An International Journal
computers & mathematics with applications

PERGAMON Computers and Mathematics with Applications 41 (2001) 407–414
www.elsevier.nl/locate/camwa

Searching Minima of an N-Dimensional Surface: A Robust Valley Following Method

W. QUAPP
Mathematical Institute
University Leipzig, Augustus-Platz
D-04109 Leipzig, Germany
quapp@server1.rz.uni-leipzig.de
http://hpleo.leopoldina.uni-leipzig.de/~quapp/index.html

(Received April 2000; accepted May 2000)

Abstract—A procedure is proposed to follow the "minimum path" of a hypersurface starting anywhere in the catchment region of the corresponding minimum. The method uses a modification of the so-called "following the reduced gradient" [1]. The original method connects points where the gradient has a constant direction. In the present letter, this is replaced by the successive directions of the tangent of the searched curve. The resulting pathway is that valley floor gradient extremal which belongs to the smallest (absolute) eigenvalue of the Hessian. The new method avoids third derivatives of the objective function. The effectiveness of the algorithm is demonstrated by using a polynomial test, the notorious Rosenbrock function in two, 20, and in 100 dimensions.

Keywords—Minima, Path following, Saddle point, Reduced gradient, Gradient extremal.

1. INTRODUCTION

The idea of this note comes from applications of path following procedures in theoretical chemistry [1–3]. There, the concept of the minimum energy path of a potential energy surface is the usual approach to theoretical kinetics of chemical systems. The search for valley pathways is an important part of the analysis. This search is not equivalent to the finding of steepest descent pathways, which is the main concept in chemistry [4,5], if not the well-known zigzagging emerges, see, for example, [6]. The usual curves of RGF (see below) can be used only in certain cases for the minimum path [1,2]. The gradient extremal [7–13] appeared to represent a suitable ansatz for a minimum path, but, with its many additional solution curves and turning points [13,14], this concept in its general form is not suited to be used as a routine program for the calculation of such paths. Additionally, the calculation of the gradient extremal needs third derivatives of the cost function.

In this letter, the combination of the gradient extremal concept with the "reduced gradient following" (RGF) [1,2] opens a manageable way to follow the streambed of the surface, $f(\mathbf{x})$. We

The work was made possible through the financial support of the Deutsche Forschungsgemeinschaft. I warmly thank M. Hirsch and D. Heidrich for many interesting discussions.

0898-1221/01/$ - see front matter Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX
PII: S0898-1221(00)00283-2

understand the term streambed [15] as the valley-floor gradient extremal of the surface following the direction of the eigenvector to the smallest (absolute) eigenvalue. This gradient extremal leaves the minimum with the gentlest ascent. RGF finds a curve where the selected gradient direction comes out at every curve point, $\mathbf{x} = \mathbf{x}(t)$:

$$\frac{\nabla f(\mathbf{x}(t))}{\|\nabla f(\mathbf{x}(t))\|} = \mathbf{r}, \tag{1}$$

where t is the curve parameter, and $\mathbf{r}$ is the unit vector of the fixed search direction [2]. $\|\mathbf{x}\|$ is the Euclid l_2-norm of vector $\mathbf{x}$. The RGF method needs gradient and (updates of) the Hessian of the objective function. There are curves, which pass all stationary points in most cases. Thus, RGF is an interesting procedure in order to determine all types of stationary points [1].

