Theor Chem Acc (2002) 107:130–139
DOI 10.1007/s00214-001-0306-x

Theoretical
Chemistry Accounts

Regular article

A new look at the reduced-gradient-following path

Ramon Crehuet[1], **Josep Maria Bofill**[2], **Josep Maria Anglada**[1]

[1] Institut d'Investigacions Químiques i Ambientals. C.I.D.–C.S.I.C., Jordi Girona Salgado 18-26, 08034 Barcelona, Catalunya, Spain
[2] Departament de Química Orgànica and Centre Especial de Recerca en Química Teòrica, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Catalunya, Spain

Received: 21 May 2001 / Accepted: 27 September 2001 / Published online: 9 January 2002

Abstract. The mathematical structure of the reduced-gradient-following (RGF) path introduced by Quapp et al. (1988 *J. Comput. Chem.* 19:1087) is reviewed and analyzed. We report two new algorithms to evaluate the RGF path. The RGF path is also compared mathematically and computationally with the gradient extremals path. An example of the evaluation of the RGF path is also reported.

Key words: Reaction path – Reduced-gradient-following path – Gradient extremals – Transition states – Turning points

1 Introduction

From a theoretical point of view the path of a chemical reaction can be defined as a curve on the potential-energy surface (PES). This curve goes from the reactant to product minima trough a saddle point of index 1. The saddle point is nothing more than the transition structure of transition-state theory. For reviews about reaction paths see Refs. [1, 2].

The reaction path can be seen as a parameterization of the PES with respect to a parameter, the so-called reaction coordinate (RC). If $\mathbf{x}$ is a coordinate vector of dimension N, then the curve associated with the reaction path is $\mathbf{x}(s)$, where s is the RC. As a consequence of this definition, the N-dimensional PES is reduced to a one-parameter problem. Using the reaction path concept it is, in principle, possible to locate both stationary points and dynamics without the calculation of the whole PES [3, 4]. There are many different mathematical ways to define a reaction path. The most used reaction path in chemistry is the one that follows the steepest-descent curve from the first-order saddle point. If this type of curve is computed using mass-weighted coordinates then we have the so-called intrinsic reaction coordinate (IRC) [5] or minimum energy path (MEP). The tangent along the IRC or MEP is defined by an autonomous system of differential equations and its solution is unique. The IRC curve does not present bifurcations except in the stationary points. The mathematical expression of the tangent to the steepest-descent curve is

$$\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}s} = -\frac{\mathbf{g}[\mathbf{x}(s)]}{|\mathbf{g}[\mathbf{x}(s)]|} \ , \tag{1}$$

where $\mathbf{x}$ and $\mathbf{g}$ are the coordinate and gradient vectors respectively and s is the arc length of the curve or RC. Another possibility to define a path between stationary points corresponding to minima and first-order saddle points is the reduced-gradient following (RGF) recently introduced by Quapp et al. [6]. In the RGF path the gradient of each point of the path has a constant direction, which is mathematically formulated as an implicit function

$$\mathbf{r} - \frac{\mathbf{g}[\mathbf{x}(t)]}{|\mathbf{g}[\mathbf{x}(t)]|} = \mathbf{0} \ , \tag{2}$$

where the unit vector $\mathbf{r}$ is the selected constant direction and t is the parameter of the resulting curve. The RGF method can be seen as a rigorous mathematical formulation of the old "distinguished coordinate method" [7]. The RGF paths possess some important properties; the branching points of these curves are valley-ridge inflection (VRI) points of the PES [8].

So far three methods have been proposed to compute the RGF path. The first of them was given by Quapp et al. [6] taking into account that the integration of the RGF curve involves at each iteration both a predictor and a corrector step. These authors use QR decomposition to evaluate the predictor step and a Newton–Raphson procedure as a corrector step. More recently Bofill and coworkers [9, 10] have proposed two new methods. Both methods use as corrector step the Newton–Raphson algorithm, but as predictor step the IRC technique is used in one method and an eigenvector-following technique in the other.

Correspondence to: J. M. Bofill
e-mail: jmbofill@qo.ub.es

Because of the interesting properties of the RGF curve we propose another method to follow it. The article is organized as follows: first, we describe the mathematical basis of the RGF curve; second, we give the description of two new algorithms to evaluate the most general RGF curve; a theoretical comparison with the gradient extremals (GE) path [11] is also reported; finally, some examples are given and analyzed.

2 Theoretical background of the RGF path

The tangent to a curve, $\mathrm{d}\mathbf{x}(t)/\mathrm{d}t$, determines the variation quotient of the points of this curve with respect to the parameter t that characterizes the curve. Consequently the first step consists of evaluating the form of the tangent vector of the proposed curve. We differentiate Eq. (2) with respect to t by taking into account the fact that the normalized vector $\mathbf{r}$ is constant,

$$\begin{aligned}\mathbf{0} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{r} - \frac{\mathbf{g}[\mathbf{x}(t)]}{|\mathbf{g}[\mathbf{x}(t)]|}\right)\\ &= -\frac{1}{|\mathbf{g}[\mathbf{x}(t)]|}\left[\mathbf{I} - \frac{\mathbf{g}[\mathbf{x}(t)]}{|\mathbf{g}[\mathbf{x}(t)]|}\left(\frac{\mathbf{g}[\mathbf{x}(t)]}{|\mathbf{g}[\mathbf{x}(t)]|}\right)^{\mathrm{T}}\right]\\ &\quad \times \mathbf{H}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} \quad .\end{aligned} \tag{3}$$

Where $\mathbf{H}[\mathbf{x}(t)]$ is the Hessian matrix evaluated at $\mathbf{x}(t)$. For points lying on the RGF path, Eq. (2) is satisfied and Eq. (3) takes the form

$$\left(\mathbf{I} - \mathbf{r}\mathbf{r}^{\mathrm{T}}\right)\mathbf{H}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{0} \quad . \tag{4}$$

Following Quapp et al. [8] we define a rectangular matrix, $\mathbf{P}_{\mathrm{r}}$, such that the columns of this matrix form an orthonormalized basis set, orthogonal to the normalized $\mathbf{r}$ vector,

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{r} = \mathbf{0}_{N-1} \quad , \tag{5}$$

where $\mathbf{0}_{N-1}$ denotes a zeroed vector of dimension $N-1$. The requirement that the $N-1$ columns of $\mathbf{P}_{\mathrm{r}}$ form a orthonormal basis can be relaxed to a set of $N-1$ linear independent direction vectors orthogonal with respect to the unit vector $\mathbf{r}$. Multiplying on the left Eq. (4) by $\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}$ we obtain

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{0}_{N-1} \quad . \tag{6}$$

The homogeneous system of $N-1$ linear equations (Eq. 6) determines the tangent vector of the RGF curve [6, 8]. Some part of the present discussion is based on the concept of the vectors $\mathbf{H}$-conjugated. The concept of conjugated vector is important and is used in optimization theory [12]. The solution of the set of linear equations (Eq. 6) gives the first-order predictor step (see Eq. 28). The corrector step should be carried out in order to eliminate the gradient vector components in the directions that are orthonormal to the unit vector $\mathbf{r}$.

The previous discussion leads to an important consequence: the gradient vector, $\mathbf{g}[\mathbf{x}(t)]$ at any point of the RGF curve, $\mathbf{x}(t)$, can be expressed in the full set of N linear independent directions, $\mathbf{S} = [\mathbf{r}|\mathbf{P}_{\mathrm{r}}]$. The notation $[\mathbf{r}|\mathbf{P}_{\mathrm{r}}]$ means an $N \times N$ matrix whose first column contains the $\mathbf{r}$ vector and the others $N-1$ columns are contained in the $N \times (N-1)$ $\mathbf{P}_{\mathrm{r}}$ submatrix. If $\alpha(t)$ is the component of the gradient vector $\mathbf{g}[\mathbf{x}(t)]$ in the $\mathbf{r}$ direction then

$$\mathbf{g}[\mathbf{x}(t)] = [\mathbf{r}|\mathbf{P}_{\mathrm{r}}]\begin{pmatrix}\alpha(t)\\ \mathbf{0}_{N-1}\end{pmatrix} = \alpha(t)\mathbf{r} \quad . \tag{7}$$

By comparing Eq. (2) with Eq. (7) we see that $\alpha(t) = |\mathbf{g}[\mathbf{x}(t)]|$. From Eq. (7) we conclude that the energy is stationary with respect to the $\mathbf{P}_{\mathrm{r}}$ directions. In principle, the stationary condition has a minimum character; however, at the moment this is not important. Another important consequence of Eq. (7) is that the gradient vector preserves the direction along the RGF curve since the vector $\mathbf{r}$ is constant.

