Improved RGF Method to Find Saddle Points

MICHAEL HIRSCH, WOLFGANG QUAPP

Mathematisches Institut, Universität Leipzig, Augustus-Platz, D-04109 Leipzig, Germany

Abstract: The predictor-corrector method for following a reduced gradient (RGF) to determine saddle points [Quapp, W. et al., J Comput Chem 1998, 19, 1087] is further accelerated by a modification allowing an implied corrector step per predictor but almost without additional costs. The stability and robustness of the RGF method are improved, and the new version in addition reduces the number of gradient and Hessian calculations.

© 2002 Wiley Periodicals, Inc.   J Comput Chem 23: 887–894, 2002

Key words: saddle point; reduced gradient; gradient extremal; reaction path; continuous Newton method

Introduction

In 1998, the old distinguished coordinate method[1] was reactivated by the so-called "reduced gradient following" (RGF)[2,3,4] using another mathematical point of view. Equivalent curves are obtained also by the continuous Newton method (the Branin trajectories).[3,5,6] The RGF has been proved to be an effective tool in determining saddle points (SP)[7–13] on a potential energy surface (PES). Branin's method is additionally well adapted to exactly calculate symmetric valley-ridge-inflection (VRI) points that are bifurcation points of RGF curves.[4,12,14] As a rule, VRI points indicate the branching of reaction paths.

The idea of RGF is to define (as simply as possible) a curve that connects the minimum with the corresponding SP of interest, or with the VRI point, respectively, and to numerically follow this curve by predictor and corrector steps.[15] Starting at a minimum, every SP connected with the region of attraction of that minimum will be found. Kliesch[16] uses a similar concept, but does not consider the most adequate and simple definition of the curve to follow, compare eq. (1) in ref. 16. A modification of RGF to follow not only an arbitrary defined curve, but really the valley floor line, was termed the "tangent search concept" (TASC).[17–19] TASC allows the calculation of the gradient extremal (GE) along the valley floor by second order methods only working in a very strict analogy to the RGF method. (A GE is defined by $\mathbf{Hg} = \lambda\mathbf{g}$ with the Hessian matrix, $\mathbf{H}$, with the gradient, $\mathbf{g}$, and λ is the eigenvalue of the Hessian.) For TASC the improvements shown here are valid as well.

The basic strategy of any path following is the continuation of a curvilinear curve by a predictor step along the tangent, and the subsequent correction of the error of that linear step by a corrector step (by a Newton-Raphson-like method). The corrector step is orthogonal to the tangent and brings the predicted point back to the solution curve.[15] Here we propose the linear combination of the predictor and the corrector step at the current point in one step,[6] forming a new kind of predictor step.

Our methods are programmed in FORTRAN as independent modular codes. The programs can be obtained on request or downloaded (http://www.mathematik.uni-leipzig.de/MI/quapp). Any comments are welcome.

Algorithm

$E(\mathbf{x})$ is the function of the PES, $\mathbf{g}(\mathbf{x})$ is its gradient vector in the configuration space, and $\mathbf{R}^n$ is defined by the coordinates $\mathbf{x}$ of the molecule. As usual, $n = 3N - 6$ forms the number of independent internal coordinates, the dimension of the problem. $\mathbf{x}$ and $\mathbf{g}$ are vectors of this dimension n. RGF finds a curve where the selected gradient direction is fixed at every curve point, $\mathbf{x} = \mathbf{x}(t)$:

$$\mathbf{g}(\mathbf{x}(t))/\|\mathbf{g}(\mathbf{x}(t))\| = \mathbf{r} \tag{1}$$

where t is the curve parameter, and $\mathbf{r}$ is the unit vector of the fixed search direction.[3] The search direction usually corresponds with the start direction of a chemical reaction, or it may point into the direction of the SP, or it may be a coarse estimation of those directions. We recall that RGF curves generally are not minimum energy pathways.[2,9] The RGF curve eq. (1) may be near a valley floor line or not. Nevertheless, these curves may follow a reaction path in favorable cases, at least qualitatively. The possibility of the calculation of the minimum energy path then depends on a clever definition of the search direction.[21]

To realize the requirement eq. (1), the RGF algorithm[3] uses a projection of the gradient of the PES to fulfill the equations:

$$\mathbf{P_r}\mathbf{g}(\mathbf{x}(t)) = \mathbf{0} \qquad (2)$$

where the projector $\mathbf{P_r}$ is formed by $(n-1)$ rows of unit vectors being orthogonal to the direction $\mathbf{r}$: it holds $\mathbf{P_r}\mathbf{r} = \mathbf{0}$. Then $\mathbf{P_r}\mathbf{g}$ generally gives that part of $\mathbf{g}$ that is orthogonal to $\mathbf{r}$. If that part is zero then the *reduced gradient* eq. (2) is the $(n-1)$-zero vector. The projector, $\mathbf{P_r}$, is a constant matrix of dimension $(n-1) \times n$, thus of rank $(n-1)$: that one which enforces the gradient to point at every curve point, $\mathbf{x}(t)$, into the same direction $\mathbf{r}$; or, in mathematical terms, $\mathbf{g}(\mathbf{x})$ is an element of the linear span of $\mathbf{r}$. The tangent to curve eq. (2), $\mathbf{x}'(t)$, is obtained by a solution of the following system of equations:

$$\frac{d}{dt}[\mathbf{P_r}\mathbf{g}(\mathbf{x}(t))] = \mathbf{P_r}\frac{d\mathbf{g}(\mathbf{x}(t))}{dt} = \mathbf{P_r}\mathbf{H}(\mathbf{x}(t))\mathbf{x}'(t) = \mathbf{0} \qquad (3)$$

where $\mathbf{H}$ is the Hessian of the PES. The simplicity of the RGF method is based on the constance of the $\mathbf{P_r}$ matrix. The predictor step is done along the tangent $\mathbf{t} = \mathbf{x}'(t)/\|\mathbf{x}'(t)\|$, and, orthogonal to this direction, Newton-Raphson-like steps of the corrector for a point near to curve eq. (2) are calculated. Using TASC, only the projector matrix is changed into $\mathbf{P_t}$ after the predictor steps.[17–19] Note that in refs. 8 and 10 other predictor strategies are proposed, while in ref. 13 two alternate corrector schemes are applied.

The corrector method by Newton-Raphson-like steps[15] starts with the $(n-1) \times n$ matrix equation, for $\mathbf{c}$ (see the dashed step $\mathbf{x}_2 \rightarrow \mathbf{x}_3$ in Fig. 1):

$$\mathbf{P_r}\mathbf{H}(\mathbf{x}_2)\mathbf{c} = -\mathbf{P_r}\mathbf{g}(\mathbf{x}_2) \qquad (4)$$

However, to make the solution unique, the system is augmented to a full $n \times n$ system by adding an equation with the scalar product

$$\mathbf{t}(\mathbf{x}_2)^T\mathbf{c} = 0 \qquad (5)$$

enforcing that the corrector step is orthogonal to the tangent of an RGF curve at $\mathbf{x}_2$. Diener and Schaback proposed[6] avoiding the conventional predictor step, $p\mathbf{t}$ with step length p, and instead determining a combined step τ starting with $\mathbf{t}$ at $\mathbf{x}_1$ by solving eq. (5) not orthogonal to $\mathbf{t}$ but skewed with the scalar product:

$$\mathbf{t}(\mathbf{x}_1)^T\tau = p \qquad (6)$$

where p is the step length of the former predictor step. This results in a new Newton-Raphson-like step by solution of the augmented linear system of equations:

$$\mathbf{P_r}\mathbf{H}(\mathbf{x}_1)\tau = -\mathbf{P_r}\mathbf{g}(\mathbf{x}_1)$$
$$\mathbf{t}(\mathbf{x}_1)^T\tau = p \qquad (7)$$

(see the steps depicted by full polygonal lines in Fig. 1). The step vector τ is a combined predictor-corrector step with the component step length p in the direction of $\mathbf{t}$, which should give a point $\mathbf{x}_4$ near
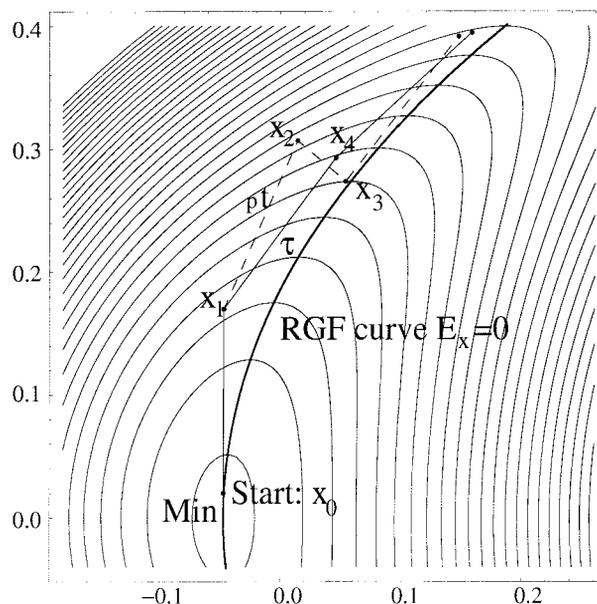
Figure 1. Model potential surface[23] of eq. (10) with the boldfaced RGF curve $E_x = 0$. Dashed lines show the former RGF version of predictor ($p\mathbf{t}$) and corrector steps ($\mathbf{c}$) from point $\mathbf{x}_1$ to $\mathbf{x}_2$ and to $\mathbf{x}_3$ and so on. The polygonal approximation given by solid lines is the improved RGF method with implied corrector proposed here.

$\mathbf{x}_3$, being consequently near the searched curve (RGF or TASC) fulfilling also

$$\|\mathbf{P_r}\mathbf{g}(\mathbf{x}_4)\| < \varepsilon \qquad (8)$$

with the threshold ε. In general, only if eq. (8) is unsatisfied do we need to take further corrector steps as defined by eqs. (4) and (5); but generally, the use of steps τ avoids these separate corrector steps $\mathbf{c}$. It holds especially for a large scaling of ε. The former proposal in the original RGF was to set $\varepsilon \approx (0.01 \ to \ 0.1) \times p$. Now, ε may be as large as p or larger. For the original RGF, the inclusion of the corrector via ε of eq. (8) is not avoided because the predictor alone goes wrong, if it is not corrected from time to time.