The idea of this letter is to modify the RGF method to intrinsically search the minimum path. We replace the constant search direction $\mathbf{r}$ in equation (1) of the RGF method by a variable direction. We take the tangent of the searched curve itself as new direction. This is iteratively realizable by predictor and corrector steps of the RGF method. Every corrector step is calculated with the tangent direction of the previous predictor. This quickly leads to self-consistency on the valley floor gradient extremal. We term the method the TAngent Search Concept: *TASC*. Practically, equation (1) is realized by a projector ansatz [2], and the curve following requires the derivation of the ansatz [16]. However, we do not derivate the projector of the "reduced gradient" to calculate the tangent of the next predictor. With this trick, we avoid the irritating third derivatives of the cost function which occur in the terms of current gradient extremal calculations [12–14]. The error due to this trick is compensated by self-consistent iterations. Usually the smallest eigenvalue of the Hessian belongs to that eigenvector which describes the streambed direction of the surface. A counterexample is the region of a Don Quixote type saddle point [17] (a saddle point with a large curvature along the valley pathway and at least one smaller curvature across the path, as one might use on an emaciated horse). While TASC is strictly limited to follow the direction of the smallest eigenvalue, such a Don Quixote region is to leave by steepest descent.

With the tangent search concept in connection with steepest descent, we propose a practicable algorithm for searching minima of complicated, rugged surfaces. The search is restricted to the catchment region of any minimum. This region is defined as the collection of all points of the $\mathbf{R}^n$ from where steepest descent paths lead (theoretically) to the minimum. Thus, no global search is the aim of this note. Given an initial point, we start downward with steepest descent. At the emergence of zigzagging we switch to TASC which leads along the streambed to the minimum. If it is necessary, there may be additionally done some Newton-Raphson steps, to exactly calculate the minimum. The method is implemented as a subroutine in our research code of the quantum chemical GAMESS-UK program [18], or as a separate FORTRAN shell. We would like to distribute it on request.

The letter is organized as follows: Section 2 shortly repeats the mathematical fundamentals of the RGF method [1–3], and defines the modified RGF by the iterative method of the "tangent search". Subsequently, the success is demonstrated by the example of the Rosenbrock function. We finally conclude in Section 4.

2. MODIFICATION OF RGF BY FOLLOWING THE TANGENT OF THE PREVIOUS PREDICTOR STEP

2.1. TASC

We consider the solution of the unconstrained minimization problem

$$\text{minimize}_{\mathbf{x}\in\mathbf{R}^n} f(\mathbf{x}). \tag{2}$$

The objective function $f(\mathbf{x}) : \mathbf{R}^n \to \mathbf{R}^1$ is assumed to be twice-continuously differentiable, and $\mathbf{g}(\mathbf{x})$ is its gradient vector. To appreciate the gist of the method employed here, it is worthwhile to briefly review general aspects of following the reduced gradient [2]. To realize the requirement (1) on a curve $\mathbf{x}(t)$, the RGF algorithm uses a projection of the gradient to fulfill the n equations or rank $n-1$

$$\mathbf{P_r g}(\mathbf{x}(t)) = 0. \tag{3}$$

This results in the zero vector of the *reduced gradient.* The projector, $\mathbf{P_r}$, was chosen to be a constant $n \times n$-matrix: that one which enforces the gradient to point at every curve point, $\mathbf{x}(t)$, into the same direction $\mathbf{r}$. The tangent to a curve (3), $\mathbf{x}'(t)$, is obtained by a solution of the system of equations:

$$\frac{d}{dt}\left[\mathbf{P_r g}(\mathbf{x}(t))\right] = \mathbf{P_r}\frac{d\,\mathbf{g}(\mathbf{x}(t))}{dt} = \mathbf{P_r H}(\mathbf{x}(t))\mathbf{x}'(t) = 0, \tag{4}$$

where $\mathbf{H}$ is the Hessian matrix of the objective function. The predictor-corrector method of RGF is the predictor step along the tangent $\mathbf{x}'(t)$, and Newton-Raphson steps of the corrector to search, orthogonally to this direction, a solution of curve (3) [16]. The simplicity of RGF is based on the constance of the $\mathbf{P_r}$ matrix. Now, we change the projector after the predictor step: the tangent direction of the previous curve point iteratively becomes the search direction for the next point being the result of corrector steps. The procedure is named the TAngent Search Concept (*TASC*). (The task is: find the minimum path! and the motto: don't run, float [19].) But all calculations of the predictor-corrector method were done by equations (3) and (4). In the derivation of (4) we assume a "constant" $\mathbf{P}_{\mathbf{x}'(t)}$ matrix in the current step.

