2J_{ij} - K_{ij} \right) \ . \tag{2}$$

The electronic energy $E^{\mathrm{HE}}_{\mathrm{ROHF}}$ is then obtained by adding a "half-electron correction" E^{HE} which corresponds to the difference $E_{\mathrm{ROHF}} - E_{\mathrm{R}}$ calculated with the use of the half-electron MOs.

$$E^{\mathrm{HE}}_{\mathrm{ROHF}} = E_{\mathrm{R}} + E^{\mathrm{HE}} \, . \tag{3}$$

This electronic energy is non-variational because the exact energy expression, Eq. (1), is evaluated from MOs that minimize Eq. (2). The errors caused by this approximation are known to be small in a semiempirical framework [23]. The variational energy contribution E_{R} in Eq. (3) has the same form as the closed-shell energy expression. Its geometrical derivatives are well-known [11,24,25] and reduce to a simple sum of atom pair contributions in the MNDO case. Therefore, only the derivatives of the non-variational correction E^{HE} need to be considered here.

Expressions for E^{HE} have been reported for several important cases [14, 26–28]. The correction can always be written as a linear combination of the Coulomb and exchange integrals over the orbitals in the open shells:

$$E^{\mathrm{HE}} = \sum_{r\le s}^{\text{open}} h_{rs} J_{rs} + \sum_{r<s}^{\text{open}} g_{rs} K_{rs} \, . \tag{4}$$

The coefficients h_{rs} and g_{rs} for the simplest cases are given in Table 1.

Table 1. Weights of the Coulomb and exchange integrals in the half-electron correction E^{HE} [14, 26, 27]

Configuration	Multiplicity	h_{rr}	h_{rs}	g_{rs}
$\phi_r\phi_s$	3	−1/4	0	−1/2
$\phi_r \pm \phi_s$	2	−1/16	−1/4	1/8
$\phi_r^2\phi_s \pm \phi_r\phi_s^2$	2	−1/16	−1/4	1/8
ϕ_r	2	−1/4		
$\phi_r\phi_s$	1	−1/4	0	3/2
$\phi_r^2 - \phi_s^2$	1	1/4	−1	−1/2
$\phi_r^2 + \phi_s^2$	1	1/4	−1	3/2

The formulation for the general high-spin case [28] considers an arbitrary combination of up to triply degenerate open subshells where each open-shell MO is characterized by the degeneracy $m_r \le 3$ and the number of α and β electrons (n_r^α, n_r^β, $n_r = n_r^\alpha + n_r^\beta$) in the corresponding open subshell. The coefficients in Eq. (4) are then defined [28] by:

$$h_{rr} = g_{rr} = -\frac{n_r^2}{4m_r^2} + \frac{n_r^\beta}{m_r} \, , \tag{5}$$

$$h_{rs} = -\frac{1}{m_r(m_r-1)}\left(\frac{n_r(2m_r-n_r)}{m_r} - n_r^\alpha + n_r^\beta\right) , \tag{6}$$

$$g_{rs} = -\frac{1}{2m_r(m_r-1)}\left(\left(n_r^\alpha - n_r^\beta\right)^2 - \frac{n_r(2m_r-n_r)}{m_r}\right) , \tag{7}$$

if the orbitals r and s belong to the same degenerate open subshell ($m_r = 2, 3$ for Eqs. (6)-(7)), and by

$$h_{rs} = 0 \; , \tag{8}$$

$$g_{rs} = -\frac{1}{2m_r m_s}\left(n_r^\alpha - n_r^\beta\right)\left(n_s^\alpha - n_s^\beta\right) \; , \tag{9}$$

if they belong to different open subshells. Although only the upper triangles of h_{rs} and g_{rs} enter Eq. (4), we define h_{rs} and g_{rs} as symmetric matrices and choose $g_{rr} = h_{rr}$ which allows an alternative (more symmetric) formulation of the half-electron correction:

$$E^{\mathrm{HE}} = \frac{1}{2}\sum_{rs}^{\text{open}}\left(h_{rs}J_{rs} + g_{rs}K_{rs}\right) \; . \tag{10}$$