Let us now to write the tangent vector, $\mathrm{d}\mathbf{x}(t)/\mathrm{d}t$, as a function of the N linearly independent directions $\mathbf{S}$,

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{S}\mathbf{b} = [\mathbf{r}|\mathbf{P}_{\mathrm{r}}]\begin{pmatrix}b_1\\ \mathbf{b}_{N-1}\end{pmatrix} = \mathbf{r}b_1 + \mathbf{P}_{\mathrm{r}}\mathbf{b}_{N-1} \quad , \tag{8}$$

where the vector $\mathbf{b}_{N-1}{}^{\mathrm{T}} = (b_2,...,b_N)$. Note that the vector $\mathbf{b}$ depends on the parameter t. Substituting Eq. (8) into Eq. (6) we obtain

$$\begin{aligned}\mathbf{0}_{N-1} &= \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{S}\mathbf{b}\\ &= \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}b_1 + \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{\mathrm{r}}\mathbf{b}_{N-1} \quad .\end{aligned} \tag{9}$$

Equation (9) possesses two types of general solution. The first one can be expressed as

$$\mathbf{b}_{N-1} = -\left(\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{\mathrm{r}}\right)^{-1}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}b_1 \quad , \tag{10}$$

given the component b_1, the $\mathbf{b}_{N-1}$ vector is determined. When the $\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{\mathrm{r}}$ matrix is nonsingular, i.e., $\det(\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{r}) \neq 0$, the tangent vector in the $\mathbf{S}$ representation vectors can be expressed as

$$\mathbf{b}^{\mathrm{T}} = (b_1,\mathbf{b}_{N-1}) = (1,(-\mathbf{A}^{-1}\mathbf{f})_1,\ldots,(-\mathbf{A}^{-1}\mathbf{f})_{N-1})b_1 \quad , \tag{11}$$

where $\mathbf{A} = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{\mathrm{r}}, \mathbf{f} = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}$ and b_1 is determined by normalization. This tangent vector corresponds at first order to the predictor step of the RGF curve.

If the matrix $\mathbf{A}$ is singular it is convenient to diagonalize it. Let $\mathbf{U}$ be the unitary matrix that diagonalizes $\mathbf{A}$, then

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{\mathrm{r}} = \mathbf{A} = \mathbf{U}\mathbf{a}\mathbf{U}^{\mathrm{T}} \quad , \tag{12}$$

where $\mathbf{a}$ is a diagonal matrix. Then, Eq. (9) can be transformed into

$$-\mathbf{a}\mathbf{h} = \mathbf{e}b_1 \quad , \tag{13}$$

where $\mathbf{h} = \mathbf{U}^{\mathrm{T}}\mathbf{b}_{N-1}$ and $\mathbf{e} = \mathbf{U}^{\mathrm{T}}\mathbf{f} = \mathbf{U}^{\mathrm{T}}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}$. Equation (13) corresponds to a set of $N-1$ decoupled equations:

$$-a_i h_i = e_i b_1 \quad (i = 1,\ldots,N-1) \quad . \tag{14}$$

Because $\det(\mathbf{A}) = 0$, at least an eigenvalue a_{i} is zero and other solutions must be investigated. Let us assume that $a_j = 0$, $e_j \neq 0$, then $b_1 = 0$, and consequently the RGF tangent is normal to the normalized vector $\mathbf{r}$. The normalized tangent vector is $\mathbf{b}^{\mathrm{T}} = (0,\ \mathbf{U}_{1,j},\ldots,\ \mathbf{U}_{N-1,j})$, where $\mathbf{U}_{i,j}$ are the elements of the $\mathbf{U}$ matrix. Note that the b_1 component represents the cosine of the angle between the RGF tangent vector and the normalized vector $\mathbf{r}$.

The curvature vector of the RGF path, $\mathrm{d}^2\mathbf{x}(t)/\mathrm{d}t^2$, must be obtained by derivation of Eq. (6) with respect to the parameter t:

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\left\{\left\langle \mathbf{F}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\right\rangle \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} + \mathbf{H}[\mathbf{x}(t)]\frac{\mathrm{d}^2\mathbf{x}(t)}{\mathrm{d}t^2}\right\} = \mathbf{0}_{N-1} \ , \tag{15}$$

where $\mathbf{F}[\mathbf{x}(t)]$ is the third energy derivative tensor with respect to the $\mathbf{x}$ variables evaluated at $\mathbf{x}(t)$. The $\langle\mathbf{F}[\mathbf{x}(t)]\mathrm{d}\mathbf{x}(t)/\mathrm{d}t\rangle$ symbolism is used to indicate a square matrix that is a contracted product of a three-index array with a vector yielding a two-index array; thus,

$$\langle\mathbf{Ab}\rangle_{ij} = \sum_k \mathbf{A}_{ijk}\mathbf{b}_k \ . \tag{16}$$

The curvature vector can be written using the set of N linear independent directions $\mathbf{S}$

$$\frac{\mathrm{d}^2\mathbf{x}(t)}{\mathrm{d}t^2} = \mathbf{Sc} = [\mathbf{r}|\mathbf{P}_{\mathrm{r}}]\begin{pmatrix} c_1 \\ \mathbf{c}_{N-1}\end{pmatrix} \ . \tag{17}$$

Substituting Eqs. (8) and (17) into Eq. (15) we get

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{Sb}\rangle\mathbf{Sb} + \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{Sc} = \mathbf{0}_{N-1} \ . \tag{18}$$

When the $\mathbf{A}$ matrix is nonsingular the curvature vector can be expressed as

$$\begin{aligned}\mathbf{c}^{\mathrm{T}} &= (1,(-\mathbf{A}^{-1}\mathbf{f})_1,\ldots,(-\mathbf{A}^{-1}\mathbf{f})_{N-1})c_1 \\ &\quad + (0,(-\mathbf{A}^{-1}\mathbf{d})_1,\ldots,(-\mathbf{A}^{-1}\mathbf{d})_{N-1}) \ ,\end{aligned} \tag{19}$$

where $\mathbf{d} = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{Sb}\rangle\mathbf{Sb}$. If the tangent vector is normalized, i.e., $[\mathrm{d}\mathbf{x}(t)/\mathrm{d}t]^{\mathrm{T}}[\mathrm{d}\mathbf{x}(t)/\mathrm{d}t] = 1$, then the curvature vector is normal to the tangent vector, i.e., $[\mathrm{d}\mathbf{x}(t)/\mathrm{d}t]^{\mathrm{T}}[\mathrm{d}^2\mathbf{x}(t)/\mathrm{d}t^2] = 0$ or in the $\mathbf{S}$ basis set representation $\mathbf{c}^{\mathrm{T}}\mathbf{b} = 0$, from this condition the c_1 component is determined. It should be noted that for a strictly quadratic approximation the curvature vector is $\mathbf{c} = \mathbf{0}$ at this point. If the actual potential energy is replaced by a local quadratic approximation, then such approximation to the potential energy will yield a good approximation to the actual curvature vector only when the elements of the third-order derivatives in the direction of the tangent vector are small enough. If this is not the case, the c_1 component will be nonzero and a correction step will be important in order to follow the correct RGF path.

When the $\mathbf{A}$ matrix is singular Eq. (18) can be transformed into

$$-\mathbf{am} = \mathbf{e}c_1 + \mathbf{n} \ , \tag{20}$$

where $\mathbf{m} = \mathbf{U}^{\mathrm{T}}\mathbf{c}_{N-1}$ and $\mathbf{n} = \mathbf{U}^{\mathrm{T}}\mathbf{d}$. Equation (20) represents a set of $N-1$ decoupled equations:

$$-a_{\mathrm{i}}m_{\mathrm{i}} = e_{\mathrm{i}}c_1 + n_{\mathrm{i}} \ (i = 1,\ldots,N-1) \ . \tag{21}$$

As in the tangent case, assuming that $a_j = 0,\ e_j \neq 0$ then $c_1 = -n_j/e_j$. Because the tangent and curvature vectors are normal, $m_j = 0$.

It is interesting to consider the situation which occurs when the Hessian matrix in a point of the RGF curve possesses an eigenvector that coincides with the normalized vector $\mathbf{r}$. The vector $\mathbf{r}$ is $\mathbf{H}$-conjugate with respect to the set of $N-1$ vectors collected in the $\mathbf{P}_{\mathrm{r}}$ matrix. In this case $\mathbf{f} = \mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r} = \mathbf{0}_{N-1}$. If the matrix $\mathbf{A}$ is nonsingular, then the normalized tangent vector in the $\mathbf{S}$ basis representation is $\mathbf{b}^{\mathrm{T}} = (1, \mathbf{0}_{N-1}{}^{\mathrm{T}})$ and consequently in the geometry coordinate representation the tangent vector is $\mathrm{d}\mathbf{x}(t)/\mathrm{d}t = \mathbf{r}$, and owing to Eq. (7) the tangent vector is parallel to the gradient vector. The curvature vector is $\mathbf{c}^{\mathrm{T}} = (0,(-\mathbf{A}^{-1}\mathbf{d})_1,\ldots,(-\mathbf{A}^{-1}\mathbf{d})_{N-1})$, and is normal to the normalized vector $\mathbf{r}$. Note that $c_1 = 0$. These kinds of points are know as turning points (TP) [8, 9].