2.2. Example

Figure 1 illustrates the action of TASC. We use the notorious Rosenbrock function in two variables [20] in a global view at the deep parabolic valley, along $y = x^2$,

$$f(x,y) = 100\left(y - x^2\right)^2 + (x-1)^2. \tag{5}$$

The minimum of f at $(1,1)^\top$ is zero, and the highest equilevel line in Figure 1 is level 100, but the value of f at $(-1,1)^\top$ near the minimum pathway is only 4. Along the parabola, we have $f(x,x^2) = (x-1)^2$. Note, the parabola is not exactly the minimum pathway! However, it is the reduced gradient curve $f_y = 0$. To give the general case, we start at the "other side" of the central ridge: the starting point at $(-1, 0.733)^\top$ is chosen beside the minimum path. The first step is steepest descent, which in the Rosenbrock mountains quickly end at the valley ground (far away from minimum). For TASC we take a simple polynomial line search for the steplength of the predictor along $\mathbf{x}'$ [21,22], but a constant threshold (of 0.25) for the corrector steps, cf. [1]. Beginning with an *ad hoc* value of the initial steplength of 0.25, we find the following automatically adapted predictor steplengths: (0.22, 0.278, 0.309, 0.343, 0.180, 0.157, 0.159, 0.166, 0.172, 0.173, 0.167, 0.152, 0.125, 0.084, and 0.031) for 15 predictor steps. We observe one corrector step before the next predictor is done. Because the valley is curvilinear, the predictor steps are somewhat skew to the valley line, and at least one corrector step is necessary to find back into the bottom region. Of course, if there is an interest at the exact bottom line, the threshold of the corrector has to be as small as necessary. The convergence of the solution to the optimum at $(1,1)^\top$ is readily apparent. The polygonal path (whose vertices are corrector points) successfully progresses along the shallow valley to the true minimum bypassing the corner. The eigenvalues of the Hessian at the minimum are $\lambda_1 = 0.4$ and $\lambda_2 = 1001.6$, thus, the condition number $\lambda_{\max}/\lambda_{\min}$ is 2.5×10^3. The contour lines, whose axes lengths are proportional to the inverse of the eigenvalues, are thus quite elongated. For a valley with bent valley floor, the precise determination of the minimum point has been rather difficult up to now. Most optimization methods converge very slowly.

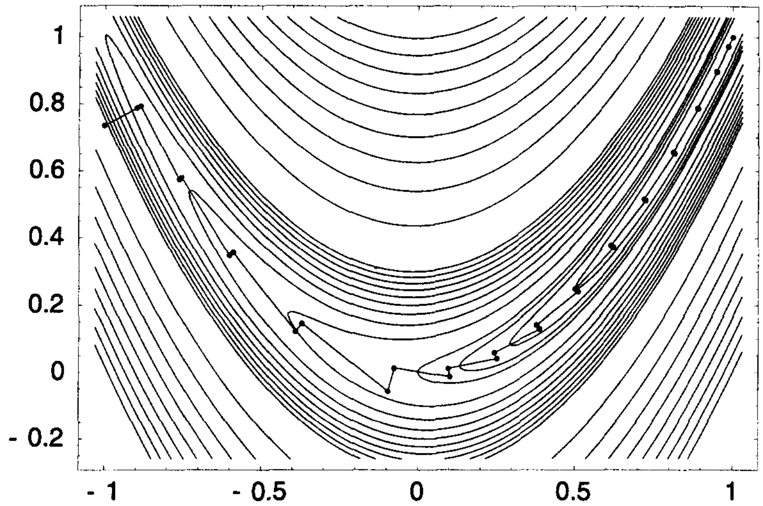

Figure 1. Convergence of TASC on the Rosenbrock surface [20] $f(x,y) = 100(y - x^2)^2 + (x-1)^2$. Contour lines are at $0.25, 0.5, 0.75, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90$, and 100. Starting at $(-1, 0.733)$, there are 15 predictor steps to find the minimum at $(1,1)^\top$.

