A quasi-Newton algorithm for first-order saddle-point location

P. Culot[1], G. Dive[1], V. H. Nguyen[2], and J. M. Ghuysen[1]

[1] Centre d'Ingénierie des Protéines, Université de Liège, Institut de Chimie, B6, B-4000 Sart Tilman (Liège 1), Belgium

[2] Department of Mathematics, Facultés Universitaires de Namur, Rempart de la Vierge, 8, B-5000 Namur, Belgium

Summary. A new algorithm for the location of a transition-state structure on an energy hypersurface is proposed. The method is compared to three other quasi-Newton step calculations available in literature. Numerical results derived from several examples are compared to those obtained by the two algorithms implemented in the Gaussian package.

1 Introduction

According to the idea of McIver and Komornicki [1], several attempts have been made to solve the problem of locating the transition-state structures. By reference to the work of Murrell et al. [2], the transition-state of a chemical reaction is assumed to be a stationary point on an energy hypersurface [3], associated to an indefinite Hessian matrix with one and only one negative eigenvalue. The first-order saddle-point can be interpreted as a maximum along the direction defined by the eigenvector corresponding to the negative eigenvalue, and a minimum in all the other orthogonal directions. The direction of the negative curvature is called the transition vector [4]. In most cases, all its components are not known a priori, and hence must be determined in the course of the iterative search procedure.

The purpose of this paper is to propose an algorithm to locate a first-order saddle-point and to compare it to the quasi-Newton algorithm of Schlegel [5], Baker [6], and Simons et al. [7].

Several applications of the new algorithm are described. The selected problems deal with simple isomerization reactions as well as with complex intramolecular rearrangements on a flat surface. The starting geometries are not necessarily taken within the quadratic region surrounding the transition structure. All the results are compared to those obtained by the two methods available in the Gaussian package [8].

2 Quasi-Newton algorithms

Let us consider the search of a first-order saddle-point of an energy function, $E(X)$. This function of N real variables is assumed to be at least twice continuously differentiable. The available methods in the literature are related to the quasi-Newton algorithm in which the user supplies an initial estimation X^1 of the first-order saddle-point.

As an approximation of the energy function at a point X^k, it is pertinent to use the energy E^k, the gradient vector G^k and an approximate Hessian matrix H^k. A Taylor series second-order expansion in the step direction D:

$$E(X^k + D) \cong Q^k(D) = E^k + G^{kT}D + \tfrac{1}{2}D^T H^k D \tag{1}$$

is thus considered. Given the stationary condition $\nabla Q^k(D) = 0$, the quasi-Newton step is equal to:

$$D^k = -H^{k-1}G^k \tag{2}$$

In terms of the eigenvectors V_i^k and the eigenvalues b_i^k of H^k, this step can be written as:

$$D^k = -\sum_{i=1}^{n} \frac{\bar{G}_i^k}{b_i^k} V_i^k \tag{3}$$

where $\bar{G}_i^k = V_i^{kT}G^k$ is the component of G^k along the eigenvector V_i^k.

For an *ab initio* SCF function, it is convenient to calculate the gradient each time the energy function is computed. However, the analytical or numerical calculation of the second derivative is expensive, about N times the computational effort of one gradient evaluation. Thus, an iterative updating of the Hessian should be performed using a suitable estimate of the second derivative matrix at the beginning of the search process.

The iterative procedure to locate a saddle-point involves the following steps:

(a) the choice of a search direction D^k,

(b) the scaling the step Euclidean norm $\|D^k\|$ below an upper bound $R_{\max}$,

(c) the calculation of $X^{k+1} = X^k + D^k$,

(d) the update of the approximate Hessian matrix H^{k+1} from H^k, D^k, G^k, and G^{k+1}.

To update this matrix, Baker [6] and Simons et al. [7] use the formula of Powell [9], while Schlegel [5] applies his own algorithm. The choice of the direction D^k in step (a) and the scaling procedure in step (b) are different in the three methods. Yet, all of them use convergence criteria based on the gradient vector and on the search direction length.

If the current point X^k lies in the region where the Hessian has the required unique negative eigenvalue, the quasi-Newton step is a good search direction. The first component of D^k, in the eigenvector V basis, is ascendant and all the other elements are descendant. In contrast, if the Hessian has not the expected local inertia, then the quasi-Newton step is not the proper one. So, the new point X^{k+1} has to leave this inadequate local area and to reach a region of the surface characterized by the right inertia property. All the methods discussed in this paper apply the quasi-Newton step in the vicinity of the solution. The way the algorithms evolve differs mainly from each other when the curvature at point X^k

is not the good one. This is one of the main problems involved in an optimization procedure.

2.1 Algorithm of Schlegel [5]

The right inertia of the approximate Hessian matrix is obtained by adjusting the sign of inadequate eigenvalues. If several negative eigenvalues occur, all of them are replaced by their absolute value, except the smallest one. If no negative eigenvalue is present, the sign of the least positive eigenvalue is changed. The quasi-Newton step is consequently modified as:

$$D^k = - \sum_{i=1}^{n} \frac{\bar{G}_i^k}{b_i'} V_i^k \tag{4}$$

where $|b_i'| = |b_i^k|$, $b_1' < 0 < b_2' < \cdots < b_n'$.

This can be related to the Greenstadt [10] proposal that is used in a minimization process. The main effect of this procedure is to reverse the ascendant/descendant character of the search direction. However, in an area characterized by a large curvature, the resulting direction is not necessarily the opposite of the initial one if the investigated region is far from an extremum and thus may be incorrect. This effect is modulated by a scaling factor, implemented in the current version of the Gaussian programs.