Static and response parts of the derivative of the half-electron correction

Formal differentiation of Eq. (4) or Eq. (10) with respect to the nuclear coordinate τ_a of nucleus A involves the corresponding derivatives of the Coulomb integrals J_{ij} and exchange integrals K_{ij} in the MO basis [20]:

$$\frac{dJ_{ij}}{d\tau^a} = J_{ij}^a + 2\left(\sum_p x_{pi}^a\left(jj|ip\right) + \sum_p x_{pj}^a\left(ii|jp\right)\right) \; , \tag{11}$$

$$\frac{dK_{ij}}{d\tau^a} = K_{ij}^a + 2\left(\sum_p x_{pi}^a\left(ij|jp\right) + \sum_p x_{pj}^a\left(ij|ip\right)\right) \; , \tag{12}$$

where x_{pi}^a denotes the derivative of an orbital coefficient in the basis of the unperturbed MOs, and $\left(ij|kl\right)$ a two-electron MO integral. In MNDO approximation, the static parts of the derivatives, J_{ij}^a and K_{ij}^a, are given by

$$J_{ij}^a = \sum_{B\neq A}\sum_{\mu\leq\nu}^{A}\sum_{\lambda\leq\sigma}^{B}\left(O_{\mu\nu}^{ii}O_{\lambda\sigma}^{jj} + O_{\lambda\sigma}^{ii}O_{\mu\nu}^{jj}\right)(\mu\nu,\lambda\sigma)^a \; , \tag{13}$$

$$K_{ij}^a = \sum_{B\neq A}\sum_{\mu\leq\nu}^{A}\sum_{\lambda\leq\sigma}^{B} 2O_{\mu\nu}^{ij}O_{\lambda\sigma}^{ij}\,(\mu\nu,\lambda\sigma)^a \; , \tag{14}$$

$$O_{\mu\nu}^{ij} = C_{\mu i}C_{\nu j} + \left(1 - \delta_{\mu\nu}\right)C_{\nu i}C_{\mu j} \; , \tag{15}$$

where $C_{\mu i}$ and $(\mu\nu,\lambda\sigma)^a$ denote the usual LCAO coefficients and the two-electron AO integral derivatives, respectively. The static part of the derivative of the half-electron correction

$$\left.\frac{\partial E^{\mathrm{HE}}}{\partial\tau^a}\right|_{\text{static}} = \sum_{r\leq s}^{\text{open}} h_{rs}J_{rs}^a + \sum_{r<s}^{\text{open}} g_{rs}K_{rs}^a \tag{16}$$

can be evaluated easily, with $O(N^2)$ operations and without any storage of AO derivative integrals.

The complete derivative expression contains a static and a response part:

$$\frac{dE^{\mathrm{HE}}}{d\tau^a} = \left.\frac{\partial E^{\mathrm{HE}}}{\partial \tau^a}\right|_{\mathrm{static}} + \left.\frac{\partial E^{\mathrm{HE}}}{\partial \tau^a}\right|_{\mathrm{response}} . \tag{17}$$

The response contribution arises from the second term in Eqs. (11) and (12), respectively. Eqs. (10)-(12) and (17) yield:

$$\left.\frac{\partial E^{\mathrm{HE}}}{\partial \tau^a}\right|_{\mathrm{response}} = \sum_i^{\mathrm{all}} \sum_r^{\mathrm{open}} q_{ir} x_{ir}^a \tag{18}$$

$$q_{ir} = 2 \sum_s^{\mathrm{open}} \left(h_{rs} \left(ss|ri\right) + g_{rs} \left(rs|si\right)\right) . \tag{19}$$

The evaluation of the two-electron quantities q_{ir} requires all two-electron MO integrals $(ij|kl)$ which involve three or four open-shell MOs. There are $O(N)$ such integrals, each of which can be computed with $O(N^2)$ operations in MNDO approximation, so that a straightforward implementation of Eq. (19) results in an $O(N^3)$ procedure.

The first-order orbital coefficients x_{ir}^a in Eq. (18) are the solutions of the coupled perturbed Hartree-Fock (CPHF) equations, which are discussed in the following section.

CPHF equations with fractional occupation numbers in MNDO

Numerous derivations of the first-order CPHF equations are available in the literature, e.g. [24, 29–31], so that we focus on the major steps which differ in the case of fractional occupation numbers. Although a slightly more compact derivation of the CPHF equations is possible with the use of the Derivative Hartree-Fock theory [31], the classical approach allows for an easier exposure of the computational complexity.

The classical derivation proceeds as follows:

- The Fock matrix $\mathscr{F}_{ij}$, the orbital coefficient matrix $\mathscr{C}_{ij}$ and the orbital energies ε_i are expanded in the basis of the unperturbed molecular orbitals as a function of a small perturbation τ^a.
- These expansions are substituted into the Hartree-Fock equations and the orthonormality conditions on the molecular orbitals.
- The terms that are of first order in τ^a are collected.