Finally we are dealing with the second kind of solution of Eq. (9). This solution is presented when $\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r} = \mathbf{0}_{N-1}$, i.e., the normalized vector $\mathbf{r}$ is $\mathbf{H}$-conjugate with respect to the $N-1$ linear independent vectors collected in the $\mathbf{P}_{\mathrm{r}}$ matrix. In order to follow a path with this condition the vector $\mathbf{r}$ and the $\mathbf{S}$ matrix are not constant. Instead, because Eq. (5) should be satisfied, the $\mathbf{r}$ vector is an eigenvector of the current $\mathbf{H}[\mathbf{x}(t)]$ matrix. The eigenvectors are a special kind of $\mathbf{H}$-conjugate vector [12]. In this case from Eq. (10) we obtain $\mathbf{b}_{N-1} = \mathbf{0}_{N-1}$ for a given b_1.

In order to understand the nature of this second type of solution, first Eqs. (5) and (6) can be compared,

$$\mathbf{r}\mu = \mathbf{H}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} \ , \tag{22}$$

since $\mathbf{P}_{\mathrm{r}}$ is a set of $N-1$ linear independent vectors and μ is a nonzero scalar. Now, substituting Eq. (8) into Eq. (22), multiplying the resulting equation on the left by $\mathbf{r}^{\mathrm{T}}$ and imposing the conjugancy condition, $\mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{P}_{\mathrm{r}} = \mathbf{0}_{N-1}{}^{\mathrm{T}}$, the following equation is obtained:

$$\mathbf{r}^{\mathrm{T}}\mathbf{r}\mu = \mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}b_1 = \mu \ . \tag{23}$$

From this equality we obtain a value for the b_1 component, $b_1 = \mu/(\mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r})$, and the tangent vector can be obtained as

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mu\frac{\mathbf{r}}{\mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}} \ , \tag{24}$$

since $\mathbf{b}_{N-1} = \mathbf{0}_{N-1}$. Substituting Eq. (24) into Eq. (23) and taking the first equality of Eq. (23), one can write

$$\mu\mathbf{r}^{\mathrm{T}}\left\{\mathbf{H}[\mathbf{x}(t)]\mathbf{r} - \mathbf{r}\mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}\right\} = 0 \ . \tag{25}$$

Finally, using the fact that the scalar $\mu \neq 0$ and the normalized vector $\mathbf{r}$ is an eigenvector of the $\mathbf{H}[\mathbf{x}(t)]$ matrix and also $\mathbf{r} \neq 0$, then

$$\mathbf{H}[\mathbf{x}(t)]\mathbf{g}[\mathbf{x}(t)] = \mathbf{g}[\mathbf{x}(t)]\frac{\{\mathbf{g}[\mathbf{x}(t)]\}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{g}[\mathbf{x}(t)]}{\{\mathbf{g}[\mathbf{x}(t)]\}^{\mathrm{T}}\mathbf{g}[\mathbf{x}(t)]} \ , \tag{26}$$

where Eq. (2) has been used. Equation (26) is the basic equation of the GE paths [11, 13, 14]. The previous deduction can be seen as a pedestrian derivation of the GE path; a more rigorous proof will be outlined later. According to this derivation any point of the GE path satisfies simultaneously both Eq. (7) and $\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r} = \mathbf{0}_{N-1}$. Owing to this fact, at each point of the GE path the **S** matrix changes. The **S** matrix is formed by a set of vectors which are conjugate with respect to the current $\mathbf{H}[\mathbf{x}(t)]$; more specifically they are eigenvectors of this matrix. This is an important difference with respect to the RGF methodology, where the **S** matrix is constant and is not forced to be conjugate with respect to the Hessian matrix. The GE path does not always connect the minimum with first-order saddle points, so they are not candidates for a reaction path. On the other hand, the RGF path goes from a minimum to another minimum through a first-order saddle point if the normalized **r** vector is properly selected.

3 RGF algorithms

3.1 Basic considerations

According to the previous discussion the RGF curve must be defined as follows: given a set of N linear independent and orthonormal direction vectors, **S**, the RGF curve is composed of the set of points on the PES such that the gradient vector, expressed in the **S**-direction vectors, has only one nonzero component in one of these directions, which has to remain the same for all this set of points. The first question in the application of the RGF methodology is how to define the search direction normalized vector **r**. Because we are interested in finding the transition structure of an elementary reaction, by comparing the molecular geometry of both the reactant and product we observe the geometry parameters that present the largest difference between these two minima. Now, we make a perturbation of the molecular geometry, either of reactant or product structure, in the geometry parameters that present the largest observed difference [9, 10]. In the next step these parameters are ket fixed and the energy of the molecular system is minimized with respect to the rest of the molecular geometry parameters. The resulting point is taken as the first point of the RGF curve and the normalized gradient vector as the **r** vector. Note that the normalized **r** vector and the set of vectors $\mathbf{P}_{\mathrm{r}}$ correspond to the $\mathbf{x}_{\mathrm{r}}$ geometry variables or driving force variables and the $\mathbf{x}_{\mathrm{p}}$ equilibrium variables, respectively, used in Refs. [9, 10]. The second question concerns the way to integrate the RGF path. The integration is done by a quadratic stepwise procedure, in other words if $\mathbf{x}_0$ is a point on the RGF path, the energy variation is approximated by a quadratic Taylor expansion around $\mathbf{x}_0$:

$$\begin{aligned}\Delta \mathrm{E}^{(2)}(\mathbf{x}_0 + \Delta\mathbf{x}_0) &= \mathrm{E}^{(2)}(\mathbf{x}_0 + \Delta\mathbf{x}_0) - \mathbf{E}(\mathbf{x}_0) \\ &= \Delta\mathbf{x}_0^{\mathrm{T}}\mathbf{g}_0 + \frac{1}{2}\Delta\mathbf{x}_0^{\mathrm{T}}\mathbf{H}_0\Delta\mathbf{x}_0 \ , \qquad (27)\end{aligned}$$

where $\mathbf{g}_0$ and $\mathbf{H}_0$ are the gradient vector and Hessian matrix, respectively, at the point $\mathbf{x}_0$ and the $\Delta\mathbf{x}_0$ vector is defined as $\Delta\mathbf{x}_0 = \mathbf{x} - \mathbf{x}_0$. The quadratic approximation of the energy to integrate the RGF path is controlled by using the restricted-step technique [15]. The restricted step ensures that $(\Delta\mathbf{x}_0{}^{\mathrm{T}}\Delta\mathbf{x}_0)^{1/2} \leq R$, where R is the "trust radius" that characterizes the "trust region". The "trust region" is the region where the quadratic expansion of the energy is valid.

Given any point on the path $\mathbf{x}_0 = \mathbf{x}(t_0)$, the path itself can be represented as a Taylor series in t expanded about $\mathbf{x}_0$,

$$\begin{aligned}\mathbf{x}(t) &= \mathbf{x}(t_0) + \left.\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\right|_{t=t_0}\Delta t_0 + \frac{1}{2}\left.\frac{\mathrm{d}^2\mathbf{x}(t)}{\mathrm{d}t^2}\right|_{t=t_0}\Delta t_0^2 + \cdots \\ &= \mathbf{x}(t_0) + \mathbf{S}\left(\mathbf{b}\Delta t_0 + \frac{1}{2}\mathbf{c}\Delta t_0^2 + \cdots\right) \ , \qquad (28)\end{aligned}$$

where $\Delta t_0 = t - t_0$ and Eqs. (8) and (17) have been used. The path is specified at first and second order in Δt_0 by knowing the tangent and the curvature vectors respectively. This Taylor series representation of the path forms the basis of the two numerical integration algorithms presented next.

3.2 Description of first-order approximation

In this subsection we outline the proposed first-order approximation (FOA) algorithm.