If the quasi-Newton search direction of Eq. (4) exceeds the maximum allowed step R_{max}, its length is set to R_{max}. This requires the addition of a shift parameter λ obtained by the search of an extremum of the quadratic function $Q^k(D)$ on the sphere with radius R_{max}, as proposed by Golab et al. [11]. In practice, the shift parameter λ is obtained by minimizing the function:

$$(\|D^k(\lambda)\| - R_{max})^2 \tag{5}$$

The radius R_{max} is updated using a trust region method, as suggested by Fletcher [12]. The step direction is:

$$D^k = - \sum_{i=1}^{n} \frac{\bar{G}_i^k}{(b_i' - \lambda)} V_i^k \tag{6}$$

In the implemented version of the trust region like method, the minimization of Eq. (5) is performed by determining the zero of its first derivative using a Newton–Raphson procedure. However, the convergence threshold of the Newton–Raphson algorithm is guided by a zero value of Eq. (5). Given that the minimum of this function is not necessarily associated with a zero function value, the procedure may fail (number of iteration steps exceeded in subroutine Redstp).

Moreover, this Newton–Raphson search of a zero value first derivative implies that the parameter λ lies in the interval $]b_1, b_2[$. Thus concerning Eq. (6), the step D^k is uphill along the first eigenvector V_1^k and downhill along all the others.

2.2 Algorithm of Simons et al. [7]

The augmented quasi-Newton step:

$$D^k = - \sum_{i=1}^{n} \frac{\bar{G}_i^k}{(b_i^k - \lambda)} V_i^k \tag{7}$$

is used by Simons et al. [7] in an uphill walk from a minimum towards a first-order saddle-point. This can be achieved by maximizing the quadratic approximation $Q^k(D)$ along an eigenvector V_1^k and minimizing it along the others. The selected parameter λ is that:

$$b_1^k < \lambda < \tfrac{1}{2}b_2^k \qquad (8)$$

This choice guarantees that the step has the same orientation as the gradient in the direction V_1^k and the opposite one along all the other directions. Moreover, $Q^k(D^k)$ increases in the direction V_1^k and decreases along the V_i^k directions $(i = 2, \ldots, N)$.

The situation $b_1^k > \tfrac{1}{2}b_2^k$ may occur. Hence, a scaling of the coordinate during the walk is performed. The question arises as to know what λ value in the desirable range should be selected. If b_1^k is negative and b_2^k is positive, then $\lambda = 0$ is a "good choice", following Simons et al. [7]. The step is scaled if it exceeds the trust radius. In the other cases, the parameter is determined in such a way that the augmented quasi-Newton step (7) lies on the boundary of the trust region:

$$\|D(\lambda)\| = R_{\max} \qquad (9)$$

However, it may not exist a parameter λ such that Eq. (9) is verified. The value:

$$\lambda = \tfrac{1}{2}(b_1^k + \tfrac{1}{2}b_2^k) \qquad (10)$$

is then suggested by Simons et al.

This method attempts to calibrate the scaling procedure on the basis of a trust region method. But an arbitrary choice of λ is to be made that results in a truncation of the D^k step. The question thus arises in the same term as in the Schlegel's algorithm, and the problem might be solved:

$$\text{Minimize } (\|D^k(\lambda)\| - R_{\max})^2$$

$$\text{subject to } b_1^k < \lambda < \tfrac{1}{2}b_2^k \qquad (11)$$

In this case, the λ value is the solution of the minimization problem and thus is more suitable than the proposition of Eq. (10). Application of Eq. (7) appears to be efficient to leave an inadequate region but the condition on λ seems to be too stringent.

2.4 Algorithm of Baker [6]

Instead of a quadratic model, a rational model of the energy $E(X)$ about X^k has been proposed by Banerjee et al. [13] and also considered by Baker [6]. This rational function known as Padé approximant is also widely applied in interpolation calculations [14]:

$$E(X^k + D) \cong R^k(D) = E^k + \frac{G^{k^T}D + \tfrac{1}{2}D^T H^k D}{(1 + D^T D)} \qquad (12)$$

Derived from Eq. (12), the augmented quasi-Newton step is given by:

$$(H^k - \lambda I)D^k + G^k = 0 \qquad (13)$$

with

$$\lambda = G^{k^T}D^k \qquad (14)$$

It can be written as:

$$D^k = -\sum_{i=1}^{n} \frac{\bar{G}_i^k}{(b_i^k - \lambda)} V_i^k \tag{15}$$

where λ satisfies:

$$\lambda = \sum_{i=1}^{n} \frac{(\bar{G}_i^k)^2}{(\lambda - b_i^k)} \tag{16}$$

The problem is splitted into a maximization of $R^k(D)$ relative to an eigenvector V_1^k, with a shift parameter λ_p, and into a minimization of $R^k(D)$ relative to the other eigenvectors, with a shift parameter λ_n. In terms of the eigenvectors, two matrix equations derived from Eqs. (13) and (14) can be written as:

$$\begin{pmatrix} b_1^k & \bar{G}_1^k \\ \bar{G}_1^k & 0 \end{pmatrix} \begin{pmatrix} D_1^k \\ 1 \end{pmatrix} = \lambda_p \begin{pmatrix} D_1^k \\ 1 \end{pmatrix} \tag{17}$$

$$\begin{bmatrix} b_2^k & & 0 & \bar{G}_2^k \\ & \ddots & & \vdots \\ 0 & & b_n^k & \bar{G}_n^k \\ \bar{G}_2^k & \cdots & \bar{G}_n^k & 0 \end{bmatrix} \begin{bmatrix} D_2^k \\ \vdots \\ D_n^k \\ 1 \end{bmatrix} = \lambda_n \begin{bmatrix} D_2^k \\ \vdots \\ D_n^k \\ 1 \end{bmatrix} \tag{18}$$

The search direction D^k is calculated from both the eigenvector of Eq. (17) corresponding to the highest eigenvalue λ_p and the eigenvector of Eq. (18) corresponding to the lowest eigenvalue λ_n. However, in practice, D^k is not calculated in that way. The parameter λ_p is the highest solution of a quadratic equation obtained from Eq. (17). λ_n is obtained iteratively via Eq. (16), and the direction D^k is calculated from Eq. (15). The step is scaled if it exceeds a fixed maximum step length R_{max}. The D^k step is an uphill walk along V_1^k and a downhill walk along all other eigenvectors if λ_p and λ_n are positive and negative, respectively.

By definition, this rational approximation leads to conic isocontours of the function values which are not necessarily concentric and similar. For example, given the function $f(X_1, X_2) = (X_1^2 + X_2 - 11)^2 + (X_1 + X_2^2 - 7)^2$, four curvature environments are displayed in Fig. 2.