The proposed use of an implied corrector step in every predictor is a kind of automatic improvement of the predictor direction. If this direction is defined by a fixed search direction for the gradient as in the RGF method, the method works well. In the case of the TASC the search direction itself is changed along the course of the calculation.[17] The search direction for TASC is the tangent $\mathbf{t}$ of the momentary RGF curve. The two processes (the calculation of τ with eq. (7) and the change of search direction in TASC) can add and may lead to an overshooting of the resulting predictor step. We found an easy and global solution to suppress this effect: we do not use the full corrector part of step τ of eq. (7) but only a certain ratio of corrector to predictor. Most experiments show that ratio 1/2 is a good setting. In the case of the very notorious example of the Rosenbrock surface, see below, we used the ratio 1/3:

$$\text{TASC\_p\_step} = \frac{1}{3}(2p\mathbf{t} + \tau) \qquad (9)$$

A scheme is given in the appendix showing the algorithms for RGF and TASC.

Possibility of Hessian Update

Beginning with the first version of RGF,[2] the Davidon-Fletcher-Powell (DFP) update of the Hessian matrix was included. The update also works in the case when the index of the Hessian at the minimum changes into the SP index.[2,22] Of course, this is the condition for an update to serve for a search of pathways from minimum to SP. In the new version of RGF/TASC we include the possibility of using Bofill's update,[20] which is well accepted in this field of computations.

Examples

Test Surface 1

The first example is the 2D polynomial PES of Lami and Villani[23] describing the ion $O_2H_5^+$:

$$E(x, y) = vx + qx^2 + rx^3 + sx^4 + (a + bx + cx^2)y^2$$
$$+ (d + ex + fx^2)y^4 \quad (10)$$

with the constants $v = 0.0066$, $q = 0.0661$, $r = -0.052$, $s = 0.0345$, $a = 0.0096$, $b = -0.1899$, $c = 0.0825$, $d = 0.1213$, $e = -0.0366$, and $f = -0.0237$. This simple test PES gives the possibility of an accurate, in-depth analysis of the proposed method. Figure 1 shows the region around the minimum (-0.047, 0.0) and the action of the RGF method. The start x_0 is near the minimum at the RGF curve $E_x = 0$. RGF makes a first step along the fixed search direction (0, 1). The dashed curve is the original RGF with predictor step length $p = 0.15$ and with the maximal aberration in eq. (8) of less than $\varepsilon = 0.008$. At x_1, the ε-condition is still fulfilled and the method makes a further predictor step. However, at x_2, the ε-criterion is not fulfilled, and RGF makes the corrector step leading to x_3. The new RGF with implied corrector (depicted by full polygonal lines) starts at x_0 with the same step, because this point belongs to the RGF curve. At x_1, the hypothetical corrector step from x_1 to the RGF curve $E_x = 0$ (the gradient component E_y is the search direction) is added to the predictor step $pt = (x_2 - x_1)$ to give the new predictor $\tau = (x_4 - x_1)$. The new RGF smoothly attends the true RGF curve: the series of the new RGF points goes very near the true RGF curve $E_x = 0$ up to the SP (1.361, 1.318), but also the original RGF does not need a further corrector step, in this case of a quite straight valley up to the SP.

If the predictor-corrector threshold is extended, for example nearly doubled to $\varepsilon = 0.015$ using further $p = 0.15$, then the former original RGF needs again one corrector step, but the new RGF strategy does not need a separate corrector step.

Figure 2 shows the analogous work of the original, first version of the TASC method as well as the TASC method with the implied corrector. The parameters are somewhat larger than in Figure 1: $p = 0.2$, $\varepsilon = 0.02$. Note that TASC automatically follows the
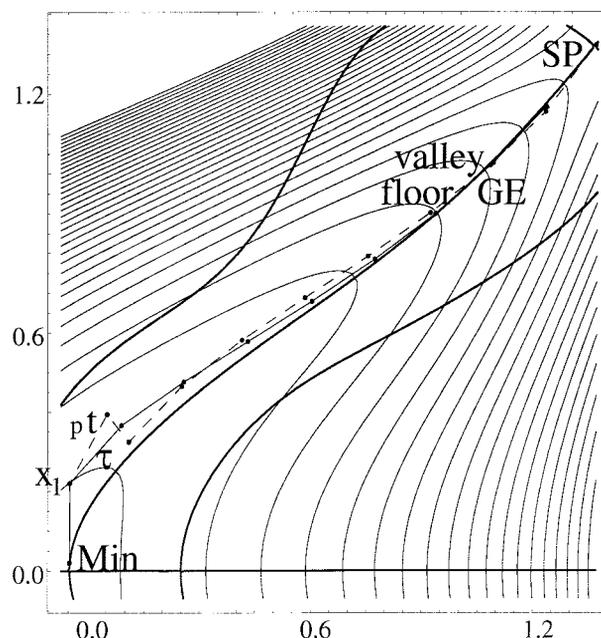
Figure 2. Model potential surface[23] of eq. (10) where the boldfaced curves are the gradient extremals. The original (dashed) TASC version of the usual predictor + corrector strategy is shown to approximate the valley floor gradient extremal, and the new TASC version (solid lines) of combined predictor and corrector steps τ [eqs. (7) and (9)] is shown to solve the same task.

valley floor line, the GE, already in its original definition.[17] This is due to the artifice to change the projector matrix into P_t along the pathway, step by step. Also, if the valley floor is strongly curvilinear, steps τ are successful, as well as in the RGF case (see the steps pt or τ from point x_1, correspondingly). In Figure 2, TASC ends at the SP on top of the valley floor line by a full-dimensional Newton-Raphsen step. Here, the internal step length search of TASC[19] does not enforce a shortening of the maximal step length because the valley floor line near the SP is quite straight.