Adapted methods give the following results [21,22]: (#fg is the number of function and gradient evaluations). Truncated Newton method: 22 iterations, #fg: 27; nonlinear conjugate gradient: 14 iterations, #fg: 31; and Quasi-Newton, full-memory BFGS: 40 iterations, #fg: 47.

2.3. The Action of TASC

In general, the resulting curve of TASC is the valley floor. After the initial steepest descent, and after a small number of steps, the method follows the streambed of the valley along the direction of the eigenvector with the smallest (absolute) eigenvalue. In the next subsection we show that the gradient extremal equation actually is the background of TASC, and the resulting curve is a numeric approximation of the valley floor gradient extremal. The reason of this behavior is an intrinsic action of the RGF method. This can be explained using the equivalent differential equation of Branin [23], which has the same solution curve as the RGF method [2]. It is

$$\frac{d\mathbf{x}}{dt} = \mathbf{x}'(t) = \pm\mathbf{A}\mathbf{g}(\mathbf{x}(t)), \tag{6}$$

where $\mathbf{x}'(t)$ is the tangent to the solution curve of the RGF ansatz, and $\mathbf{A}$ is the adjoint matrix of the Hessian. $\mathbf{A}$ is defined by $\mathbf{AH} = \mathrm{Det}(\mathbf{H})\mathbf{I}$ with the unit matrix $\mathbf{I}$. (The sign is chosen to enforce $\mathbf{Ag}$ to search for stationary points of even or odd index.) If $\mathbf{e}_1, \dots, \mathbf{e}_n$ are the eigenvectors of $\mathbf{H}$ with eigenvalues $\lambda_1, \dots, \lambda_n$, then they are also the eigenvectors of $\mathbf{A}$ but with the eigenvalues $\mu_i = \prod_{k\neq i} \lambda_k$. This is due to the equation

$$\mathbf{H}\mathbf{e}_i = \lambda_i \mathbf{e}_i, \tag{7}$$

and, by multiplication with the adjoint matrix, we get

$$\mathbf{AH}\mathbf{e}_i = \mathrm{Det}(\mathbf{H})\mathbf{e}_i = \lambda_i \mathbf{A}\mathbf{e}_i \tag{8}$$

with

$$\mathrm{Det}(\mathbf{H}) = \prod_{k=1}^{n} \lambda_k. \tag{9}$$

If a point of the solution curve of the RGF method with the search direction $\mathbf{r}$ is reached, the gradient of equation (6) points in the same direction. Expressing $\mathbf{r}$ by the eigenvectors

$$\mathbf{r} = \sum_{i=1}^{n} r_i \mathbf{e}_i, \tag{10}$$

we obtain the relation for the tangent direction

$$\mathbf{x}' = \mathbf{A}\mathbf{r} = \sum_{i=1}^{n} r_i \left(\prod_{k \neq i} \lambda_k \right) \mathbf{e}_i. \tag{11}$$

Let λ_1 be the smallest (absolute) eigenvalue. The $\mathbf{e}_1$ component of the preceding search direction $\mathbf{r}$ is enforced, if in the next step the new direction $\mathbf{x}'$ (11) is used in equations (3),(4). Thus, if the search direction is the tangent of an RGF curve, this direction is now turned to the $\mathbf{e}_1$ direction. The action is larger the larger the differences of the eigenvalues $\lambda_2, \ldots, \lambda_n$ are against λ_1. The rate of convergence of an initial direction against the minimum path depends on the ratio of the extremal eigenvalues of $\mathbf{H}$, but it is also dependent on the entire matrix spectrum. Formula (11) allows an effective procedure of eigenvector following to the smallest eigenvalue λ_1, which is automatically realized by TASC. Using TASC, the diagonalization of the Hessian to calculate the eigenvectors is not necessary. In contrast, we may use the product [12,13,24]

$$\lambda_1 = \frac{\mathbf{g}^\top \mathbf{H} \mathbf{g}}{\|\mathbf{g}\|^2} \tag{12}$$

to guess the smallest eigenvalue on a TASC pathway, and the gradient is the corresponding eigenvector; cf. also formula (17) below. The aspect becomes computationally important for very large systems [25].