In the D range -5 to $+5$, the rational approximation isocontours of the function $f(X_1, X_2)$ are drawn near (a) a maximum: 1, (b) a minimum: 2, (c) a first-order saddle-point: 6, and (d) between a minimum: 2, and a first-order saddle-point: 6. Figure 1 displays the isocontours of the analytical function calculated on a regular mesh within the range -5, $+5$ of the two variables. The two drawings (Fig. 1a,b) clearly point out the existence of 1 maximum: 1, 4 minima: 2, 3, 4, 5, and 3 first-order saddle-points: 6, 7, 8, on this analytical surface. Given the rational approximation calculated around the X_1, X_2 values at the center of the pictures (Fig. 2a,b,c,d), the conic characteristic of the contours could be useful to guide the D^k calculation towards the expected first-order saddle-point region.

3 The proposed quasi-Newton algorithm

The quadratic model of the energy $E(X)$ around X^k is derived from the truncated Taylor series expansion:

$$E(X^k + D) \cong Q^k(D) = E^k + G^{k^T}D + \tfrac{1}{2}D^T H^k D \tag{19}$$
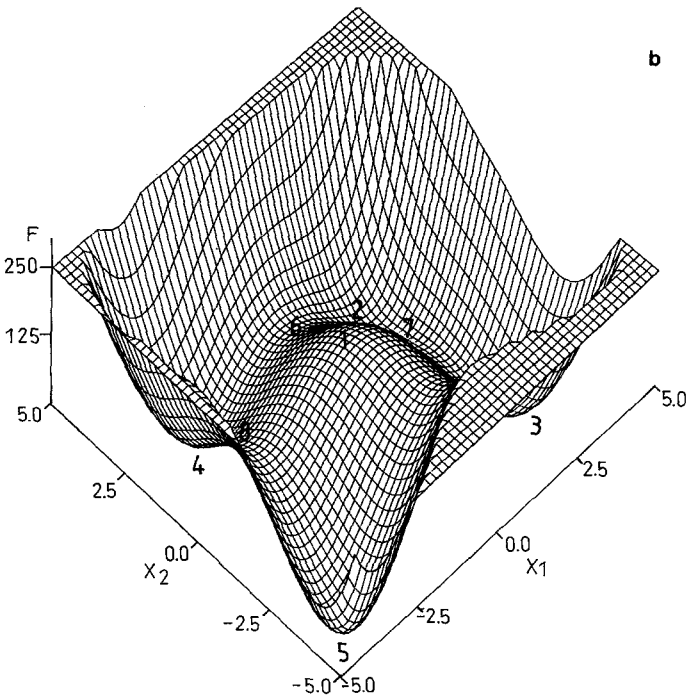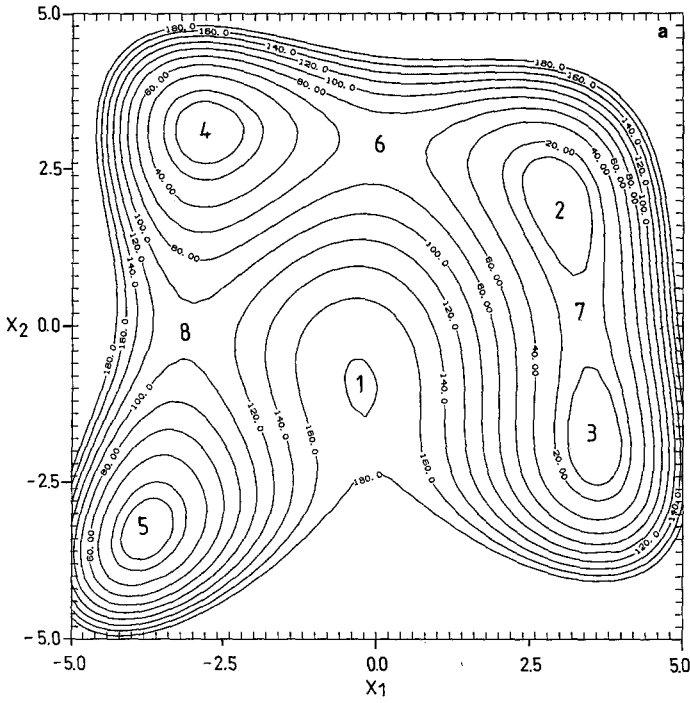
Fig. 1a,b. Isocontours of the analytical function $f(X_1, X_2) = (X_1^2 + X_2 - 11)^2 + (X_1 + X_2^2 - 7)^2$

where E^k, G^k, and H^k are the energy, the gradient and the updated matrix using Powell's formula [9] at point X^k. The confidence in the approximation is warranted by a restricted step method, with a trust radius R_{max} defined as:

$$\|D\|^2 \leqslant R_{max}^2 \tag{20}$$

The transition structure search is performed via a maximization of the quadratic approximation along an eigenvector, V_1^k, and a minimization of the approximation along the other eigenvectors. Following Simons et al. [7], the eigenvector V_1^k, updated at each step, is the eigenvector with the largest overlap on the chosen eigenvector at the previous step, V_1^{k-1}.

The mathematical formalization of the above considerations can be written as:

$$\text{Minimize } F(D') \qquad D' \in \mathbb{R}^{n-1}$$

$$\text{subject to } D'^T D' \leqslant R_2^{k2} \tag{21}$$

where

$$F(D') = \underset{D_1}{\text{Max}} \left\{ Q(D_1, D') \mid D_1^2 \leqslant R_1^{k2} \text{ and } D_1 \in \mathbb{R}^1 \right\}$$

with $D' = (D_2, \ldots, D_n)$. R_1^k and R_2^k are positive scalars. The suggested solution of those conditions can be formulated as an augmented quasi-Newton step with a positive shift parameter λ:

$$D^k = -\frac{\bar{G}_1^k}{(b_1^k - \lambda)} V_1^k - \sum_{i=2}^{n} \frac{\bar{G}_i^k}{(b_i^k + \lambda)} V_i^k \tag{22}$$

By comparison with the previous solution of the quadratic approximation given by Eqs. (6) and (7), and the rational approach of Baker of Eq. (15), D^k calculation is split into two parts. The first part is relative to the maximization of $Q^k(D)$ along one direction; the other concerns the minimization in the supplementary subspace.