If the corrector threshold is reduced, for example to $\varepsilon = 0.01$, using further $p = 0.2$, then the original TASC needs two more corrector steps to reach the SP, but also the new TASC strategy needs two corrector steps.

The PES of HCP

Points of the PES of HCP are calculated by the restricted Hartree-Fock method with the 6-31G* basis set using the GAMESS_UK program.[24] The starting point is the global minimum at the linear HCP geometry, and the search direction of RGF is the pure bending direction of the H atom. The pathway along this valley is not very ambitious. The test calculations for the comparison of the two different RGF versions are done with five different predictor step lengths, p, and also five different thresholds, ε, for the predictor/corrector decision. The results are given in Table 1. For higher values of p or ε, the original RGF method often runs into a problem: RGF stops before the SP, anywhere near the valley pathway. We use for the stopping criterion the smallness of a

Table 1. RGF Tests for the Saddle Point Search (for the CPH Linear Structure) on the PES of HCP.

$p = 0.08^a$	RGF w.i.C.[b]		Old RGF	
Threshold	P	C	P	C
0.008	122	0	122	5
0.016	122	0	122	3
0.024	122	0	122	2
0.032	122	0	122	2
0.040	122	0	122	3

$p = 0.16$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.008	61	0	61	5
0.016	61	0	61	2
0.024	61	0	61	2
0.032	61	0	61	1
0.040	61	0	61	1

$p = 0.24$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.008	41	0	41	4
0.016	41	0	41	2
0.024	41	0	—	—[c]
0.032	41	0	41	2
0.040	41	0	41	3

$p = 0.32$	RGF w.i.C.[b]		Old RGF	
Threshold	P	C	P	C
0.008	31	1	31	5
0.016	31	0	—	—[c]
0.024	31	0	31	1
0.032	31	0	31	0
0.040	31	0	31	0

$p = 0.40$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.008	25	2	25	7
0.016	25	0	—	—[c]
0.024	25	0	—	—[c]
0.032	25	0	—	—[c]
0.040	25	0	—	—[c]

[a]p: predictor step length, P: # predictor steps, C: # corrector steps.
[b]RGF w.i.C.: with implied Corrector.
[c]RGF stops anywhere on the valley pathway before the SP.

Newton step to the next stationary point, in comparison to the predictor step length (here $0.6 \times p$), to avoid the fact that the method jumps over the stationary point. The steps of the original

RGF not so near its "true" RGF curve seem to include a more accidental calculation of a corresponding Newton step to a hypothetical next stationary point. A reason for that behavior may be the flatness of some regions of the pathway to the SP of the PES of HCP.

The test shows the expected result that the new version with the implied corrector step gives the better results, throughout, in every case of the predictor step length.

PES of H_2CO

Points of the PES of H_2CO are calculated by the method RHF/STO-3G using the GAMESS_UK program.[24] The starting point is again the global minimum **M1** (see ref. 2 Table 3), and the search direction is the symmetric diminution of the two symmetric angles of the H atoms. This direction leads into a region where the PES is geometrically difficult. The test calculations of RGF are done with five different predictor step lengths, p, and also five different thresholds, ε, for the predictor/corrector decision. The results are given in Table 2. The pathway we searched for is an RGF curve to the saddle point **T1** of index 3. The pathway was found to be quite unstable in our previous tests of the first version of RGF. The curve has to cross two bifurcation points as well as one turning point, and along the way the modifications of the coordinates r_{CO} and $r_{CH_1} = r_{CH_2}$ often change their direction. Figure 3 shows the calculated pathway in a 3D system of symmetry coordinates. The curvatures orthogonal to the pathway also change greatly. The RGF with implied corrector finds the pathway in all tested cases, and the method also correctly reports the crossing of the two bifurcation points as well as the turning point. Table 2 shows again that a sufficiently large threshold, ε, enables us to totally avoid separate corrector steps. The original RGF only finds the point **T1** in accidental cases, but in these cases the bifurcation points are ordinarily reported.

To compare the VRI points of Figure 3 with the those of Figures 5 and 6 of ref. 14 we have $\alpha = 180° - 2\beta$, and r_2 in au corresponds to R in Å, where r_1 in au corresponds to r_3 in Å. The fixed planes of Figures 5 and 6 of ref. 14 are between $r_1 = 2.30$ au and $r_1 = 2.76$ au. The calculated RGF path of Figure 3 traverses this region. The first bullet of a VRI point of Figure 3 is lying at a plane analogous to the plane in Figure 5 of ref. 14, but the r_1 level, 2.315, is about that of Figure 5. This VRI point is at $R = 1.11$ Å and $\beta = 52.5°$, obviously at a VRI_1-manifold there. The second VRI point bullet of Figure 3 is at the plane $r_1 = 2.315$ au, thus at a plane nearer to Figure 6 of ref. 14. The VRI point is at $R = 1.09$ Å and $\beta = 48.0°$ at the corresponding VRI_2-manifold there. Note that the quantum mechanical levels are different in this article and in ref. 14.