2.4. Proof that TASC Yields a Gradient Extremal

Instead of RGF (3) we may use the equivalent Branin method, equation (6). There is the tangent, $\mathbf{x}'$, of a solution of RGF, as well. The projector of the original RGF was $\mathbf{P}_\mathbf{r}$ to the search direction, $\mathbf{r}$. We replace this constant direction after every predictor step by the direction of the tangent, $\mathbf{x}'$. Thus, the projector for the next step must be constructed with the vector $\mathbf{Ag}$. This may be done by the dyadic product

$$\mathbf{P}_{\mathbf{Ag}} = \mathbf{I} - \frac{(\mathbf{Ag})(\mathbf{Ag})^\top}{\|\mathbf{Ag}\|^2} \tag{13}$$

(because it realizes $\mathbf{P}_{\mathbf{Ag}}(\mathbf{Ag}) = \mathbf{0}$). If we search for a solution curve of (3), this becomes

$$\mathbf{P}_{\mathbf{Ag}}\mathbf{g} = 0 = \mathbf{g} - \mathbf{Ag}\,\frac{(\mathbf{g}^\top \mathbf{Ag})}{\|\mathbf{Ag}\|^2}. \tag{14}$$

Multiplication by $\mathbf{H}$ from the left-hand side gives

$$\mathbf{Hg} = \mathrm{Det}(\mathbf{H})\mathbf{Ig}\,\frac{(\mathbf{g}^\top \mathbf{Ag})}{\|\mathbf{Ag}\|^2}. \tag{15}$$

$\mathrm{Det}(\mathbf{H})\mathbf{I}$ is commutative with all other terms and can change its place into the product $(\mathbf{g}^\top \mathbf{Ag})$, and then we can replace it back to $\mathbf{AH}$. Thus, the expression on the right-hand side, without one $\mathbf{g}$, is a scalar. If we denote it by $\lambda(\mathbf{x}(t))$, we obtain the known eigenvector equation

$$\mathbf{Hg} = \lambda \mathbf{g}, \tag{16}$$

which is the equation of the gradient extremal [9]. Note that with $\mathbf{Ag} = \mathbf{v}$ the eigenvalue becomes, cf. [13],

$$\lambda = \frac{\mathbf{v}^\top \mathbf{H} \mathbf{v}}{\|\mathbf{v}\|^2}. \tag{17}$$

3. PERFORMANCE

3.1. Rosenbrock Function, Dimension $n = 20$

We test TASC with the high dimensional, extended Rosenbrock function in a version where all dimensions are coupled [26]

$$f(\mathbf{x}) = \sum_{i=1}^{n-1} \left(100\left(x_{i+1} - x_i^2\right)^2 + (x_i - 1)^2\right). \tag{18}$$

As well as in the two-dimensional case, the minimum is the point $\mathbf{x}_{\min} = (1, \dots, 1)^\top$. A good approximation of the minimum pathway is again the "super-parabola" $x_i = t^{2^{i-1}}$, $i = 1, \dots, 20$. It leads from the global minimum over a saddle point at (rounded)

$$\mathbf{x}_{\mathrm{SP}} = (-0.555, 0.322, 0.115, 0.024, 0.011, 14 \times 0.010, 0.0001)^\top$$