To solve the Min-Max problem of Eq. (21), the parameter λ is chosen so that the three following conditions are fulfilled:

(i) the search direction vector lies within a fixed trust region of radius R_{max},

(ii) $(b_1^k - \lambda)$ is negative to obtain an uphill step along the eigenvector V_1^k,

(iii) $(b_i^k + \lambda)$ is positive to obtain a downhill step along the other eigenvectors V_i^k, $i = 2, \ldots, N$.

Under these conditions, the quadratic approximation increases along the direction V_1^k and decreases along all the other directions.

If the (approximate) Hessian has the expected structure and if the quasi-Newton step lies within the boundary of the trust region, then the quasi-Newton step is selected. Otherwise, the step is chosen on the boundary of the trust region, with λ as the solution of:

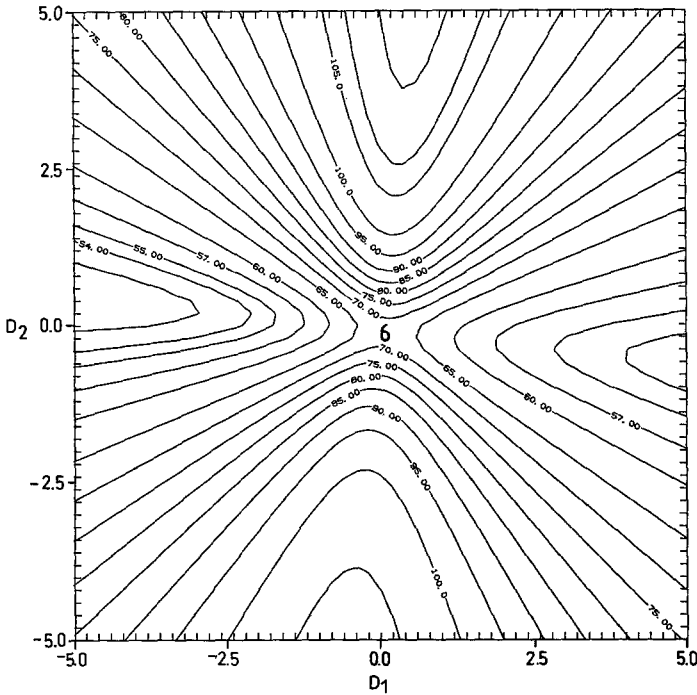$$\|D^k(\lambda)\|^2 = R_{max}^2 \tag{23}$$

This step calculation, well suited to the trust region model, is appropriate for a first-order saddle-point search and can be easily generalized to higher order saddle-point search.
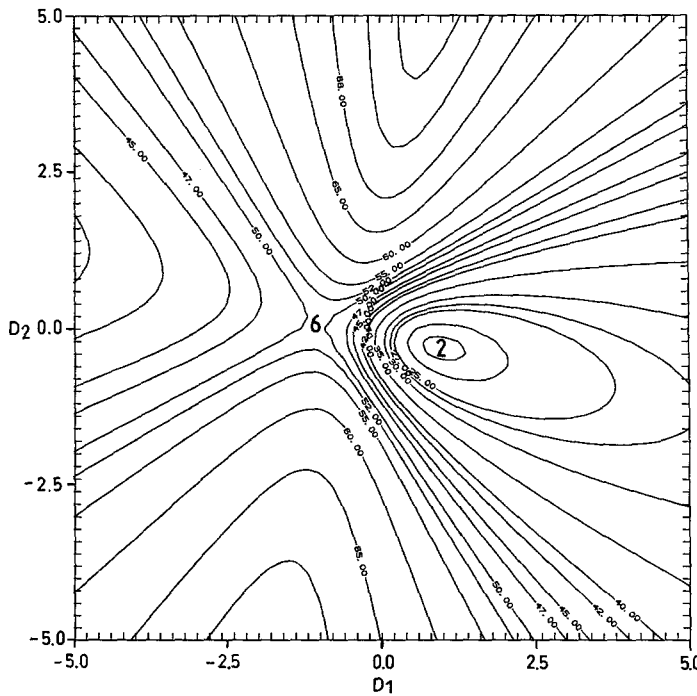
a