Isomerization of Butane, C_4H_{10}

Points of the PES of C_4H_{10} are calculated by the method RHF/6-31G* using the GAMESS_UK program.[24] The starting point is the global minimum m_1 (see ref. 16 Fig. 2), and the search direction is the turning of one half of the molecule around the central bond. The test calculation of RGF is done with our predictor step lengths for medium molecules, $p = 0.5$ au, chosen intuitively and proven by experience,[2] and the usual RGF threshold, $\varepsilon = 0.005$, for the

Table 2. RGF Tests for the Saddle Point Search[a] (for the Point T1) on the PES of H_2CO.

$p = 0.08$[b]	RGF w.i.C.[c]		Old RGF	
Threshold	P	C	P	C
0.015	66	0	—	—
0.030	66	0	—	—
0.045	66	0	—	—
0.060	66	0	—	—
0.075	66	0	—	—

$p = 0.16$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.015	33	0	—	—
0.030	33	0	33	5
0.045	33	0	—	—
0.060	33	0	—	—
0.075	33	0	—	—

$p = 0.24$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.015	23	4	—	—
0.030	23	0	—	—
0.045	23	0	—	—
0.060	23	0	—	—
0.075	23	0	—	—

$p = 0.32$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.015	17	5	17	6
0.030	17	0	—	—
0.045	17	0	—	—
0.060	17	0	—	—
0.075	17	0	—	—

$p = 0.40$	RGF w.i.C.		Old RGF	
Threshold	P	C	P	C
0.015	14	4	14	7
0.030	14	1	—	—
0.045	14	0	—	—
0.060	14	0	—	—
0.075	14	0	—	—

[a] M1: $r_{CO} = 1.22$ Å, $r_{CH_1} = r_{CH_2} = 1.10$ Å, $\alpha_{OCH_1} = \alpha_{OCH_2} = 122.7°$, and $\theta = 180°$; T1: $r_{CO} = 1.77$ Å, $r_{CH_1} = r_{CH_2} = 1.09$ Å, $\alpha_{OCH_1} = \alpha_{OCH_2} = 65.6°$, and $\theta = 180°$.
[b] p: predictor step length. P: # predictor steps, C: # corrector steps.
[c] RGF w.i.C.: with implied Corrector.

predictor/corrector decision. RGF results in the isomerization pathway to the saddle point s_2 and explores the pathway by 15 steps (14 predictor steps and one corrector), as well as exploring the downward pathway to the second minimum m_2 by 12 steps. (In comparison to ref. 16 where between 60 and 100 steps are needed to go uphill, or 111 steps to go back downhill to m_1 and 51 steps down to m_2.) The difference in energy between the m_1 and the saddle point s_2 is 0.005 8 H, and the difference in energy between the staggered antiminimum m_2 and the gauche m_1 is 0.001 5 H or 0.94 kcal/mol.

Ring Opening of Sym-Tetrazine

Points of the PES of $H_2C_2N_4$ are calculated by the 3-21G basis using the GAMESS_UK program.[24] The starting point is the global minimum (cf. ref. 25 Scheme 10 for a more adapted initial point), and the natural search direction is the stretching of three "opposite" bonds of the ring. If the internal coordinates start with the ring distances, the search vector is $\mathbf{r} = 1/\sqrt{3}\{1, 0, 1, 0, 1, 0$ and $12 \times 0\}$. The test calculation of RGF is done with the predictor step lengths[2] for medium molecules, $p = 0.5$ au, and the threshold, $\varepsilon = 0.01$, for the predictor/corrector decision. RGF results in the pathway to the saddle point of the concerted unimolecular triple dissociation and explores the pathway by 11 steps (10 predictors and one corrector). (The comparison with ref. 25 shows
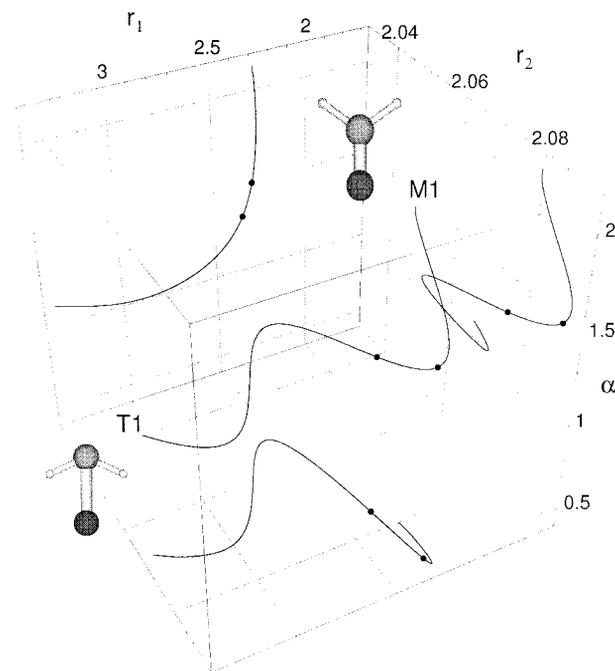
Figure 3. RGF curve of the symmetric lowering of the two H angles of H_2CO. Coordinates are given in atomic units. The pathway leads from the global minimum **M1** to the SP of index 3 **T1** and passes two valley-ridge inflection points depicted by bullets. The *ab initio* method is RHF with basis STO-3G. The coordinates are the symmetry coordinates: distances $r_1 = r_{CO}$ and $r_2 = r_{CH_1} = r_{CH_2}$ and the angle $\alpha = \angle H1CO = \angle H2CO$. Three projections of the pathway into sectional planes are also shown.
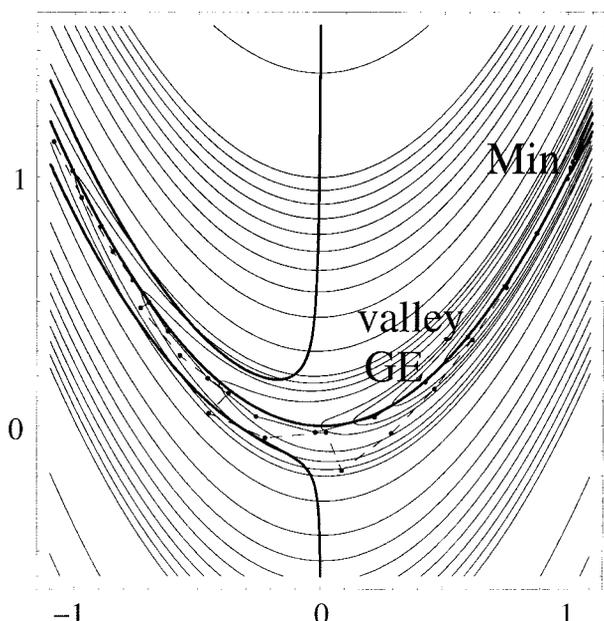
Figure 4. Rosenbrock surface[27] of eq. (11) where the boldfaced curves are the gradient extremals. The original TASC version (dashed lines) of the usual predictor + corrector strategy is shown to approximate the valley floor gradient extremal, as well as the improved TASC version (solid lines) of the implied 1/3-corrector step τ [see text, eqs. (7) and (9)]. The valley uphill is perfectly traced.