The resulting equations for the first-order orbital coefficients x_{ij}^a and energies ε_i^a in the MNDO approximation are:

$$\left(\varepsilon_j - \varepsilon_i\right) x_{ij}^a = \frac{d\mathscr{F}_{ij}}{d\tau^a} , \tag{20}$$

$$x_{ij}^a + x_{ji}^a = 0 , \tag{21}$$

$$\varepsilon_i^a = \frac{d\mathscr{F}_{ii}}{d\tau^a} \, . \tag{22}$$

The first-order Fock matrix $\frac{d\mathscr{F}_{ij}}{d\tau^a}$ depends on the first-order density matrix which is linearly related to x_{ij}^a:

$$\frac{dP_{\mu\nu}}{d\tau^a} = -\sum_{i<j} \left(C_{\mu i} C_{\nu j} + C_{\mu j} C_{\nu i}\right) \left(n_i - n_j\right) x_{ij}^a \, . \tag{23}$$

The derivative of the MNDO expression for the Fock matrix [1,7] can be partitioned into a static part ($\mathscr{F}_{ij}{}^a$) and a response part ($\mathscr{R}_{ij}^a$):

$$\frac{d\mathscr{F}_{ij}}{d\tau^a} = \mathscr{F}_{ij}{}^a + \mathscr{R}_{ij}^a \, , \tag{24}$$

$$\mathscr{F}_{ij}{}^a = \sum_{\mu\nu} C_{\mu i} C_{\nu j} F_{\mu\nu}^a \, , \tag{25}$$

$$F_{\mu\nu}^a = \sum_{B \neq A} \left(\sum_{\lambda\sigma \in B} P_{\lambda\sigma} (\mu\nu, \lambda\sigma)^a - Z_B (\mu\nu, q_{\text{core}}^B)^a \right) \text{ if } \mu, \nu \in A, \tag{26}$$

$$F_{\mu\nu}^a = \sum_{\lambda\sigma \in A} P_{\lambda\sigma} (\mu\nu, \lambda\sigma)^a - Z_A (\mu\nu, q_{\text{core}}^A)^a \text{ if } \mu, \nu \in B \neq A, \tag{27}$$

$$F_{\mu\lambda}^a = \beta_{\mu\lambda} S_{\mu\lambda}^a - \frac{1}{2} \sum_{\nu \in A} \sum_{\sigma \in B} P_{\nu\sigma} (\mu\nu, \lambda\sigma)^a \text{ if } \mu \in A, \, \lambda \in B \neq A, \tag{28}$$

$$\mathscr{R}_{ij}^a = \sum_{\mu\nu} C_{\mu i} C_{\nu j} R_{\mu\nu}^a \, , \tag{29}$$

$$R_{\mu\nu}^a = -\frac{1}{2} \sum_{\lambda\sigma \in B} \frac{dP_{\lambda\sigma}}{d\tau^a} (\nu\lambda, \nu\sigma) + \sum_B \sum_{\lambda\sigma \in B} \frac{dP_{\lambda\sigma}}{d\tau^a} (\mu\nu, \lambda\sigma) \text{ if } \mu, \nu \in A, \tag{30}$$

$$R_{\mu\lambda}^a = -\frac{1}{2} \sum_{\nu \in A} \sum_{\sigma \in B} \frac{dP_{\nu\sigma}}{d\tau^a} (\mu\nu, \lambda\sigma) \text{ if } \mu \in A, \, \lambda \in B \neq A, \tag{31}$$

where the standard MNDO notation [1,7] has been employed for the terms in Eqs. (24)–(31). $(\mu\nu, \lambda\sigma)^a$ and $S_{\mu\nu}^a$ denote the derivatives of the two-electron integrals and the overlap integrals in the AO basis, respectively, with regard to τ_a.

Substitution of Eqs. (23)-(31) into Eqs. (20) and (22) yields two systems of linear equations for x_{ij}^a and ε_i^a (CPHF equations). The non-redundant CPHF variables are those that appear in the derivatives of the density matrix, Eq. (23), i.e. x_{ij}^a with $i < j$ and $n_i \neq n_j$. Other CPHF variables are redundant in the sense that they are available from Eqs. (20) and (22) as soon as the non-redundant CPHF variables are known.