b

c
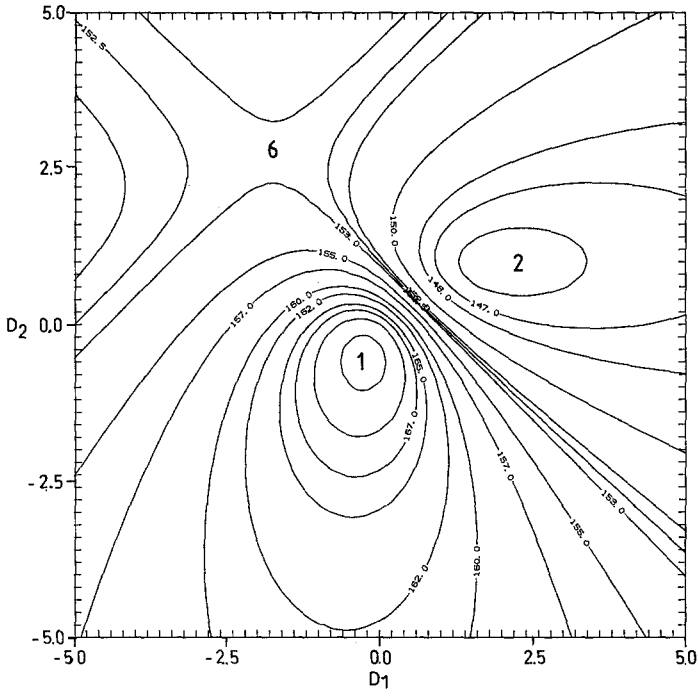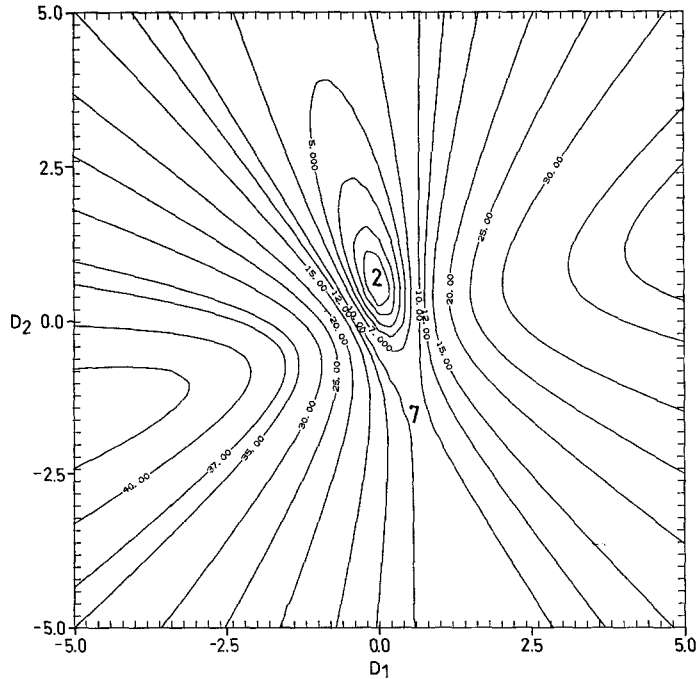
Fig. 2a–d. Isocontours of the rational approximation $R(D)$ of the function $f(X_1, X_2)$. **a** near the maximum 1 with $X_1 = 0.0$, $X_2 = 0.0$; **b** near the minimum 2 with $X_1 = 3.0$, $X_2 = 1.5$; **c** near the saddle-point 6 with $X_1 = 0.0$, $X_2 = 3.0$; **d** between the minimum 2 and the saddle-point 6 with $X_1 = 1.719$, $X_2 = 2.75$

d

4 Analysis of the numerical procedure

The proposed algorithm, called QA, based on the augmented quasi-Newton step of Eq. (22), is implemented in the Gaussian package. The flowchart (Fig. 3) illustrates the organization of the levels involved in the iterative procedure.

1. Initialization
- Choose a guess $X°$ and an initial trust radius $R°$.
- At the point $X°$, calculate the energy function $E°$, the gradient vector $G°$, and the Hessian matrix $H°$ or an approximation of it.
- Set $k = 0$.

2. Eigenvalues and eigenvectors
- Calculate the eigenvalues b_i^k and the eigenvectors V_i^k of H^k, $i = 1, \ldots, N$.
- Calculate the components of the gradient vector $\bar{G}_i^k$ in the eigenvector basis, $i = 1, \ldots, N$.

3. Choice of an eigenvector
- Determine the eigenvector along which the energy is to be maximized: V_j^k, $j \in [1, N]$:
  - at the first iteration, the eigenvector is associated with the smallest eigenvalue, except in the case of a user choice;
  - at the subsequent steps, V_j^k is the eigenvector with the largest overlap on the chosen eigenvector at the previous step, V_j^{k-1}.
- Determine the inertia of $H^k = (n_-^k, n_0^k, n_+^k)$, where n_-, n_0, n_+ are the numbers of negative, zero, and positive eigenvalues.
- Set $b_{\min} = \min\{b_i^k; i \neq j, i = 1, \ldots, N\}$.

4. Quasi-Newton step
- If $\|H^{k-1}G^k\| \leqslant R^k$ and $n_-^k = 1$ then take the quasi-Newton step D^k as:

$$D^k = -\sum_{i=1}^{n} \frac{\bar{G}_i^k}{b_i^k} V_i^k$$

- Else select the augmented quasi-Newton step D^k as:

$$D^k(\lambda) = -\frac{\bar{G}_j^k}{(b_j^k - \lambda)} V_j^k - \sum_{\substack{i=1 \\ i \neq j}}^{n} \frac{\bar{G}_i^k}{(b_i^k + \lambda)} V_i^k$$

where the parameter λ is the zero of the function:

$$\|D^k(\lambda)\|^2 - R^{k2}$$

in the interval $]\max\{0, b_j^k, -b_{\min}\}, +\infty[$.