14 steps that are needed to find the SP, or compare ref. 26 which uses 30 steps.)

Test Surface 2

The second example is the 2D polynomial surface of the Rosenbrock function[27] (see Fig. 4):

$$E(x, y) = 100(y - x^2)^2 + (x - 1)^2 \tag{11}$$

The level lines are drawn at {0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300}. The solution of $E_y = 0$ (RGF curve) would be a good minimum path trajectory that follows the parabola $y = x^2$ around the "corner" at the y-axis. However, the parabola is not exactly the minimum pathway. The GE to the smallest eigenvalue follows the streambed of the surface very well, and TASC follows the GE. (The GEs are drawn in boldface.) The global minimum is $E(1, 1) = 0$, and the highest equilevel line in Figure 4 is level 300, but the value of E at point $(-1, 1)$ near the minimum pathway is only 4. The polygonal pathways of the two different TASC methods are shown. Both of them successfully progress along the shallow valley uphill, by-passing the corner. The original TASC method is depicted by a dashed line, where the new one with implied corrector eq. (9) is shown with a solid line. Using the fixed predictor step length $p = 0.25$ and the very large threshold 12.5, both methods need 13 predictor steps, where the correctors need five steps for the original method, and two steps for the new one.

Test Surface 2 with n = 4

This example is the Rosenbrock function for dimension, $n = 4$. We test TASC with the higher dimensional, extended Rosenbrock function in a version where all dimensions are coupled:

$$E(\mathbf{x}) = \sum_{i=1}^{n-1} (100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2) \tag{12}$$

As in the 2D case, the minimum is the point $\mathbf{x}_M = (1, 1, 1, 1)$. An approximation of the minimum pathway is the "super-parabola":

$$x_i = t^{2^{i-1}}, \quad i = 1, \ldots, 4, \quad t \in [-1, 1]$$

with parameter t. However, the "true" minimum energy path is again the GE. It leads from the global minimum with start direction $(-0.12, -0.23, -0.44, -0.86)$ to the saddle point at $\mathbf{x}_{SP} = (-0.656, 0.443, 0.204, 0.042)$ along the eigenvector to the smallest eigenvalue. Table 3 gives the results of tests using the original and the new TASC method starting at the global minimum. We used two versions of the predictor step length: $p = 0.1$, and $p = 0.25$, and the predictor/corrector threshold, ε, varies over five powers of ten, from 0.0005 to 100. Using the small values, we can be sure to explore the GE very exactly. For higher values we follow the valley more or less along the outer side of the arc, going across the cirque. Because the Rosenbrock function has such a deep valley, it seems to nearly be independent of the chosen threshold to follow the valley successfully: the new TASC with the implied corrector (9) in every case finds the SP. If the method is not enforced to calculate the valley ground exactly, then TASC can go uphill to the SP without any explicit corrector step. The former TASC method works well for the threshold up to $\varepsilon = 10.0$ in the small step length case ($p = 0.1$), and to $\varepsilon = 0.05$ in the larger step length case ($p = 0.25$). However, it needs a higher number of separate corrector steps, as expected. Behind the given values of ε, the method also follows uphill the deep Rosenbrock valley; however, the very shallow SP is failed and passed by without notice. The reason is again the stopping criterion: we use as the stopping criterion the smallness of a Newton step to the next stationary point, in comparison to the predictor step length. However, this notorious function is so flat along the valley ground that the stopping criterion has to be chosen two orders smaller than the predictor step length. (It is crit_stop = 0.025. Usually, we recommend that crit_stop should be larger than $0.5 \times p$.) The larger deviation of a predictor step of the original TASC from the "true" GE curve leads to the bypassing of the SP. For the largest tested $\varepsilon = 100$ in the step length case $p = 0.25$, the original TASC method finally goes totally wrong.