to a second minimum $\mathbf{x}_{\mathrm{M2}}$ near the point $(-1, 1, \dots, 1)^\top$. The initial point of the test search reported here lies at the saddle side, but in opposition to the global minimum. A typical result is reported in Figure 2. We start near the saddle point, however not directly on the valley floor. We choose $x_1 = -0.45$, and the other coordinates are put to the values of the saddle given above. This is $\mathbf{x}_{\mathrm{start}}$. The first step is done by steepest descent down to the minimum pathway. There the method switches to TASC. The corrector threshold 1 is used. The predictor steplength of the TASC is started with a value of 0.3, and it is extended if possible by a rough steplength stretch as long as the function (18) decreases along the step. This is done in the saddle point region where the streambed profile is not convex, where, in contrast, a region of indefinite curvature is encountered (λ_1 is very small, but negative). The influence of the saddle point region extends over 22 predictor steps where the steplength could not be larger then 0.45. Then the minimum bowl is really reached, where all the eigenvalues become positive. The predictor steplength is now dynamically calculated by a polynomial line search [21,22]. The method TASC needs a totality of 30 predictor steps to reach the minimum, and 50 corrector steps are used to follow the valley ground more or less exactly. Figure 2 shows the behavior of the coordinates: the diamond symbol $\diamond$ represents the coordinates $x_{1 \bmod 5}$, the star symbol $\star$ represents the coordinates $x_{2 \bmod 5}$, the square $\square$ represents the coordinates $x_{3 \bmod 5}$, the triangle $\triangle$ represents the coordinates $x_{4 \bmod 5}$, and the bullet $\bullet$ represents $x_{0 \bmod 5}$. The steepest descent step 1 is shown, and the final convergence of the iteration steps to the minimum. The coupling of the coordinates in the objective function (18) enforces a successive search across the 20 dimensions of the problem, which can be seen by the

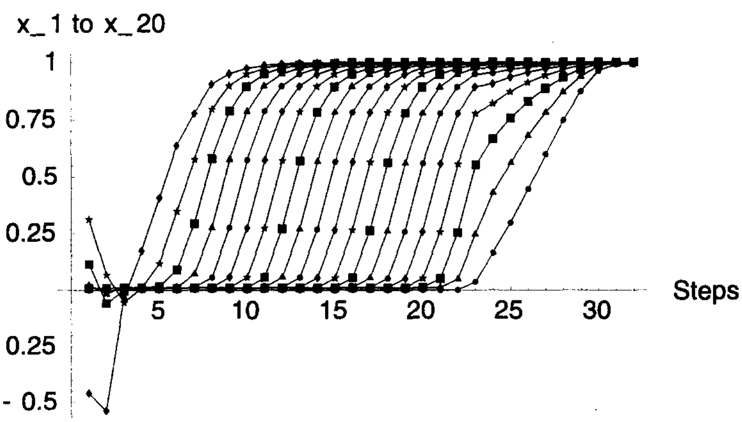

Figure 2. Convergence of TASC on the 20-dimensional Rosenbrock function in 30 iteration steps to the minimum $(1, \dots, 1)^\top$, see text.

successive convergence of the different coordinates to the final value 1. The intrinsic curvilinearity of the surfaces valley ground is the notorious property of this special test function. Near the minimum, the steplength used by the predictor is successively reduced downwards. At the minimum, the ratio of λ_1 to λ_2 is ≈ 0.049, and λ_1/λ_{20} is ≈ 0.017, thus, TASC is very well adapted to do the minimum search in such a rugged hypersurface along the direction of the smallest eigenvalue. Note, steepest descent will die of zigzagging, and Newton's method (with $\mathbf{x}_{\text{start}}$) accidentally found the second minimum $\mathbf{x}_{\text{M2}}$. An actual genetic method needs quite more steps (using only function calls) and does not reach the accuracy of TASC (in the cases $n = 3$, 11, and 21: see [26] where the special character of the coupling in the objective function (18) is explicitely used).