5. Prediction
- Calculate the predicted change in energy following the quadratic energy function approximation:

$$Ce = Q^k(D^k) - E^k = G^{kT}D^k + \tfrac{1}{2}D^{kT}H^kD^k$$

6. Trust region
- Calculate the energy E^* at the point $X^k + D^k$.
- Set $r = (E^* - E^k)/Ce$.
- If $r \leqslant r_{\min}$ or $r \geqslant (2 - r_{\min})$ then $R^{k+1} = R^k/Sc$.
- If $r \geqslant r_{\text{good}}$ and $r \leqslant (2 - r_{\text{good}})$ and $D^k \neq -H^{k-1}G^k$ then $R^{k+1} = R^k \cdot (Sc)^{1/2}$.
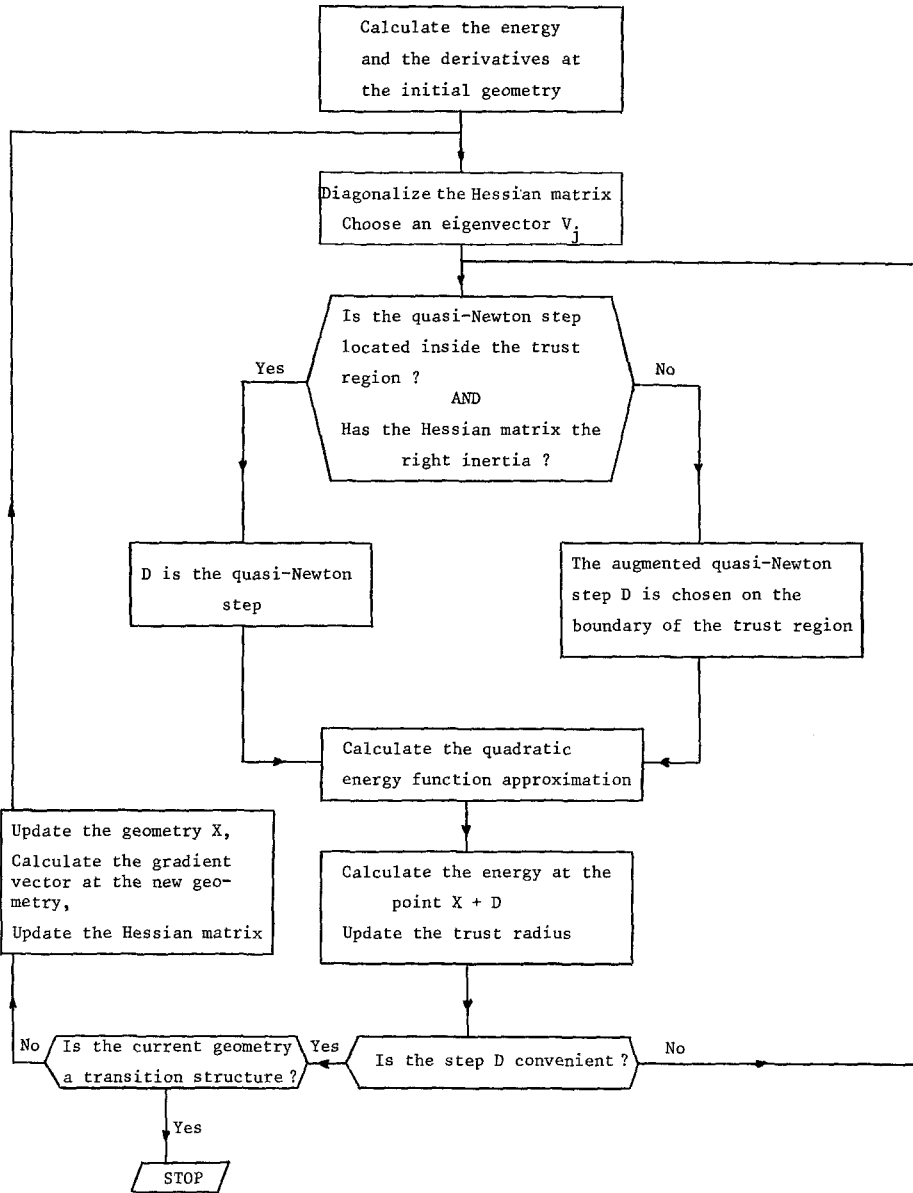$r_{\min} = 0.75$, $r_{\text{good}} = 0.8$ and $Sc = 2$ are the implemented values.

Calculate the energy
and the derivatives at
the initial geometry

Diagonalize the Hessian matrix
Choose an eigenvector V_j

Is the quasi-Newton step
located inside the trust
region ?
AND
Has the Hessian matrix the
right inertia ?

Yes   No

D is the quasi-Newton
step

The augmented quasi-Newton
step D is chosen on the
boundary of the trust region

Calculate the quadratic
energy function approximation

Update the geometry X,
Calculate the gradient
vector at the new geo-
metry,
Update the Hessian matrix

Calculate the energy at the
point X + D
Update the trust radius

No   Is the current geometry   Yes   Is the step D convenient ?   No
a transition structure ?

Yes

STOP

Fig. 3. QA flowchart

7. Adequation of step D^k
- If $r < 0$ or $r > 2$

7a, then, estimate a new D^k step at the same X^k coordinate but using the
updated R^{k+1} trust radius:
- $X^{k+1} = X^k$,
- $E^{k+1} = E^k$,
- Set $k = k + 1$,
- Goto #4;

7b. else, update the X^k vector and the iteration count.

  • Check the four critera on convergence:

$$\|G^k\|_1 \leqslant \varepsilon_1, \qquad \|G^k\|_{RMS} \leqslant \varepsilon_2$$

$$\|D^k\|_1 \leqslant \varepsilon_3, \qquad \|D^k\|_{RMS} \leqslant \varepsilon_4$$

  • If they are fulfilled, then STOP: X^k is a first-order saddle-point.
  • $X^{k+1} = X^k + D^k$,
  • $E^{k+1} = E^*$,
  • Calculate the gradient vector G^{k+1} at X^{k+1}.
  • Update the approximate Hessian matrix using the symmetric formula of Powell:

$$H^{k+1} = H^k + \frac{1}{D^{k^T}D^k} \left\{ yD^{k^T} + D^k y^T - \frac{(y^T D^k)(D^k D^{k^T})}{(D^{k^T}D^k)} \right\}$$

  where $y = G^{k+1} - G^k - H^k D^k$,
  • Set $k = k + 1$.
  • Goto #2.

5 Application

QA algorithm applications have been compared to the results obtained by the procedures TS by Schlegel [5], and EF by Baker [6], implemented in the Gaussian packages [8]. The QA subroutines are included in the link 113 of the GAUSSIAN 86 version running on a FPS 264 attached processor and the GAUSSIAN 88 running on a VAX 11/780.

In all the analyzed structures, the first estimation of the Hessian matrix is computed analytically. A default maximum stepsize or an initial trust radius of 0.3 Bohr/Radian is considered. The convergence thresholds are kept as the standard default values on the maximum and RMS gradients and the maximum and RMS displacements are scaled to 450×10^{-6}, 300×10^{-6}, 1800×10^{-6}, 1200×10^{-6} a.u., respectively. The initial starting and final transition state internal coordinates as well as the number of cycles to reach convergence are given for each of the selected examples shown below.

In most cases, convergence using the Schlegel's original algorithm (Eq. (4)) requires a greater number of iterations than QA and EF. Moreover, using the non-Newton−Raphson scale option (Eqs. (5) and (6)) of the program, the walking search diverges and stops within the Redstp procedure (see Sect. 2.1).
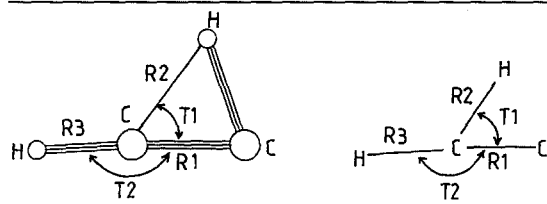
The first two examples, i.e., the acetylene-ethenylidine rearrangement and the isomerization of the methoxy radical, are often used in the literature [15] to illustrate the different transition state search performance.