For the case of the 4D Rosenbrock function we give results of tests of Bofills update[20] in Table 3. It seems that the updated Hessian also gives pretty good information for this notoriously complicated function.

Table 3. TASC Tests for Saddle Point Search[a] on the 4D Rosenbrock Model Surface.[27]

$p = 0.1$	T wiC[b]		T wiC + U[c]		Old TASC		Old T + U	
Threshold	P	C	P	C	P	C	P	C
0.0005	32	39	30	105	30	59	27	104
0.005	32	31	30	83	30	45	28	83
0.05	32	24	29	62	30	31	27	64
0.5	32	6	30	33	31	30	29	40
1.0	32	1	26	22	31	24	27	33
5.0	32	0	33	17	31	7	32	23
10.0	32	0	—	—	33	4	—	—
50.0	32	0	—	—	(34	0)[d]	—	—

$p = 0.25$	T wiC		T wiC + U		Old TASC		Old T + U	
Threshold	P	C	P	C	P	C	P	C
0.0005	14	22	11	57	13	31	9	66
0.005	14	19	11	50	13	27	—	—
0.05	14	14	14	43	13	23	—	—
0.5	14	10	16	48	(14	16)[d]	—	—
1.0	14	7	16	39	(14	13)[d]	10	35
5.0	14	1	—	—	(14	11)[d]	—	—
10.0	14	0	—	—	(14	7)[d]	—	—
50.0	14	0	—	—	(15	2)[d]	—	—
100.0	14	0	—	—	TASC	fails	—	—

[a]Minimum at $\{1, 1, 1, 1\}$, saddle point at $\{-0.66, 0.44, 0.20, 0.04\}$. p: predictor step length, P: # predictor steps, C: # corrector steps.
[b]T wiC: with 1/3 implied corrector, see text.
[c]+U: the Hessian is updated by Bofills update.[20]
[d]TASC follows the valley pathway up to the SP but fails to locate the SP.

Discussion

RGF[2,3]/TASC[17] methods are effective tools to locate SPs of a PES, and additionally, TASC detects the valley floor line of a chemical reaction. The original RGF/TASC methods follow a curvilinear leading line by tangential steps, but those steps deviate more or less from their leading curve and corrector steps have to be done to satisfy the threshold (8). Of course, the ratio of predictor and corrector steps depends on the curvilinearity of the leading line and on the parameters of the method: the predictor step length, p, and the corrector/predictor threshold, ε. The predictor strategy proposed here follows the leading curve more smoothly and often reduces the total number of steps required. In favorable cases it is possible to explore the pathway from minimum to SP without any additional corrector step! This allows us to choose a large threshold, ε, for the corrector step decision, and, in addition, leads to more stability when following the path. It costs nearly nothing to additionally solve the system (7) at every predictor step, in comparison to a new calculation of gradient and Hessian of the PES for a separate corrector step, $\mathbf{c}$.

Acknowledgments

Appendix: Scheme of the Algorithms

RGF—Reduced Gradient Following

1. Initialization, preparation, and execution of first step along given direction.
   Calculate constant projector $\mathbf{P_r}$ and set $k: = 0$.
2. Set $k: = k + 1$.
   Transform internal to Cartesian coordinates.
   Calculate gradient and Hessian matrix or update of the Hessian.
   Mass weighting.
   Calculate metric tensor $\mathbf{g}$ and coordinate transformation B.
   Transform gradient and Hessian into internal coordinates: $\mathbf{g}_k$ and $\mathbf{H}_k$.
   Calculate unit vector of the new tangent by decomposition of $\mathbf{P_r H}_k\mathbf{t}_k = 0$.
3. Test of tolerance.
   If $|\mathbf{P_r g}_k| < \varepsilon$ go to step 4 to predictor.
   Or else do corrector step $\mathbf{c}$ orthogonally to tangent
   $$\mathbf{P_r H}_k\mathbf{c} = -\mathbf{P_r g}_k$$
   $$\mathbf{t}_k^T\mathbf{c} = 0$$

   Execution: $\mathbf{x}_{k+1} =: \mathbf{x}_k + \mathbf{c}$ and go to step 5.

4. Predictor step.

     Original RGF   RGF with implied corrector

                 Solve with $\mathbf{t}_k$: $\mathbf{P_r}\mathbf{H}_k\tau_k = -\mathbf{P_r}\mathbf{g}_k$

$$\mathbf{t}_k^T\tau_k = p$$

     Execution: Execution:

     $\mathbf{x}_{k+1} =: \mathbf{x}_k + p\mathbf{t}_k$ $\mathbf{x}_{k+1} =: \mathbf{x}_k + \tau_k$

5. Store actual values.

Repeat steps 2 to 4 until STOP criterion is satisfied: Newton-Raphson step $< 0.6 \ p$.

TASC—TAngent Search Concept

1. Initialization, $k := 0$, preparation, and execution of first step along given direction by start projector $\mathbf{P_{t0}}$ with $\mathbf{t_0} =: \mathbf{r}$ or lowest eigenvector.