A compact representation of the CPHF equations in MNDO approximation can be obtained in matrix notation. Since $x_{ii}^a = 0$ in MNDO due to Eq. (21) we may formally introduce the vector $\mathbf{X^a}$ of unknowns with the components:

$$\mathbf{X^a}_{ii} = \varepsilon_i^a \,, \tag{32}$$

$$\mathbf{X^a}_{ij} = x_{ij}^a \text{ for } i \neq j. \tag{33}$$

The CPHF equations can then be written as a single matrix equation

$$(\Gamma - \mathbf{K})\,\mathbf{N}\mathbf{X^a} = \mathcal{F}^{\mathbf{a}} \,, \tag{34}$$

where Γ and $\mathbf{N}$ are diagonal matrices (see below), $\mathbf{K}$ is a square matrix with two-electron MO integrals (see below), and $\mathcal{F}^{\mathbf{a}}$ is a vector containing the static Fock matrix derivatives (see Eq. (25)). All vectors and matrices in Eq. (34) have compound indices (ij and kl), and their dimension is equal to the number of unknowns. They can be partitioned into redundant and non-redundant parts:

$$\Gamma = \begin{pmatrix} \Gamma_{\mathrm{NR}} & 0 \\ 0 & \Gamma_{\mathrm{R}} \end{pmatrix} , \tag{35}$$

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_{\mathrm{NR}} & 0 \\ 0 & \mathbf{N}_{\mathrm{R}} \end{pmatrix} , \tag{36}$$

$$\mathbf{X^a} = \begin{pmatrix} \mathbf{X}^{\mathbf{a}}_{\mathrm{NR}} \\ \mathbf{X}^{\mathbf{a}}_{\mathrm{R}} \end{pmatrix} , \tag{37}$$

$$\mathcal{F}^{\mathbf{a}} = \begin{pmatrix} \mathcal{F}^{\mathbf{a}}_{\mathrm{NR}} \\ \mathcal{F}^{\mathbf{a}}_{\mathrm{R}} \end{pmatrix} , \tag{38}$$

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{\mathrm{NR}} & 0 \\ \mathbf{K}_{\mathrm{R}} & 0 \end{pmatrix} . \tag{39}$$

The definition of Γ, $\mathbf{N}$ and $\mathbf{K}$ explicitly accounts for the occupation numbers (n_i, n_j, etc.) and, in particular, allows for fractional occupations as they occur in the half-electron treatment:

$$(\Gamma_{\mathrm{NR}})_{ij,kl} = \delta_{ik}\delta_{jl}\frac{\varepsilon_j - \varepsilon_i}{n_j - n_i} \,, \tag{40}$$

$$(\Gamma_{\mathrm{R}}\ )_{ij,kl} = \delta_{ik}\delta_{jl}\left(\varepsilon_j - \varepsilon_i\right) \ , \text{ if } i \neq j \tag{41}$$

$$(\Gamma_{\mathrm{R}}\ )_{ii,ii} = 1 \,, \tag{42}$$

$$(\mathbf{N}_{\mathrm{NR}})_{ij,kl} = \delta_{ik}\delta_{jl}\left(n_j - n_i\right) \,, \tag{43}$$

$$(\mathbf{N}_{\mathrm{R}}\ )_{ij,kl} = \delta_{ik}\delta_{jl} \,, \tag{44}$$

$$(\mathbf{K}_{\mathrm{NR}})_{ij,kl} = \left(1 - \delta_{n_i,n_j}\right)\left(1 - \delta_{n_k,n_l}\right)\left(4\left(ij|kl\right) - \left(ik|jl\right) - \left(il|jk\right)\right) , \tag{45}$$

$$(\mathbf{K}_{\mathrm{R}}\ )_{ij,kl} = \delta_{n_i,n_j}\left(1 - \delta_{n_k,n_l}\right)\left(4\left(ij|kl\right) - \left(ik|jl\right) - \left(il|jk\right)\right) , \tag{46}$$

Note that the definition chosen for $\mathbf{N}$ ensures that $\mathbf{K}_{\mathrm{NR}}$ and $(\Gamma - \mathbf{K}_{\mathrm{NR}})$ are symmetric matrices which is relevant for the actual solution of the linear system (see below). By contrast, $\mathbf{K}$ is not symmetric because all columns corresponding to redundant variables $\mathbf{X^a}_{ij}$ are zero. The formal solution of Eq. (34) can therefore be written as:

$$\mathbf{X}^{\mathbf{a}}_{\mathrm{NR}} = \mathbf{N}_{\mathrm{NR}}{}^{-1} \left(\Gamma_{\mathrm{NR}} - \mathbf{K}_{\mathrm{NR}}\right)^{-1} \mathcal{F}^{\mathbf{a}}_{\mathrm{NR}} \tag{47}$$

$$\mathbf{X}^{\mathbf{a}}_{\mathrm{R}} = \Gamma_{\mathrm{R}}{}^{-1} \left(\mathcal{F}_{\mathrm{R}}^{\mathbf{a}} + \mathbf{K}_{\mathrm{R}}\mathbf{N}_{\mathrm{NR}}\mathbf{X}^{\mathbf{a}}_{\mathrm{NR}}\right) \ . \tag{48}$$

Eq. (48) is valid for any subset of redundant variables as long as all non-redundant variables are kept. In the following section, we shall consider only the subset of redundant variables which belong to the active set (see below).