The other examples are related to enzymatic reactions. Several Hessian inertia conditions and curvature amplitudes are considered.

5.1 Acetylene-ethenylidine rearrangement RHF/3-21G

The initial guess of the geometry optimization is characterized by a higher energy value than the energy calculated at the transition-state solution [6]. At the starting geometry, the Hessian matrix has the right inertia. The location of the

Table 1. Acetylene-ethenylidine rearrangement

	Starting	Final
R1	1.24054	1.247
R2	1.65694	1.428
R3	1.06318	1.056
T1	60.3568	54.2
T2	150.3568	176.6
First eigenvalue	−0.35638	−0.39251

Number of iterations TS No NRSCale: failed EF: 7 QA: 10
TS NRSCale: 21

equilibrium structure thus requires the decrease of the energy function while the required negative eigenvalue is retained.

Both QA and EF reach this goal after 10 and 7 iterations, respectively (Table 1). With the default option of GAUSSIAN 88, TS converges to the solution within 21 calculations, while Redstp problems occur when the non-Newton–Raphson scale option is used.

## 5.2 Methoxy radical isomerization RHF/STO-3G

In this compound, the starting geometry used by Baker et al. [16] is described as a "branching" point with one null eigenvalue. On the way to the saddle-point, the different algorithms have to force the appearance of one negative eigenvalue. In this case, the three results are similar to those obtained with the HCCH ↔ CCH$_2$ isomerization (Table 2).

5.3 Formamide OH$^-$ complex RHF/STO-3G

The third example concerns the reaction model of the formamide hydrolysis by the hydroxide ion. In view to investigate the mechanism by which an enzyme catalyzes chemical reactions, amidic bond cleavage has been intensively studied in the gas phase and in aqueous solution by Wiener et al. [17].

The first approach proposed by Alagona et al. [18] describes the chemical pathway going from a tetrahedral intermediate to a formate-ammonia adduct through a four-membered ring transition-state structure.

For the three analyzed algorithms, the starting internal coordinates are those of Alagona et al. At this point, the Hessian matrix has one negative eigenvalue, and its associate eigenvector includes the degrees of freedom of the cycle

Table 2. Methoxy radical isomerization
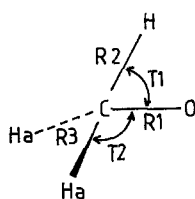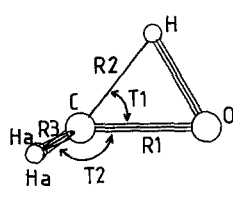
P1 = HCOHa

	Starting	Final
R1	1.423	1.423
R2	1.484	1.326
R3	1.087	1.089
T1	42.7	51.8
T2	117.5	117.5
P1	105.4	106.4
First eigenvalue	−0.00969	−1.36810

Number of iterations TS No NRSCale: failed EF: 11 QA: 14
                              TS NRSCale: 10

rearrangement. Again, the numerical performance of the three algorithms is in agreement with previous results (Table 3).

5.4 Formamide-water complex RHF/3-21G

The energy hypersurface generated by the formamide-water complex is very flat. In 3-21G [19], one conformational minimum given by Jasien [20] describes a hydrogen bond between the carbonyl oxygen of formamide and a water hydrogen. This geometry is used as a starting point on the search towards another local minimum with a hydrogen bond between the water oxygen and the formamide hydrogen. So the initial Hessian matrix has only positive eigenvalues and the energy function value is lower than the value at the transition state structure. TS does not converge to a solution whatever the selected optimization options. Both QA and EF reach the transition structure after 9 and 12 iterations respectively (Table 4).

5.5 Methanolysis of protonated methyl-formic-ester RHF/STO-3G

Given the extremely low curvature of the energy hypersurface, the location of the critical points generated by a protonated system is particularly difficult. In the reaction studied, a hydrated methanol transfers a proton to the methoxy ester group through the flip-flop of the water molecule. Several geometrical arrangements can be trapped depending on the dihedral angles that defined the six-membered ring at the transition state. Two of them have been investigated. In the first one (Fig. 4a), QA converges to the solution after 20 iterations. The eigenvector

Table 3. Formamide OH⁻ complex
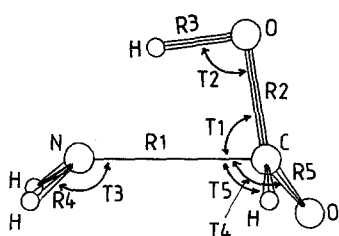
	Starting	Final
R1	2.1	2.117
R2	1.385	1.409
R3	1.096	1.028
R4	1.035	1.045
R5	1.232	1.234
T1	84.29	80.2
T2	84.94	91.5
T3	126.61	120.9
T4	111.52	115.3
T5	89.47	90.7
First eigenvalue	−0.21042	−0.23683

Number of iterations TS No NRSCale: failed EF: 10 QA: 10
　　　　　　　TS NRSCale: 15

Table 4. Formamide-water complex
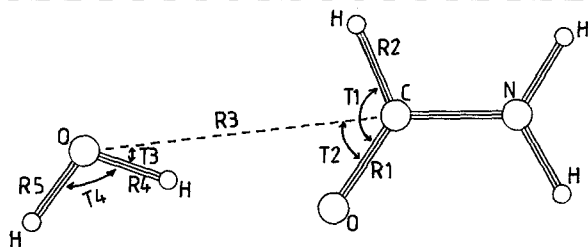
	Starting	Final
R1	1.2236	1.221
R2	1.0779	1.080
R3	2.8900	3.461
R4	0.9719	0.973
R5	0.9641	0.966
T1	121.116	122.0
T2	69.890	49.4
T3	49.568	25.3
T4	109.692	107.9
First eigenvalue	0.00894	−0.00045