2. Set $k =: k + 1$.

Transform internal to Cartesian coordinates.

Calculate energy, gradient, and Hessian matrix or update of the Hessian.

Mass weighting.

Calculate metric tensor $\mathbf{g}$ and coordinate transformation B.

Transform gradient and Hessian into internal coordinates: $\mathbf{g}_k$ and $\mathbf{H}_k$.

Vibrational analysis.

Calculate unit vector of the new tangent by decomposition of $\mathbf{P_r}\mathbf{H}_k\mathbf{t}_k = 0$.

3. Test of tolerance.

If $|\mathbf{P_r}\mathbf{g}_k| < \varepsilon$ go to step 4 to predictor.

Or else do corrector step $\mathbf{c}$ orthogonally to tangent

$$\mathbf{P_r}\mathbf{H}_k\mathbf{c} = -\mathbf{P_r}\mathbf{g}_k$$
$$\mathbf{t}_k^T\mathbf{c} = 0$$

Execution: $\mathbf{x}_{k+1} =: \mathbf{x}_k + \mathbf{c}$ and go to step 5.

4. Predictor step.

     Original TASC   TASC with implied corrector

                 Solve with $\mathbf{t}_k$: $\mathbf{P_r}\mathbf{H}_k\tau_k = -\mathbf{P_r}\mathbf{g}_k$

$$\mathbf{t}_k^T\tau_k = p$$

     Execution: Execution:

     $\mathbf{x}_{k+1} =: \mathbf{x}_k + p\mathbf{t}_k$ $\mathbf{x}_{k+1} =: \mathbf{x}_k + \dfrac{1}{3}\left(\tau_k + 2p\mathbf{t}_k\right)$

Set $\mathbf{r} =: \mathbf{t}_k$ and calculate next projector $\mathbf{P_r} = \mathbf{P}_{\mathbf{t}k}$.

5. Store actual values.

Repeat steps 2 to 4 until STOP criterion is satisfied: Newton-Raphson step $< 0.6 \ p$.

References

1. (a) Rothmann, M. J.; Lohr Jr., L. L. Chem Phys Lett 1980, 70, 405; (b) Scharfenberg, P. J Comput Chem 1982, 3, 277; (c) Williams, I. H.; Maggiora, G. M. J Mol Struct (Theochem) 1982, 89, 365.
2. Quapp, W.; Hirsch, M.; Imig, O.; Heidrich, D. J Comput Chem 1998, 19, 1087.
3. Quapp, W.; Hirsch, M.; Heidrich, D. Theoret Chem Acc 1998, 100, 285.
4. Hirsch, M.; Quapp, W.; Heidrich, D. Phys Chem Chem Phys 1999, 1, 5291.
5. Branin, F. H. IBM J Res Develop 1972, 504.
6. Diener, I.; Schaback, R. J Optimiz Theory Appl 1990, 67, 87.
7. Schiele, K.; Hemmecke, R. Z Angew Math Mech 2001, 81, 291.
8. Anglada, J. M.; Besalu, E.; Bofill, J. M.; Crehuet, R. J Comput Chem 2001, 22, 387.
9. Quapp, W. J Comput Chem 2001, 22, 537.
10. Bofill, J. M.; Anglada, J. M. Theoret Chem Acc 2001, 105, 463.
11. Dallos, M.; Lischka, H.; Ventura do Monte, E.; Hirsch, M.; Quapp, W. J Comput Chem 2002, 23, 576.
12. Quapp, W.; Heidrich, D. J Mol Struct (Theochem), in press.
13. Crehuet, R.; Bofill, J. M.; Anglada, J. M. Theoret Chem Acc 2002, 107, 130.
14. Quapp, W.; Melnikov, V. Phys Chem Chem Phys 2001, 3, 2735.
15. Allgower, E. L.; Georg, K. Numerical Continuation Methods—An Introduction; Springer: Berlin, 1990.
16. Kliesch, W. J Comput Chem 2001, 22, 1801.
17. Quapp, W.; Hirsch, M.; Heidrich, D. Theoret Chem Acc 2000, 105, 145.
18. Quapp, W. Comput Math Appl 2001, 41, 407.
19. Quapp, W. Appl Math Optim, submitted.
20. Bofill, J. M. J Comput Chem 1994, 15, 1.
21. Lauvergnat, D.; Nauts, A.; Justum, Y.; Chapuisat, X. J Chem Phys 2001, 114, 6592.
22. Heidrich, D.; Kliesch, W.; Quapp, W. Properties of Chemically Interesting Potential Energy Surfaces; Springer: Berlin, 1991, p. 48.
23. (a) Lami, A.; Villani, G. Chem Phys Lett 1995, 238, 137; (b) Lami, A.; Villani, G. J Mol Struct (THEOCHEM) 1995, 330, 307.
24. Guest, M. F.; Fantucci, P.; Harrison, R. J.; Kendric, J.; van Lenthe, J. H.; Schoeffel, K.; Sherwood, P. GAMESS-UK program; CFS Ltd.: Daresbury Lab., 1993.
25. Baker, J.; Chan, F. J Comput Chem 1996, 17, 888.
26. Bofill, J. M. Chem Phys Lett 1996, 260, 359.
27. Rosenbrock, H. H. Comput J 1960, 3, 175.