3.2. Rosenbrock Function, Dimension $n = 100$

The objective function (18) with $n = 100$ has an analogous saddle as in the case of 20 dimensions, at (rounded) $\mathbf{x}_{\text{SP}} = (-0.555, 0.322, 0.115, 0.024, 0.011, 94 \times 0.010, 0.0001)^{\text{T}}$.

Again, we start at $x_1 = -0.45$, and put all other coordinates as given above. We use the corrector threshold 2, and an initial steplength of 0.25. Seen over the minimum path, the (calculated) arclength down to the global minimum is quite long: 39.28 against 9.20 in the 20-dimensional case, and the slope along the minimum path is also quite flat: the function value difference between saddle point and minimum is 98.697 for $n = 100$, but for $n = 20$ it is 19.505. The curvature of the minimum path between the coordinates from i to $i+1$ is analogous to the lower-dimensional cases. The minimum path is near to the parabolic relation $x_{i+1} = x_i^2$, $i = 1, \dots, 99$, but because of the longer list of coordinates, it is a little more coupled. The steplength for predictor steps which guarantees descent is the small value of 0.423. It is tested from step to step, and it is decreased by the polynomial steplength search [21,22] in the last 10 steps only. Resulting, TASC finally finds the minimum, but it needs 84 predictor and 171 corrector steps to find through the rugged Rosenbrock mountains in this high-dimensional case.

4. CONCLUSIONS AND PERSPECTIVES

We demonstrate the workability of the TASC algorithm for following the streambed gradient extremal. We test a highly coupled problem with strong nonlinearity: the Rosenbrock function. The method performs well in practice. We start at any point in the catchment region of a minimum and follow the gradient down the slope to the "minimum path". Then we follow this path in direction of the smallest eigenvector. We use the evaluation of gradient and Hessian per iteration step, and an additional gradient for the line search of the predictor.

The procedure is a potent method for studying the streambed of a multidimensional surface which is calculable as exactly as we need it. Its success is based on the tracing of the minimum path, which we geometrically understand as the valley floor gradient extremal. The original RGF [1,2] forms an interesting tool to find minima or saddle points, where the choice of the search direction is quite arbitrary. But RGF can diverge more or less from the minimum path even if it starts in eigenvector direction. The choice of the actual tangent in TASC now overcomes the arbitrariness of the direction choice used in the RGF method.

Because the fundamental property of TASC to explore the minimum path usually does not depend on the direction—downhill or uphill—the method can also be used to search for saddle points of index one starting at a given minimum. This is reported in a forthcoming paper [27].

REFERENCES

1. W. Quapp, M. Hirsch, O. Imig and D. Heidrich, Searching for saddle points of potential energy surfaces by following a reduced gradient, *J. Computat. Chem.* **19**, 1087–1100, (1998).
2. W. Quapp, M. Hirsch and D. Heidrich, Bifurcation of reaction pathways: The set of valley-ridge-inflection points of a simple 3-dimensional potential energy surface, *Theoret. Chem. Acc.* **100**, 285–299, (1998).
3. M. Hirsch, W. Quapp and D. Heidrich, The set of valley-ridge-inflection points of the potential energy surface of the water molecule, *Phys. Chem. Chem. Phys.* **1**, 5291–5299, (1999).