Z-vector evaluation of the response contributions

Using the symmetry relation $x^a_{ij} = -x^a_{ji}$ (see Eq. (21)) which holds in MNDO approximation, the response contribution to the derivative of the half-electron correction (see Eqs. (18)-(19)) can be rewritten as follows:

$$\left.\frac{\partial E^{\mathrm{HE}}}{\partial \tau^a}\right|_{\mathrm{response}} = \sum_{i \le j}^{\text{active set}} \mathbf{Q}_{ij}\mathbf{X}^{\mathbf{a}}{}_{ij} = \mathbf{Q}^{\dagger}\mathbf{X}^{\mathbf{a}} \ , \tag{49}$$

$$\mathbf{Q}_{ir} = q_{ir} \qquad \text{if } i \in \text{closed}, \ r \in \text{open}, \tag{50}$$

$$\mathbf{Q}_{ri} = -q_{ir} \qquad \text{if } r \in \text{open}, \ j \in \text{vacant}, \tag{51}$$

$$\mathbf{Q}_{rs} = q_{rs} - q_{sr} \ \text{if } r, s \in \text{open}. \tag{52}$$

The active set includes all non-redundant CPHF variables and those redundant variables $\mathbf{X}^{\mathbf{a}}{}_{rs}$ where both r and s refer to an open-shell MO ($r < s$). The other redundant variables do not contribute to the derivative (see Eq. (18)) and are therefore excluded. The number of redundant variables in the active set is usually rather small, e.g. zero for a nondegenerate doublet and one for the other cases in Table 1.

The two-electron quantities $\mathbf{Q}_{ij}$ in Eq. (49) are nonzero only if at least one of the indices corresponds to an open-shell MO. Since $\mathbf{Q}$ is shared between all Cartesian derivatives and since different vectors $\mathbf{X}^{\mathbf{a}}$ enter Eq. (49) linearly, the Z-vector technique [32] may be used to improve the computational efficiency. However, the direct application of the Z-vector method to Eq. (49) would require the solution of the unsymmetrical linear problem $(\Gamma - \mathbf{K})^{\dagger}\,\mathbf{Z} = \mathbf{N}^{-1}\mathbf{Q}$. As unsymmetrical systems of linear equations are generally handled less efficiently than symmetrical ones by the existing iterative algorithms [33], a symmetric formulation is desirable.

Substitution of $\mathbf{X}^{\mathbf{a}}$ from Eqs. (47-48) into Eq. (49) yields, due to the block structure of $\mathbf{K}$:

$$\begin{aligned}\mathbf{Q}^{\dagger}\mathbf{X}^{\mathbf{a}} = {} & \mathbf{Q}^{\dagger}_{\mathrm{R}}\Gamma_{\mathrm{R}}{}^{-1}\mathcal{F}_{\mathrm{R}}^{\mathbf{a}} + \\ & \left(\mathbf{Q}^{\dagger}_{\mathrm{NR}}\mathbf{N}_{\mathrm{NR}}{}^{-1} + \mathbf{Q}_{\mathrm{R}}{}^{\dagger}\Gamma_{\mathrm{R}}{}^{-1}\mathbf{K}_{\mathrm{R}}\right) \left(\Gamma_{\mathrm{NR}} - \mathbf{K}_{\mathrm{NR}}\right)^{-1} \mathcal{F}^{\mathbf{a}}_{\mathrm{NR}} \ .\end{aligned} \tag{53}$$

Introducing the auxiliary symmetrical linear problem:

$$\left(\Gamma_{\mathrm{NR}} - \mathbf{K}_{\mathrm{NR}}\right) \mathcal{Y} = \mathbf{N}_{\mathrm{NR}}{}^{-1}\mathbf{Q}_{\mathrm{NR}} + \mathbf{K}_{\mathrm{R}}{}^{\dagger}\Gamma_{\mathrm{R}}{}^{-1}\mathbf{Q}_{\mathrm{R}} \ , \tag{54}$$

we obtain:

$$\left.\frac{\partial E^{\mathrm{HE}}}{\partial \tau^a}\right|_{\mathrm{response}} = \begin{pmatrix} \mathscr{Y} \\ \Gamma_{\mathrm{R}}{}^{-1}\mathbf{Q}_{\mathrm{R}} \end{pmatrix}^{\dagger} \mathscr{F}^{\mathrm{a}} = \mathscr{Z}^{\dagger}\mathscr{F}^{\mathrm{a}} . \tag{55}$$