1. Given a point $\mathbf{x}_0$, a "trust radius" $\mathbf{R}_0$ and a normalized vector **r**, which is the search direction, compute the N−1 orthonormalized vectors, $\mathbf{P}_{\mathrm{r}}$, orthogonal to the **r** vector. The orthonormalization procedure is carried out by using a Gram–Schmidt procedure by taking as initial N−1 independent vectors the columns of the unit matrix **I**. Store these N orthonormal basis vectors in a matrix $\mathbf{S} = [\mathbf{r}|\mathbf{P}_{\mathrm{r}}]$. Compute the energy, E_0, the gradient vector $\mathbf{g}_0$ and the Hessian matrix $\mathbf{H}_0$. Set $\mathrm{k} = 0$.
2. Project the Hessian matrix, $\mathbf{S}^{\mathrm{T}}\mathbf{H}_k\mathbf{S}$. Calculate the normalized tangent vector. Using Eq. (10) compute the $\mathbf{b}_{N-1}$ vector. Set $\mathbf{A}_k = \mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}_k\mathbf{P}_{\mathrm{r}}$; check the value of the $\det(\mathbf{A}_k)$. The b_1 component is evaluated by normalization, see Eq. (11). The Δt_0 parameter is computed as follows: $\mathrm{nwr} = -\mathbf{r}^{\mathrm{T}}\mathbf{g}_k/\{[\mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{r} - \mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{P}_{\mathrm{r}}\mathbf{A}_k{}^{-1}\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}_k\mathbf{r}]\mathrm{b}_1\}$ is the Newton–Raphson step following the RGF curve, then $\Delta t_0 = \mathrm{nwr}$ if $|\Delta t_0| = (\Delta\mathbf{x}_k{}^{\mathrm{T}}\Delta\mathbf{x}_k)^{1/2} < R_k$, otherwise $|\Delta t_0| = R_k$. If $[\mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{r} - \mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{P}_{\mathrm{r}}\mathbf{A}_k^{-1}\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}_k\mathbf{r}] < 0$ the sign of Δt_0 is the same sign as the nwr step otherwise take –nwr.
3. Ensure that the quadratic expansion of the energy is valid. The "trust radius" is modified, $R_k \to R_{k+1}$, according to some algorithm described elsewhere [9].
4. Compute the predictor step $\mathbf{x} = \mathbf{x}_k + \mathbf{S}\mathbf{b}\Delta t_0$. The arc length for this step is $|\Delta t_0|$. Compute the energy, E, and the gradient vector $\mathbf{g}(\mathbf{x})$ of this new predictor point. If $|\mathbf{g}(\mathbf{x})| \leq \varepsilon$ a stationary point has been found, stop. Otherwise update the Hessian matrix by using the Murtagh–Sargent–Powell (MSP) formula [16, 17], $\mathbf{H}_k \to \mathbf{H}$.

5. If $|\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{g}| > \varepsilon$ a correction step is needed, otherwise go to step 6. The correction step consists of a minimization of the energy with respect to the $\mathbf{b}_{N-1}$ vector components by fixing $b_1 = 0$. The minimization is carried out by using a standard quasi-Newton–Raphson algorithm [15], i.e., at each iteration the energy and the gradient vector are evaluated and the Hessian matrix is updated using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) formula [15]. At each Newton–Raphson step the projected Newton equation

$$\mathbf{0}_{N-1} = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{g} + \mathbf{A}\mathbf{b}_{N-1} \tag{29}$$

is solved. The minimization is converged when $|\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{g}| \leq \varepsilon$, obtaining a new point $\mathbf{x}$, energy E, gradient vector $\mathbf{g}$ and Hessian matrix $\mathbf{H}$.
6. Change the point $\mathbf{x} \rightarrow \mathbf{x}_{k+1}$, $E \rightarrow E_{k+1}$, $\mathbf{g} \rightarrow \mathbf{g}_{k+1}$, and $\mathbf{H} \rightarrow \mathbf{H}_{k+1}$. Set $k = k + 1$, go to step 2.

In the previously described algorithm the resulting updated Hessian matrix is checked by the Eckert and Werner procedure [18]. Normally the correction step is evaluated very few times, and it converges within two or three iterations. The BFGS formula is used in the correction step rather than the MSP one since the BFGS formula is much more efficient for minimization; however, in this case, the BFGS formula updates only the part of the Hessian matrix associated with the subspace defined by the set of column vectors collected in the $\mathbf{P}_{\mathrm{r}}$ matrix and, consequently, the inertia of the full Hessian matrix is preserved during the iterative process. The inertia of a matrix is the number of positive and negative eigenvalues. Two alternative algorithms to compute the correction step are given in the Appendix.

3.3 Description of second-order approximation

Because both algorithms are closely related, we only describe the steps of the second-order approximation (SOA) algorithm which are different with respect to the FOA algorithm already described. These are

3. Once the normalized tangent vector, $\mathbf{b}$, is evaluated, a new Hessian matrix is computed in the direction of this tangent. The matrix $\langle\mathbf{F}[\mathbf{x}(t)\mathbf{Sb}\rangle \approx (\mathbf{H}'_k - \mathbf{H}_k)/\delta$, where $\mathbf{H}'_k$ is the Hessian matrix evaluated at the point $\mathbf{x}'_k = \mathbf{x}_k + \mathbf{Sb}\delta$ and δ is a sufficiently small number. Compute the vector $\mathbf{d} = \mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\langle\mathbf{F}[\mathbf{x}(t)\mathbf{Sb}\rangle\mathbf{Sb}$. Evaluate the curvature vector using Eq. (19). The Δt_0 parameter is computed by solving the second-order equation: $0 = \mathbf{r}^{\mathrm{T}}\mathbf{g}_k/[\mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{r} - \mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{P}_{\mathrm{r}}\mathbf{A}_k{}^{-1}\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}_k\mathbf{r}]$ $+\mathrm{b}_1\Delta t_0 + 1/2\mathrm{c}_1(\Delta t_0)^2$. If $[(\Delta t_0)^2 + 1/4\mathbf{c}^{\mathrm{T}}\mathbf{c}(\Delta t_0)^4]^{1/2} = (\Delta\mathbf{x}_k{}^{\mathrm{T}}\ \Delta\mathbf{x}_k)^{1/2} < R_k$, we take this Δt_0 as the current parameter, otherwise the Δt_0 is evaluated by solving the equality $(\Delta t_0)^2 + 1/4\mathbf{c}^{\mathrm{T}}\mathbf{c}(\Delta t_0)^4 = R_k^2$. If $\mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{r} - \mathbf{r}^{\mathrm{T}}\mathbf{H}_k\mathbf{P}_{\mathrm{r}}\mathbf{A}_k{}^{-1}\mathbf{P}_{\mathrm{r}}{}^{\mathrm{T}}\mathbf{H}_k\mathbf{r} < 0$ the sign of Δt_0 is positive otherwise it takes a negative sign.
4. Compute the predictor step $\mathbf{x} = \mathbf{x}_k + \mathbf{S}[\mathbf{b}\Delta t_0 + 1/2\mathbf{c}(\Delta t_0)^2]$. The arc length for this step is $1/2\{(\mathbf{c}^{\mathrm{T}}\mathbf{c})^{1/2}\Delta t_0 (\mathbf{c}^{\mathrm{T}}\mathbf{c}(\Delta t_0)^2 + 1)^{1/2} + \log[(\mathbf{c}^{\mathrm{T}}\mathbf{c})^{1/2}\Delta t_0 + (\mathbf{c}^{\mathrm{T}}\mathbf{c}(\Delta t_0)^2 + 1)^{1/2}]\}$ $(\mathbf{c}^{\mathrm{T}}\mathbf{c})^{-1/2}$.

Using this algorithm, the number of iterations employed in the correction step is lower that in the first-order algorithm. The correct evaluation of the RGF curvature vector only requires the evaluation of a second Hessian matrix in the direction of the tangent vector. The TP is detected when $b_1 \approx 1$ and $|c_1| \approx 0$. At this point better performance is obtained if the predictor step is evaluated by the expression $\mathbf{x} = \mathbf{x}_k + \mathbf{r}[1 - 1/6\mathbf{c}^{\mathrm{T}}\mathbf{c}\ (\Delta t_0)^2]\ \Delta t_0 + 1/2\mathbf{P}_{\mathrm{r}}\mathbf{c}_{N-1}(\Delta t_0)^2$. With this predictor step, the corrector step converges with a few number of iterations. The term $-1/6\mathbf{c}^{\mathrm{T}}\mathbf{c}(\Delta t_0)^3$ is the $\mathbf{r}$ component of the $\mathrm{d}^3\mathbf{x}(t)/\mathrm{d}t^3$ vector at the TP. The TP problem is better handled by the SOA algorithm than with the FOA algorithm, because the FOA may be divergent in this situation. Finally the arc length formula for each step is derived from the expression

$$\begin{aligned}\Delta\mathrm{s}(\Delta t_0) &= \int_0^{\Delta t_0} \sqrt{\left(\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\right)^{\mathrm{T}}\left(\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}\right)}\mathrm{d}t \\ &= \int_0^{\Delta t_0} \sqrt{\left(1 + \mathbf{c}^{\mathrm{T}}\mathbf{c}(\Delta t)^2\right)}\mathrm{d}\Delta t\ ,\end{aligned} \tag{30}$$

where the first derivative of Eq. (28) with respect to t up to first order has been used.