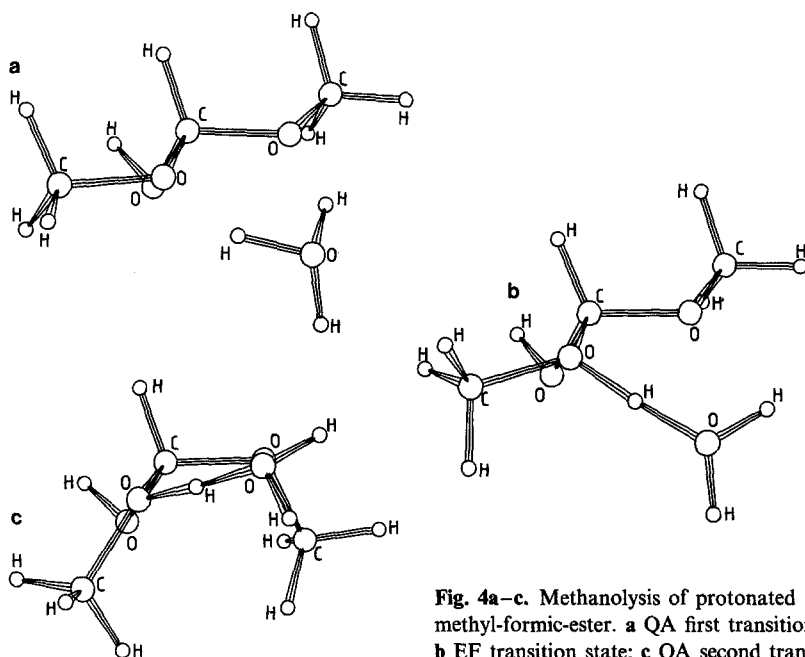Number of iterations TS No NRSCale: failed EF: 12 QA: 9
　　　　　　　TS NRSCale: failed

Fig. 4a–c. Methanolysis of protonated methyl-formic-ester. a QA first transition state; b EF transition state; c QA second transition state

components associated to the -0.00139 eigenvalue well retain the important parameters involved in the cycle reorganization. After 94 iterations, EF converges to a transition state in which only the torsional angles of the methyl rotation are concerned. Hence this critical point resembles a complex between protonated ester and the couple methanol-water (Fig. 4b).

Similar results are obtained using the second conformational geometry (Fig. 4c) in which mainly the torsional angles of the cycle differ from the preceding geometry. The optimization solution is obtained by QA after 64 iterations with the right eigenvector components associated to a -0.00107 eigenvalue while, after 116 iterations, EF slowly slips again to the previously described complex (Fig. 4b). Going from iteration 30 to iteration 70, several negative eigenvalues appear and the main components of the first eigenvector combine at once both the cycle motion and the methyl rotation. Forcing the algorithm to retain the right Hessian inertia, EF selects the methyl rotation components in the eigenvector. This numerical behaviour gives rise to a drastic geometric change in the conformation of the H_3O^+ entity.

With respect to the low curvature of the surface, the barrier height determined by the two transition states is 0.893 kcal for the first structure and 3.81 kcal for the second one, respectively.

6 Conclusion

This paper deals with algorithms involved in a transition-state structure location on an energy hypersurface. Available quadratic and rational models are analyzed in detail and a new quasi-Newton method is proposed. Regarding the efficient Baker algorithm, this newly developed algorithm solves the problem of step

estimation as an augmented quasi-Newton displacement by adding a positive shift parameter.

In practice, the selected examples emphasize the stability of the method using Hessian matrix whatsoever its inertia. The efficiency is well illustrated by the 18-atom system associated with a very low curvature surface. The convergence in those three reactions where covalent bonds are formed and/or broken is remarkable.

Two problems require further investigation. The first one is related to the stepsize calibration in order to avoid too long displacement, particularly, on a flat hypersurface. The second one is inherent to the matrix inertia. This feature is critical when more than one negative eigenvalue occur with a very weak absolute value. This nearly rank loss may be caused by redundant geometrical parameters or local symmetry conditions.

References

1. McIver JW, Komornicki A (1972) J Am Chem Soc 94:2625–2633
2. Murrell JN, Laidler KJ (1968) Trans Faraday Soc 64:371–377; and Murrell JN, Pratt GL (1970) Trans Faraday Soc 66:1680–1684
3. Eyring H, Polanyi M (1931) Z Physikal Chem 12B:279–311
4. Fukui K (1981) Acc Chem Res 14:363–368
5. Schlegel HB (1982) J Comp Chem 3:214–218
6. Baker J (1986) J Comp Chem 7:385–395
7. Simons J, Jorgensen P, Taylor H, Ozment J (1983) J Phys Chem 87:2745–2753
8. GAUSSIAN 86–GAUSSIAN 88, Frisch MJ, Head-Gordon M, Schlegel HB, Raghavachari K, Binkley JS, Gonzalez C, Defrees DJ, Fox DJ, Whiteside RA, Seeger R, Melius CF, Baker J, Kahn LR, Stewart JJP, Fluder EM, Topiol S, Pople JA, Gaussian Inc, Pittsburgh, PA
9. Powell MJD (1971) Math Prog 1:26
10. Greenstadt J (1967) Math Comp 21:360–367; and Bard Y (1974) Nonlinear parameter estimation. Academic Press, NY, p 91–94
11. Golab JT, Yeager DL, Jorgensen P (1983) Chem Phys 78:175–199
12. Fletcher R (1987) Practical methods of optimization. Wiley, NY
13. Banerjee A, Adams N, Simons J, Shepard R (1983) J Phys Chem 89:52–57
14. Cuyt A (ed) (1988) Nonlinear numerical methods and rational approximation. Reidel Publ, Dordrecht
15. Cerjan CJ, Miller WH (1981) J Chem Phys 75:2800–2806; Poppinger D (1975) Chem Phys Letts 35:550–554; and Muller K, Brown LD (1979) Theoret Chim Acta 55:75–93
16. Baker J, Gill PMW (1988) J Comp Chem 9:465–475
17. Wiener SJ, Chandra Singh U, Kollman PA (1985) J Am Chem Soc 107:2219–2229
18. Alagona G, Scrocco E, Tomasi J (1975) J Am Chem Soc 97:6976–6983
19. Dehareng D, Dive G, Ghuysen JM (1991) Theoret Chim Acta 81:281–290
20. Jasien PG, Stevens WJ (1986) J Chem Phys 84:3271–3277