Since Eq. (54) is of the same form as the closed-shell CPHF equations, the algorithms employed in the solution of the latter (e.g. [30, 34]) can be directly applied to Eq. (54). Finally, substituting the expression for the static derivative $\mathscr{F}^{\mathrm{a}}$ of the Fock matrix (Eq. (25)) into Eq. (55), and interchanging the summation indices, an alternative expression in the AO basis is obtained:

$$\left.\frac{\partial E^{\mathrm{HE}}}{\partial \tau^a}\right|_{\mathrm{response}} = \sum_{\mu,\nu} F^a_{\mu\nu} Z_{\mu\nu} , \tag{56}$$

$$Z_{\mu\nu} = \sum_{i \le j}^{\text{active set}} C_{\mu i} C_{\nu j} \mathscr{Z}_{ij} . \tag{57}$$

Since the transformation for the static derivative $\mathscr{F}^{\mathrm{a}}$ of the Fock matrix into the MO basis (Eq. (25)) is $O(N^3)$ in MNDO, while the construction of its AO representation (Eqs. (26–28)) is only $O(N^2)$, the calculation in the AO basis is significantly more efficient.

3 Implementation

The analytical derivatives of the energy in the half-electron open-shell treatment have been implemented in the MNDO94 [35] program. This section describes the major steps performed during gradient computation. Actual execution times for each step are given in parentheses for illustration (computation of the $3N$ Cartesian derivatives in $C_{150}H_{30}$ (see below) on a 100MHz SGI Indigo^2 with 64MB of main memory):

1 (30.83s) The required two-center integrals are computed. The static parts of the derivatives of the Coulomb and exchange integrals (Eqs. (13-15)), the derivatives of the closed-shell energy term E_{R} (Eq. (3)) and the pair contributions to the static derivatives of the Fock matrix (Eqs. (26-28)) are determined using the derivative two-center AO integrals, which are discarded thereafter.

2 (20.06s) The two-electron quantity $\mathbf{Q}$ is evaluated (Eqs. (19, 50–52)) and the right-hand side of the auxiliary linear system (54) is constructed. Depending on the number of the non-zero redundant components of $\mathbf{Q}$, the vector $\mathrm{K_R}^{\dagger}\Gamma_{\mathrm{R}}{}^{-1}\mathbf{Q}_{\mathrm{R}}$ is evaluated by using transformed MO integrals (Eq. (46)), or by performing a direct transformation of the two-electron terms employing Eqs. (29)-(31), with $\frac{dP_{\lambda\sigma}}{d\tau^a}$ replaced by the auxiliary quantity:

$$t_{\lambda\sigma} = - \sum_{i<j}^{\text{redundant}} \left(C_{\lambda i} C_{\sigma j} + C_{\lambda j} C_{\sigma i}\right) \mathbf{Q}_{ij} . \tag{58}$$

3 (652.62s) The auxiliary linear system (54) is solved using a modified conjugate gradient algorithm [33, 34]. The conjugate gradient algorithm solves a linear system $\mathbf{Ax} = \mathbf{b}$ by generating a series of the successive approximations $\mathbf{x}^{(i)}$ to the solution, corresponding residuals $\mathbf{r}^{(i)} = \mathbf{b} - \mathbf{Ax}^{(i)}$ and search directions $\mathbf{d}^{(i)}$, which are updated on each iteration according to [33]:

$$\mathbf{x}^{(i)} = \mathbf{x}^{(i-1)} + \alpha_i \mathbf{d}^{(i)} , \tag{59}$$

$$\mathbf{p}^{(i)} = \mathbf{Ad}^{(i)} , \tag{60}$$

$$\mathbf{r}^{(i)} = \mathbf{r}^{(i-1)} - \alpha_i \mathbf{p}^{(i)} , \tag{61}$$

$$\mathbf{d}^{(i)} = \mathbf{r}^{(i)} + \beta_{i-1} \mathbf{d}^{(i-1)} . \tag{62}$$

The scalars α_i and β_{i-1} are selected such that each two subsequent search directions and residual vectors are orthogonal. For infinitely precise arithmetics, it can then be proved that all search directions (and residuals) are mutually orthogonal [33], provided that $\mathbf{A}$ is symmetric positive definite. In practice, for the linear problems typically encountered in CPHF (e.g., with the dimension of 99 852 in the case of $C_{150}H_{30}$), the global orthogonality property tends to be lost after more than ten iterations. However, this does not result in a significant increase in the number of iterations compared with other algorithms [30, 36] which have significantly larger memory requirements. The evaluation of the response vectors (Eq. (60)) using the transformation to atomic orbitals [36] is the most time-consuming part of the above procedure (with an execution time of 605.21 sec. in the case of $C_{150}H_{30}$, where 16 response vectors are required to converge to 10^{-5}a.u.).