4 Theoretical and computational comparison with GE methodology

First of all, we outline the basic theory and computational procedure of the GE method. According to Hoffmann et al. [13] and Sun and Ruedenberg [14] the GE are curves on the PES such that at each point of the path the square of the gradient norm is stationary with respect to the variations within the contour subspace, $E(\mathbf{x})$ = constant, passing through this point of the path. The gradient vector, $\mathbf{g}[\mathbf{x}(t)]$, associated to each point of the GE path should satisfy Eq. (26) that now can be rewritten as

$$\begin{aligned}\mathbf{k}[\mathbf{x}(t)] &= \left(\mathbf{I} - \frac{\mathbf{g}[\mathbf{x}(t)]\{\mathbf{g}[\mathbf{x}(t)]\}^{\mathrm{T}}}{\{\mathbf{g}[\mathbf{x}(t)]\}^{\mathrm{T}}\mathbf{g}[\mathbf{x}(t)]}\right)\mathbf{H}[\mathbf{x}(t)]\mathbf{g}[\mathbf{x}(t)] \\ &= \left(\mathbf{I} - \mathbf{r}\mathbf{r}^{\mathrm{T}}\right)\mathbf{H}[\mathbf{x}(t)]\mathbf{g}[\mathbf{x}(t)] = \mathbf{0}\ .\end{aligned} \tag{31}$$

Note that in GE, $\mathbf{g}[\mathbf{x}(t)] = \alpha\mathbf{r}$, where $\alpha = |\mathbf{g}[\mathbf{x}(t)]|$ and consequently Eq. (31) must be written as

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{g}[\mathbf{x}(t)] = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}\alpha = \mathbf{0}_{N-1}\ , \tag{32}$$

with

$$\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{g}[\mathbf{x}(t)] = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{r}\alpha = \mathbf{0}_{N-1}\ , \tag{33}$$

for $\alpha \neq 0$. The $\mathbf{S}$ matrix is built by the set of eigenvectors of the $\mathbf{H}[\mathbf{x}(t)]$ matrix. Using Eq. (31) Sun and Ruedenberg [14] showed that the GE path are curves which connect those points where the steepest-descent curves have zero curvature. This is because the curvature vector of any steepest descent–ascent curve line is $\mathbf{k}[\mathbf{x}(t)]/\{\{\mathbf{g}[\mathbf{x}(t)]\}^{\mathrm{T}}\mathbf{g}[\mathbf{x}(t)]\}$. In order to obtain the tangent vector of the GE path, we derive the implicit function given in Eq. (31) with respect to the t parameter

$$\mathbf{0} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{k}[\mathbf{x}(t)] = \nabla_{\mathbf{x}}^{\mathrm{T}}\mathbf{k}[\mathbf{x}(t)]\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} \quad . \tag{34}$$

After some algebraic manipulations Eq. (34) gives [14]

$$\mathbf{0} = \left(\mathrm{I} - \mathbf{r}\mathbf{r}^{\mathrm{T}}\right)\Big\{\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{r}\rangle\alpha + (\mathbf{H}[\mathbf{x}(t)])^2 \\ -\mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}\mathbf{H}[\mathbf{x}(t)]\Big\}\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} \quad , \tag{35}$$

where $\mathbf{F}[\mathbf{x}(t)]$ is the third-order energy derivative tensor with respect to the $\mathbf{x}$ variables evaluated at $\mathbf{x}(t)$. Multiplying Eq. (35) on the left by $\mathbf{S}^{\mathrm{T}}$ and substituting it into Eq. (8) as well as using Eqs. (32) and (33) the following equation is obtained:

$$\mathbf{0}_{N-1} = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{r}\rangle\mathbf{r}\,\alpha\, b_1 + \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\{\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{r}\rangle\alpha \\ + (\mathbf{H}[\mathbf{x}(t)])^2 - \mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}\mathbf{H}[\mathbf{x}(t)]\Big\}\mathbf{P}_{\mathrm{r}}\mathbf{b}_{N-1} \quad . \tag{36}$$

From Eq. (36) we obtain the $\mathbf{b}_{N-1}$ vector:

$$\mathbf{b}_{N-1} = -\{\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\Big\{\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{r}\rangle\alpha + (\mathbf{H}[\mathbf{x}(t)])^2 \\ - \mathbf{r}^{\mathrm{T}}\mathbf{H}[\mathbf{x}(t)]\mathbf{r}\mathbf{H}[\mathbf{x}(t)]\}\mathbf{P}_{\mathrm{r}}\}^{-1}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\langle\mathbf{F}[\mathbf{x}(t)]\mathbf{r}\rangle\mathbf{r}\,\alpha\, b_1 . \tag{37}$$

Consequently the GE normalized tangent vector in the $\mathbf{S}$-basis representation takes the form $\mathbf{b}^{\mathrm{T}} = (b_1,\ \mathbf{b}_{N-1}{}^{\mathrm{T}})$, where the b_1 component is determined by normalization. It is interesting to compare this equation with Eq. (11). The GE predictor step is evaluated as $\mathbf{x} = \mathbf{x}_0 + \mathbf{S}\mathbf{b}\Delta t_0$. Normally $\mathbf{x}$ lies slightly off the GE path, so a corrector step is needed to search for the new point on the GE path [14]. From a mathematical point of view both the GE and RGF paths are defined by continuous implicit functions, $\Phi[\mathbf{x}(t)] = \mathbf{0}$; however, the RGF implicit function, Eq. (2), is much simpler that the GE implicit function, Eq. (31). This difference in simplicity is reflected in the computational evaluation of the tangent vector. Whereas in the RGF path the tangent vector is a function of the Hessian matrix, in the GE path it involves both the Hessian matrix and the third-order energy derivatives with respect to the geometry parameters. Note that in both cases the tangent vector is obtained from an equation like $\{\nabla_{\mathbf{x}}^{\mathrm{T}}\Phi[\mathbf{x}(t)]\}(\mathrm{d}\mathbf{x}(t)/\mathrm{d}t) = \mathbf{0}$. See Eqs. (4) and (35) for the RGF path and GE path, respectively. Also the corrector step is much easier to compute in the RGF method than in the GE path. In the RGF method the corrector step involves a restricted Newton–Raphson procedure. In summary, since the RGF method is easy to evaluate and for a selected $\mathbf{r}$ vector describes a path that connects the minimum with the transition state, it becomes an important tool to locate first-order saddle points.

5 Analysis, examples and discussion

In order to illustrate the algorithms presented previously, an application to the Müller–Brown surface [19] was performed. This is a two-dimensional surface that has three minima and two first-order saddle points and is described by the equation

$$\mathrm{E}(x,y) = \sum_{i=1}^{4} A_i \exp\Big[a_i\big(x - x_i^0\big)^2 + b_i\big(x - x_i^0\big)\big(y - y_i^0\big) \\ + c_i\big(y - y_i^0\big)^2\Big] \quad , \tag{38}$$

where $\mathbf{A}^{\mathrm{T}} = (-200, -100, -170, 15)$, $\mathbf{a}^{\mathrm{T}} = (-1, -1, -6.5, 0.7)$, $\mathbf{b}^{\mathrm{T}} = (0,\ 0,\ 11,\ 0.6)$, $\mathbf{c}^{\mathrm{T}} = (-10,\ -10,\ -6.5,\ 0.7)$, $\mathbf{x}_0^{\mathrm{T}} = (1, 0, -0.5, -1)$, and $\boldsymbol{y}_0^{\mathrm{T}} = (0, 0.5, 1.5, 1)$. We will focus the present study on the path from the minimum at (–0.558, 1.442) to the saddle point at (–0.822, 0.624) as this is the most challenging path owing to its curvature. The results reported of this section were obtained with the Mathematica program and the implementation differs in two points from the one described earlier. Firstly, exact analytic Hessian matrices were used all the time. Secondly, a fixed trust radius was used, without rescaling. The reason for such differences is that the aim of this section is to compare the algorithms emphasizing only the differences in the way the predictor and corrector are calculated.

To start, different choices for the vector $\mathbf{r}$ are possible. A sensible one is to choose it as the direction connecting the two minima. This relates the present method to the one previously described in Ref. [9]. Another simple guess would be to choose the direction towards the TS, but the TS is precisely what one usually tries to locate. This might be the optimal choice, but the path cannot be known a priori. Other possible choices are $\mathbf{r} = (1, 0)$ and $\mathbf{r} = (0,\ 1)$. These possibilities make the RGF curve equivalent to the primitive definitions where only one of the components of the gradient vector was forced to be zero [8]. Although they are quite meaningless because the basis that expresses the gradient vector is arbitrary, they converge to the desired TS because it is the only first-order saddle point surrounding the minimum. Anyway, they present interesting features, such as a TP. Different RGF curves, according to the choice of the $\mathbf{r}$ vector, are shown in Fig. 1. One clearly sees that considerably fewer points are needed to get to the TS if $\mathbf{r}$ is chosen cleverly; however, because the present objective is to test the algorithms in all possible situations, the choice of $\mathbf{r} = (0,\ 1)$ was made. Such a value of $\mathbf{r}$ leads to a RGF curve with a TP.