4. K. Fukui, A formulation of the reaction coordinate, *J. Phys. Chem.* **74**, 4161–4163, (1974).
5. W. Quapp and D. Heidrich, Analysis of the concept of minimum energy path on the potential energy surface of chemically reacting systems, *Theoret. Chim. Acta* **66**, 245–260, (1984).
6. B.L. Garrett, M.J. Redmon, R. Steckler, D.G. Truhlar, K.K. Baldridge, D. Bartol, M.V. Schmidt and M.S. Gordon, Algorithms and accuracy requirements for computing reaction paths by the method of steepest descent, *J. Phys. Chem.* **92**, 1476–1488, (1988).
7. M.V. Basilevsky and A.G. Shamov, The local definition of the optimum ascent path on a multidimensional potential energy surface and its practical application for the location of saddle points, *Chem. Phys.* **60**, 347–358, (1981).
8. M.V. Basilevsky, The topography of potential energy surfaces, *Chem. Phys.* **60**, 337–346, (1981).
9. D.K. Hoffman, R.S. Nord and K. Ruedenberg, Gradient extremals, *Theoret. Chim. Acta* **69**, 265–280, (1986).
10. W. Quapp, Gradient extremals and valley floor bifurcation on potential energy surfaces, *Theoret. Chim. Acta* **75**, 447–460, (1989).
11. D. Heidrich, W. Kliesch and W. Quapp, Properties of chemically interesting potential energy surfaces, In *Lecture Notes in Chemistry, Vol. 56*, Springer, Berlin, (1991).
12. J.-Q. Sun and K. Ruedenberg, Gradient extremals and steepest descent lines on potential energy surfaces, *J. Chem. Phys.* **98**, 9707–9714, (1993).
13. W. Quapp, O. Imig and D. Heidrich, Gradient extremals and their relation to the minimum energy path, In *The Reaction Path in Chemistry: Current Approaches and Perspectives*, (Edited by D. Heidrich), pp. 137–160, Kluwer, Dordrecht, (1995).
14. F. Jensen, Locating transition structures by mode following: A comparison of 6 methods on the Ar_8 Lennard-Jones potential, *J. Chem. Phys.* **102**, 6706–6718, (1995).
15. K. Ruedenberg and J.-Q. Sun, Gradient fields of potential energy surfaces, *J. Chem. Phys.* **100**, 5836–5848, (1994).
16. W.C. Rheinboldt, *Numerical Analysis of Parametrized Nonlinear Equations*, Wiley, New York, (1986).
17. P.G. Mezey, *Potential Energy Hypersurfaces*, Elsevier, Amsterdam, (1987).
18. M.F. Guest, P. Fantucci, R.J. Harrison, J. Kendric, J.H. van Lenthe, K. Schoeffel and P. Sherwood, *GAMESS-UK Program*, CFS Ltd., Daresbury Lab., (1993).
19. Adidas, Sales promotion for sport shoes, (1996).
20. H.H. Rosenbrock, An automatic method for finding the greatest or least value of a function, *The Computer Journ.* **3**, 175–184, (1960); cf. *Towards Global Optimization, Part 2*, (Edited by L.C.W. Dixon and G.P. Szegö), p. 84, North-Holland, Amsterdam, (1975).
21. T. Schlick, Optimization methods in computational chemistry, In *Reviews in Computational Chemistry, Vol. III*, (Edited by K.B. Lipkowitz and D.B. Boyd), pp. 1–71, VCH Publishers, New York, (1992).
22. The Computational Science Education Project, Mathematical Optimization, `http://csep1.phy.ornl.gov/CSEP/MO/NODE11B.html`, (1995).
23. F.H. Branin, Jr., Widely convergent methods for finding multiple solutions of simultaneous nonlinear equations, *IBM J. Res. Develop.* **16**, 504–522, (1972).
24. D.J. Rowe and A. Ryman, Valleys and fall lines on a Riemannian manifold, *J. Math. Phys.* **23**, 732–735, (1982).
25. A.J. Turner, V. Moliner and I.H. Williams, Transition state structural refinement with GRACE and CHARMM: Flexible QM/MM modeling for lactate dehydrogenase, *Phys. Chem. Chem. Phys.* **1**, 1323–1331, (1999).
26. P. Bouvry, F. Arbab and F. Seredynski, Distributed evolutionary optimization, in Manifold: Rosenbrock's function case study, *Inform. Sci.* **122**, 141–159, (2000).
27. W. Quapp, M. Hirsch and D. Heidrich, Following the streambed reaction on potential energy surfaces: A new robust method, *Theoret. Chem. Acc.*, (submitted).