4 (33.98s) The AO representation of the $\mathbf{Z}$ vector is computed according to Eqs. (55) and (57).

5 (56.81s) The static derivatives of the Fock matrix in the AO basis are constructed from the stored pair contributions (Eqs. (26-28)) and contracted with the $\mathbf{Z}$-vector to give the response part of the derivatives of the half-electron correction (Eq. (56)).

Throughout the code, care is taken to use the BLAS Level 3 routines [37] for the major part of the computations which should ensure a near-optimal performance of the program on most modern computer architectures.

4 Results and discussion

To assess the performance of the present implementation relative to the numerical differentiation and the previously reported analytical implementation [21], Cartesian energy gradients were computed for a series of condensed aromatic hydrocarbons [38] in the lowest triplet state, i.e. for C_6H_6, $C_{24}H_{12}$, $C_{54}H_{18}$, $C_{96}H_{24}$ and $C_{150}H_{30}$. The convergence criterion for the SCF energy was 10^{-6} e.V., while the auxiliary linear system (54) was required to converge to within 10^{-5} a.u. Enforcing more stringent convergence criteria results in insignificant changes in the

Table 2. Comparison of the SCF and gradient computation times[a] (in seconds)

System	Orbitals	SCF	Numerical gradient	Analytical gradient, [21][b]	Analytical gradient, present work
C_6H_6 [c]	30	0.19	7.71	1.74	0.20
$C_{24}H_{12}$ [c]	108	7.31	714.45	178.95	4.55
$C_{54}H_{18}$ [c]	234	82.18	10 484.92	4 116.91	36.49
$C_{96}H_{24}$ [d]	408	605.41	[e]	100 139.32	221.99
$C_{150}H_{30}$ [d]	630	3 325.14	[f]	[f]	794.44

[a] On an SGI Indigo^2 with a 100MHz R4000 CPU and 64MB of main memory
[b] As implemented in the MOPAC7 [39] program
[c] At the optimized MNDO geometry
[d] At an idealized D_{6h} geometry with R_{CC} = 1.4066Å and R_{CH} = 1.0904Å
[e] No SCF convergence in the gradient computation
[f] Not computed, estimated execution time too large

computed gradients. The gradient components from all three approaches deviate by much less than 1 kcal/Angstrom, and thus need not be shown here.

The execution times in Table 2 show that, in the present implementation, the gradient computation requires only a fraction of the effort necessary to obtain the SCF solution for all test molecules (except C_6H_6). Employing the conjugate gradient algorithm, the memory usage does not depend on the number of iterations in the solution of the auxiliary linear system (54). Overall, the memory requirements in the gradient evaluation are comparable to those in the SCF treatment. The combination of high computational efficiency and modest memory requirements makes it possible to routinely perform geometry optimizations for large open-shell systems where the UHF treatment has previously been the only practical alternative. Thus, the full geometry optimization (in Cartesian coordinates) of the lowest triplet state of C_{60} starting from the singlet I_h geometry takes 990 seconds on a 100MHz Indigo^2 using the half-electron approach, while 1201 seconds are necessary with the use of the UHF procedure.

The previously reported implementation [21] of analytical gradients employs an $O(N^4)$ algorithm, as opposed to the $O(N^3)$ algorithm in the present work. This difference causes much lower computation times for the present implementation, particularly for the larger molecules (see Table 2), e.g. for $C_{96}H_{24}$ with a speedup factor of about 450. In addition, the relative efficiency of the previous analytical implementation [21] with respect to numerical differentiation was found to decrease with the system size, and special care had to be taken to control the growth of the memory requirements with the number of iterations, thereby effectively limiting applications in large molecules [21].

Although the derivative of the half-electron correction (4) was the major objective of the present work, the treatment can easily be extended to handle small CI expansions typically encountered in semiempirical computations. Since the Hellmann-Feynman theorem is rigorously satisfied for the gradient of a limited CI energy expression in an incomplete basis, the gradient can be written as a linear combination of the derivatives of the two-electron MO integrals and the first-order orbital energies [40]. The static part of the derivative can then be computed using Eqs. (13)-(15). The response part reduces to an expression in the form of Eq. (49) (although with a different definition for the vector **Q**), provided that the derivatives of the orbital energies are included in the redundant set of parameters (Eq. (32)). Hence, the expressions (53)-(57) for the response part of the derivative still hold and can be evaluated using the same computer program. Therefore, computation of the gradient of the energy in a small MNDO CI expansion can be implemented as an $O(N^3)$ process, completely analogous to the gradient of the half-electron term. Work along these lines is in progress.