5.1 Behavior of the FOA algorithm. Comparing corrector steps

The RGF curve associated with $\mathbf{r} = (0,\ 1)$ present a TP. This TP is located at (–1.113, 0.637) and corresponds to zero change in the $\mathbf{P}_{\mathrm{r}}$ coordinates in the predictor step, i.e., $\mathbf{b}_{N-1} = \mathbf{0}_{N-1}$. This comes from the zero vector of $\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{r} = \mathbf{0}_{N-1}$ of Eq. (10) and can be checked during the iterative procedure. A second interesting point occurs when $\det(\mathbf{A}) = 0$. Then, following Eq. (14), $b_1 = 0$. This point is located at (–1.038, 0.566) and also shows the importance of the corrector step for the FOA algorithm. According to the different properties of these types of

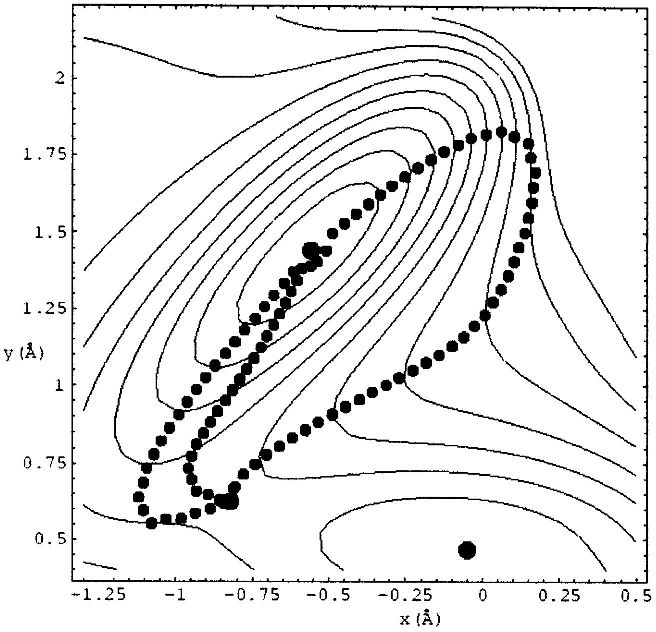


Fig. 1. Different reduced-gradient-following paths from the minimum located at (–0.558, 1.442) to the transition state located at (–0.822, 0.624) of the Müller–Brown surface [19]. From the right, $\mathbf{r} = (0,\ 1)$ (which is very similar to an **r** ponting to the transition state), **r** pointing to the following minima, and $\mathbf{r} = (1, 0)$. The coordinates can be assumed to be in angstroms and the energy contour lines, which are the *solid lines*, to be in millihartrees

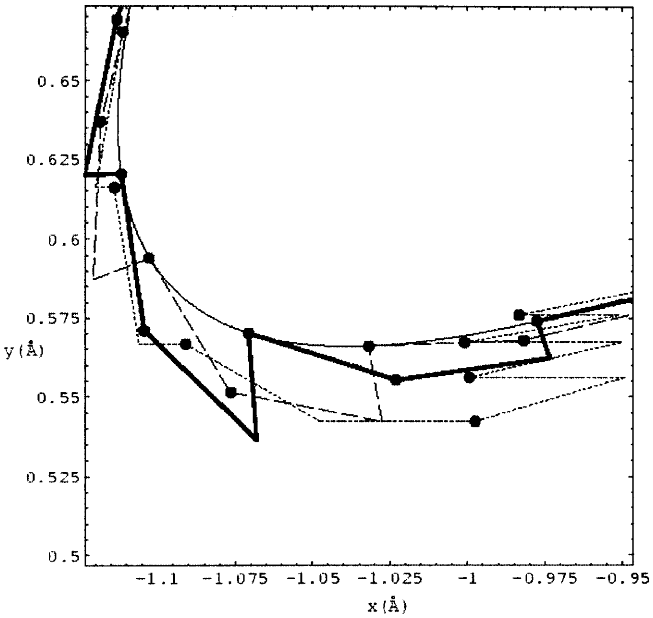


Fig. 2. First-order approximation algorithm with different corrector steps for the most curved segment of the path. The exact path is also shown (*solid line*). *Dotted line*: algorithm described in the main text. *Bold line*: first algorithm in the Appendix. *Dashed line*: second algorithm in the Appendix. The region displayed corresponds to the area around the point (–1.038, 0.566) of the Müller–Brown surface [19]

points, Δt_0 can be easily modified to make the appropriated step.

The behavior of the corrector steps is best displayed in the area of the TP and the point where $\det(\mathbf{A}) = 0$ (Fig. 2.) We used a deliberately large trust radius to force the need of a considerably large corrector step. This is also achieved by using a large threshold for the switch of the corrector step so that the points deviate considerable from the exact trajectory. From Fig. 2 several points can be made. The two corrector step algorithms described in the Appendix posses considerably better performance than the corrector step algorithm described in the main text; however, all these algorithms are iterative and therefore computationally time consuming. The main difference between them shows up near the point where $\det(\mathbf{A}) = 0$. There the predictor step has $b_1 \approx 0$ and a corrector should clearly have $b_1 \neq 0$. The corrector step algorithm described in the main text, however, is forced to optimize in the $\mathbf{P}_\mathrm{r}$ subspace only, having $b_1 = 0$ and showing a poor performance. On the other hand, near a TP, i.e., a point where $\mathbf{P}_\mathrm{r}^\mathrm{T}\mathbf{H}\mathbf{r} = \mathbf{0}_{N-1}$, the three corrector step algorithms behave almost indistinguishably. In any case, one has to bear in mind that the corrector step will usually take place in an N–1 dimensional subspace, which will usually be of dimension much larger than 1. Then, the restriction of $b_1 = 0$, corresponding to the algorithm described in the main text, should be less dramatic than for this two-dimensional surface, and, in fact, the three corrector step algorithms behave similarly for chemical reactions such as the one explained later. For the sake of completeness, we should say that in the iterative corrector step proposed, the first iteration plays the main role, so convergence is normally achieved with a few iterations.

5.2 Behavior of the SOA algorithm

The behavior of the SOA algorithm is shown in Fig. 3. This algorithm takes into account the curvature at each point and consequently is much more effective. To emphasize this fact, the points in Fig. 3 were calculated without a corrector step. In fact, as can be seen, the curvature vector enters the step in a second-order term that is very similar to the corrector step, simply because the curvature vector is orthogonal to the tangent vector appearing in the first-order term. However, a main difference is that Δt_0 includes first- and second-order terms, making the calculation of the total path length much more precise.

Although from these examples one could infer that the FOA and SOA algorithms have similar costs, this is not the case. The corrector step is only needed in strong curvature regions such as the one plotted, and therefore the FOA is faster than the SOA for the whole path.

5.3 A chemical reaction: the ring opening of cyclopropyl radical rearrangement

The two RGF algorithms described in Sect. 4 were implemented in a modified version of the MOPAC program [21, 22]. The transition-state searches of all the chemical reactions described here were calculated using the corresponding wavefunction and the Austin model 1 [23] semiempirical Hamiltonian. The convergence criteria were on both the maximum component of the gradient vector, 10^{-3} kcalmol^{-1}Å^{-1} or 10^{-3} kcalmol^{-1}rad^{-1}, and the step length, 10^{-1} Å or 10^{-1} rad. The correction step was computed using the first algorithm described in the Appendix.