5 Conclusion

The Z-vector formulation of the analytical derivatives of the energy in the semiempirical half-electron open-shell treatment provides a substantial increase in the computational efficiency over both the numerical and the conventional analytical implementation. The computation time formally scales as $O(N^3)$ with the size of the system, and is typically a small fraction of the SCF computation time, with memory requirements growing as $O(N^2)$. The methodology can easily be extended to cover small CI expansions.

Acknowledgements. This work was supported by the Schweizerischer Nationalfonds.

References

1. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4899, 4907
2. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107:3902
3. Stewart JJP (1989) J Comp Chem 10:209, 221
4. Dewar MJS (1985) J Phys Chem 89:2145
5. Thiel W (1988) Tetrahedron 44:7393
6. Stewart JJP (1990) In: Lipkowitz JB, Boyd DB (eds) Reviews in Computational Chemistry, vol 1 VCH Publishers, New York, pp 45-81
7. Thiel W, Voityuk AA (1992) Theor Chim Acta 81:391
8. Thiel W, Voityuk AA (1992) Int J Quantum Chem 44:807
9. Thiel W, Voityuk AA (1994) J Mol Struct (Theochem) 313:141
10. Thiel W, Voityuk AA J Phys Chem (in press)
11. Dewar MJS, Yamaguchi Y (1978) Comput Chem 2:25
12. Dewar MJS, Olivella S, Stewart JJP (1986) J Am Chem Soc 108:5771
13. Roothaan CCJ (1960) Rev Mod Phys 32:179
14. Dewar MJS, Hashmall JA, Venier CG (1968) J Am Chem Soc 90:1953
15. Longuet-Higgins HC, Pople JA (1955) Proc Phys Soc A68:591
16. Schaefer III HF, Yamaguchi Y (1986) J Mol Struct (Theochem) 135:369
17. Amos RD (1987) Adv Chem Phys 67:99
18. Pulay P (1987) Adv Chem Phys 69:241
19. Jørgensen P, Simons J (eds) (1985) Geometrical derivatives of energy surfaces and molecular properties, D Reidel Publishing Company, Dordrecht
20. Yamaguchi Y, Osamura Y, Goddard JD, Schaefer III HF (1994) A new dimension to quantum chemistry: Analytic derivative methods in *ab initio* molecular electronic structure theory, Oxford University Press, New York

21. Dewar MJS, Liotard DA (1990) J Mol Struct (Theochem) 206:123
22. Bobrowicz FW, Goddard WA (1977) In: Schaefer HF (ed) Modern theoretical chemistry, vol 3, Plenum, New York, pp 79-129
23. Dewar MJS, Olivella S (1979) J Chem Soc Farad Trans II 75:829
24. Gerratt J, Mills IM (1968) J Chem Phys 49:1719
25. Pulay P (1969) Mol Phys 17:197
26. Dewar MJS, Trinajstic N (1970) Chem Commun 646
27. Ellison FO, Matheu FM (1971) Chem Phys Lett 10:322
28. Voityuk AA (1983) Zh Strukt Khim (Russ) 24(3):18
29. Dodds JL, McWeeny R, Sadley AJ (1977) Mol Phys 34:1779
30. Pople JA, Krishnan R, Schlegel HB, Binkley JS (1979) Int J Quantum Chem S13:225
31. Dykstra CE, Jasien PG (1984) Chem Phys Lett 109:388
32. Handy NC, Schaefer III HF (1984) J Chem Phys 81:5031
33. Barrett R, Berry M, Chan TF, Demmel J, Donato J, Dongarra J, Eijkhout V, Pozo R, Romine C, Van der Vorst H (1994) Templates for the solution of linear systems: building blocks for iterative methods, CIAM, Philadelphia
34. Wormer PES, Visser F (1982) J Comput Phys 48:23
35. Thiel W (1994) Program MNDO94, version 4.0
36. Osamura Y, Yamaguchi Y, Saxe P, Fox DJ, Vincent MA, Schaefer III HF (1983) J Mol Struct (Theochem) 103:183
37. Dongarra J, Du Croz J, Duff I, Hammarling S (1988) A set of level 3 basic linear algebra subprograms, Argonne National Laboratory preprint No. 1
38. Bakowies D, Thiel W (1991) J Am Chem Soc 113:3704
39. Stewart JJP (1993) MOPAC7: A Public-Domain General Molecular Orbital Package (QCPE 455 7.0), see QCPE Bull 13:42
40. Tachibana A, Yamashita K, Yamabe T, Fukui K (1978) Chem Phys Lett 59:255

This article was processed by the author
using the Springer-Verlag PJour1g macro package 1991.