The ring opening of cyclopropyl radical rearrangement was studied by Olivella et al. [24] at the ab initio level of theory. We present the behavior of both the FOA and the SOA algorithms to locate the corresponding transition state. The molecular geometry parameters of the reactant, the slightly perturbed reactant, and the transition state located are shown in Fig. 4. The perturbed geometry parameters are the bond angle and the dihedral angle $H_4C_1C_2C_3$. The normalized gradient vector of the slightly perturbed reactant geometry corresponds to the vector **r**. The components of **r** are given in Table 1. The behavior of the FOA and SOA algorithms are shown in Tables 2 and 3, respectively. The FOA algorithm needs two steps more in order to converge to the first-order saddle point and it needs more iterations in the correction step than the SOA algorithm. The total arc length of the path from the slightly perturbed geometry to the transition state is 0.67 Å – rad using the FOA method and 0.80 Å – rad using the SOA algorithm. Note that in the FOA method the step arc length is $|\Delta t_0|$. This difference in the total arc length is due to the fact that in the FOA algorithm the arc length is evaluated as a straight line, whereas in the

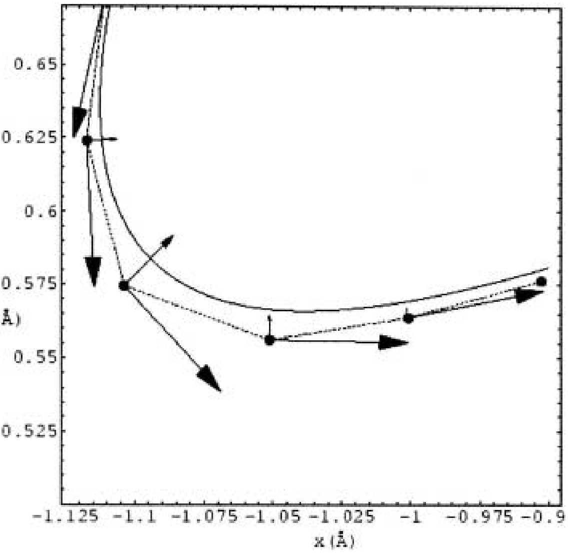
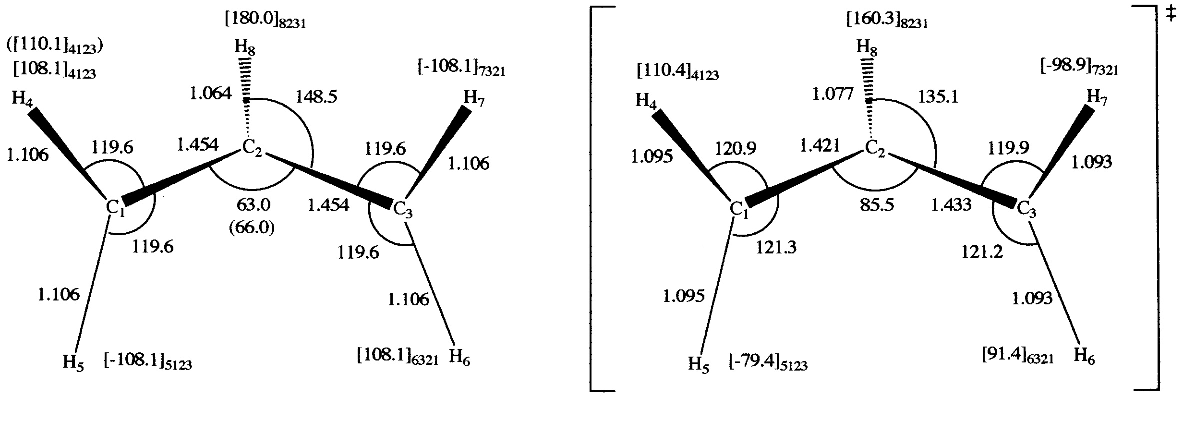


Fig. 3. Second-order approximation algorithm. For each point the (scaled) tangent and curvature vectors are plotted. No corrector step was used to calculate these points. The *longer arrows* are the tangent vectors, the *shorter arrows* are the curvature vectors. The **r** vector is (0, 1). The region displayed corresponds to the area around the point (−1.038, 0566) of the Müller–Brown surface [19]

Table 1. The components of the normalized **r** vector. The components are in internal coordinates, distances, bond angles and dihedral angles. The units of the **r** vector components are kcalmol^{-1} Å^{-1} for distances and kcalmol^{-1} rad^{-1} for bond angles and dihedrals. The atom numbering is that given in Fig. 4

Component	Value
C_2C_1	6.30×10^{-4}
C_3C_2	3.70×10^{-4}
$C_3C_2C_1$	9.95×10^{-1}
H_4C_1	0.00
$H_4C_1C_2$	-1.30×10^{-4}
$H_4C_1C_2C_3$	1.05×10^{-1}
H_5C_1	-1.00×10^{-4}
$H_5C_1C_2$	-1.00×10^{-4}
$H_5C_1C_2C_3$	1.20×10^{-4}
H_6C_3	0.00
$H_6C_3C_2$	-1.10×10^{-4}
$H_6C_3C_2C_1$	0.00
H_7C_3	0.00
$H_7C_3C_2$	0.00
$H_7C_3C_2C_1$	3.60×10^{-4}
H_8C_2	-1.70×10^{-4}
$H_8C_2C_3$	0.00
$H_8C_2C_3C_1$	0.00



Reactant

Transition State

Fig. 4. The geometry parameters of the reactant and transition state for the ring opening of cyclopropyl radical rearrangement. The perturbation of the selected geometry parameters is in *parantheses*. The bond lengths are given in angstroms and the bond and dihedral angles are in degrees

SOA algorithm it is computed in a nonlinear way. Thus, in SOA algorithm the total arc length will be much longer that in the FOA algorithm. By evaluating the steepest-descent path form the transition sate to the minimum associated with the cyclopropyl radical, the total arc length is 1.42 Å – rad. This difference with respect to the RGF path clearly shows that both paths are different; however, both connect the minimum with the first-order saddle point. In this RGF path no TP and VRI points were found. This example supports the proposition of Quapp et al. [25] which states that if a RGF curve directly connects two stationary points without crossing a VRI point, the indices of the stationary point differ by 1.

6 Conclusions

We have presented a theoretical comparative study between the RGF curve and the GE curve. We have also proposed two different algorithms to integrate the RGF curve: both are based on the interplay of a predictor with a corrector step.

Table 2. Behavior of the first-order approximation algorithm to follow the reduced-gradient-following path for the ring opening cyclopropyl radical rearrangement

Step	$\mathbf{r}^T\mathbf{g}[\mathbf{x}(t)]$[a]	Energy[b]	Δt_0[c]	Iterations[d]
1	38.7	60.3	5.0×10^{-2}	2
2	54.4	61.7	7.1×10^{-2}	2
3	71.3	64.4	1.0×10^{-1}	2
4	84.9	69.1	1.4×10^{-1}	2
5	87.3	76.2	2.0×10^{-1}	4
6	28.3	83.3	1.0×10^{-1}	2
7	1.83	83.7	1.0×10^{-2}	0
8	-1.0×10^{-1}	83.7	-5.6×10^{-4}	0
9	5.9×10^{-3}	83.7	3.2×10^{-5}	0
10	3.6×10^{-3}	83.7	1.9×10^{-5}	0
11	1.8×10^{-3}	83.7	9.0×10^{-6}	0
12	-1.2×10^{-3}	83.7	6.0×10^{-6}	0

[a] The **r** component of the gradient vector in kcalmol^{-1} Å^{-1} – kcalmol^{-1} rad^{-1}
[b] In kcalmol^{-1}
[c] In Å – rad
[d] Number of iterations employed in the step correction

Acknowledgements. We are indebted to S. Olivella for valuable suggestions. This research was supported by the Spanish DGICYT (grants PB98-1240-C02-01 and PB98-1240-C02-02).

Alternative algorithms to compute the corrector step

In this appendix we report two algorithms for the corrector step. After the predictor step is evaluated, we assume that the new **x** point lies slightly off the RGF curve and we search for a point on the RGF curve, **x**', which is as close as possible to **x**. We assume that the first-order Taylor expansion of the RGF function, Φ(**x**, **r**) = **r**−**g**(**x**)/|**g**(**x**)|, Eq. (2), around the point **x**, which is off the RGF curve, is still valid at the point **x**′, which is on the RGF curve,

$$\begin{aligned}\mathbf{0} &= \Phi(\mathbf{x}',\mathbf{r}) = \Phi(\mathbf{x},\mathbf{r}) + \nabla_{\mathbf{x}}^{T}\Phi(\mathbf{x},\mathbf{r})\Delta\mathbf{x} \\ &= \mathbf{r} - \frac{\mathbf{g}(\mathbf{x})}{|\mathbf{g}(\mathbf{x})|} - \frac{1}{|\mathbf{g}(\mathbf{x})|}\left(\mathbf{I} - \frac{\mathbf{g}(\mathbf{x})[\mathbf{g}(\mathbf{x})]^{T}}{|\mathbf{g}(\mathbf{x})|^{2}}\right)\mathbf{H}(\mathbf{x})\Delta\mathbf{x} \quad ,\end{aligned} \tag{A1}$$

where Δ**x** = **x**′−**x**. We remark that $\nabla_{\mathbf{x}}^{T}\Phi(\mathbf{x}, \mathbf{r})$ is the first derivative of Φ(**x**, **r**) = **r**−**g**(**x**)/|**g**(**x**)| at **x**, a point that is off the RGF curve. If we build a $\mathbf{P}_r{}'$ matrix such that $(\mathbf{P}_r{}')^T\mathbf{g}(\mathbf{x})/|\mathbf{g}(\mathbf{x})| = \mathbf{0}_{N-1}$, then multiplying from the left of Eq. (A1) by $(\mathbf{P}_r{}')^T$ we get

$$\mathbf{0}_{N-1} = \left(\mathbf{P}_r'\right)^T\mathbf{r}|\mathbf{g}(\mathbf{x})| - \left(\mathbf{P}_r'\right)^T\mathbf{H}(\mathbf{x})\Delta\mathbf{x} = \mathbf{q}(\Delta\mathbf{x}) \quad . \tag{A2}$$

Note that the $\mathbf{P}_r{}'$ matrix depends on the point and is evaluated at the point **x**. Equation (A2) is a linear equation of the **x**′ variables and represents the straight-line approximation to the RGF curve near **x**. Following Sun and Ruedenberg [14] we determine the shortest distance between the point **x** and the straight line given in Eq. (A2) by using the following Lagrange multiplier method:

$$L(\Delta\mathbf{x}, \boldsymbol{\lambda}) = \frac{1}{2}\Delta\mathbf{x}^T\Delta\mathbf{x} - \boldsymbol{\lambda}^T\mathbf{q}(\Delta\mathbf{x}) \quad , \tag{A3}$$

Table 3. Behavior of the second-order approximation algorithm to follow the reduced-gradient-following path for the ring opening cyclopropyl radical rearrangement

Step	$\mathbf{r}^T\mathbf{g}[\mathbf{x}(t)]$[a]	Energy[b]	Δt_0[c]	$\Delta s(t_0)$[d]	Curvature[e]	Iterations[f]
1	38.7	60.3	5.0×10^{-2}	5.0×10^{-2}	1.255	1
2	54.4	61.7	7.1×10^{-2}	7.1×10^{-2}	1.330	1
3	71.3	64.4	1.0×10^{-1}	1.0×10^{-1}	1.438	2
4	84.7	69.0	1.4×10^{-1}	1.4×10^{-1}	1.537	2
5	87.6	75.7	1.9×10^{-1}	2.0×10^{-1}	2.935	2
6	69.6	81.8	2.0×10^{-1}	2.0×10^{-1}	0.961	2
7	7.80	83.7	3.6×10^{-2}	3.7×10^{-2}	3.125	1
8	-1.6×10^{-1}	83.7	-8.4×10^{-4}	8.4×10^{-4}	2.940	0
9	1.2×10^{-2}	83.7	6.5×10^{-5}	6.5×10^{-5}	2.930	0
10	1.2×10^{-3}	83.7	5.0×10^{-6}	5.0×10^{-6}	2.930	0

[a] The **r** component of the gradient vector in kcalmol^{-1} Å^{-1} – kcalmol^{-1} rad^{-1}
[b] In kcalmol^{-1}
[c] In Å – rad
[d] The arc length for this step in Å – rad
[e] The curvature is defined as $(\mathbf{c}^T\mathbf{c})^{1/2}$, where **c** is the curvature vector defined in Eq. (18)
[f] Number of iterations employed in the step correction

where the $\mathbf{q}(\Delta\mathbf{x})$ vector is defined in Eq. (A2). Note that the $\boldsymbol{\lambda}$ vector has the dimension $N-1$. The solution of Eq. (A3) is

$$\Delta\mathbf{x} = \mathbf{H}(\mathbf{x})\mathbf{P}'_{\mathrm{r}}\left\{\left(\mathbf{P}'_{\mathrm{r}}\right)^{\mathrm{T}}[\mathbf{H}(\mathbf{x})]^{2}\mathbf{P}'_{\mathrm{r}}\right\}^{-1}\left(\mathbf{P}'_{\mathrm{r}}\right)^{\mathrm{T}}\mathbf{r}|\mathbf{g}(\mathbf{x})| \ . \qquad (A4)$$

By an iterative procedure we find a point such that $(\mathbf{P}_{\mathrm{r}}')^{\mathrm{T}}\mathbf{r} = \mathbf{0}_{N-1}$ and consequently $\Delta\mathbf{x} = \mathbf{0}$. This point is on the RGF curve since the $\mathbf{P}_{\mathrm{r}}'$ matrix is orthogonal to the normalized $\mathbf{r}$ vector that characterizes the RGF curve. By using either the FOA algorithm or the SOA algorithm this correction step procedure normally converges in two or three iterations.

The second algorithm is based on Eq. (29) where the component b_1 is not forced to be zero. We rewrite Eq. (29) as

$$\mathbf{0}_{N-1} = \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{g} + \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{r}b_1 + \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{P}_{\mathrm{r}}\mathbf{b}_{N-1} \ . \qquad (A5)$$

Equation (A5) represents a system of $N-1$ equations with N variables. From Eq. (A5) we obtain

$$\mathbf{b}_{N-1} = -\left(\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{P}_{\mathrm{r}}\right)^{-1}\left(\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{g} + \mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{r}b_1\right) \ . \qquad (A6)$$

Now the distance $1/2\Delta\mathbf{x}^{\mathrm{T}}\Delta\mathbf{x} = 1/2\mathbf{b}^{\mathrm{T}}\mathbf{b} = 1/2(b_1^2 + \mathbf{b}_{N-1}{}^{\mathrm{T}}\mathbf{b}_{N-1})$ is minimized with respect to b_1,

$$\frac{\mathrm{d}}{\mathrm{d}b_1^2}\left(\frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}}\Delta\mathbf{x}\right) = 0 \ . \qquad (A7)$$

Note that $\Delta\mathbf{x} = \mathbf{S}\mathbf{b} = [\mathbf{r}|\mathbf{P}_{\mathrm{r}}]\mathbf{b}$ has been used. On substituting Eq. (A6) into Eq. (A7) after some straighforward algebra we get

$$b_1 = -\frac{\mathbf{r}^{\mathrm{T}}\mathbf{H}\mathbf{P}_{\mathrm{r}}\left(\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{P}_{\mathrm{r}}\right)^{-2}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{g}}{1 + \mathbf{r}^{\mathrm{T}}\mathbf{H}\mathbf{P}_{\mathrm{r}}\left(\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{P}_{\mathrm{r}}\right)^{-2}\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{H}\mathbf{r}} \ . \qquad (A8)$$

Equations (A6) and (A8) are the basic equations for this algorithm. At the iteration such that $\mathbf{P}_{\mathrm{r}}^{\mathrm{T}}\mathbf{g} = \mathbf{0}_{N-1}$, $b_1 = 0$ and consequently form Eq. (A6) $\mathbf{b}_{N-1} = \mathbf{0}_{N-1}$ and the finding point is on the RGF curve. This algorithm also needs few iterations to converge.

References

1. Schlegel HB (1994) In: Yarkony DR (ed) Modern electronic structure theory. World Scientific, Singapore, p 459
2. Quapp W, Imig O, Heidrich D (1995) In: Heidrich D (ed) The reaction path in chemistry: current approaches and perspectives. Kluwer, Dordrecht, p 137
3. Miller WH, Handy NC, Adams JE (1980) J Chem Phys 72: 99
4. González J, Giménez X, Bofill JM (2001) J Phys Chem A 105: 5022
5. (a) Fukui K (1970) J Phys Chem 74: 4161; (b) Truhlar DG, Kuppermann A (1971) J Am Chem Soc 93: 1840
6. Quapp W, Hirsch M, Imig O, Heidrich D (1998) J Comput Chem 19: 1087
7. (a) Rothman MJ, Lohr LL Jr, Ewig CS, van Wazer JR (1981) In: Truhlar DG (ed) Potential energy surfaces and dynamics calculations. Plenum, New York, p 653; (b) Williams IH, Maggiora GM (1982) J Mol Struct (THEOCHEM) 89: 365
8. Quapp W, Hirsch M, Heidrich D (1998) Theor Chem Acc 100: 285
9. Anglada JM, Besalú E, Bofill JM, Crehuet R (2001) J Comput Chem 22: 387
10. Bofill JM, Anglada JM (2001) Theor Chem Acc 105: 463
11. Basilevsky MV, Shamov AG (1981) Chem Phys 60: 347
12. Hestenes MR (1980) Conjugate directions methods in optimization. Springer, Berlin Heidelberg New York
13. Hoffman DK, Nord RS, Ruedenberg K (1986) Theor Chim Acta 69: 265
14. Sun JQ, Ruedenberg K (1993) J Chem Phys 98: 9707
15. Fletcher R (1987) Practical methods of optimization. Wiley, Chichester
16. Bofill JM (1994) J Comput Chem 15: 1
17. Bofill JM, Comajuan M (1995) J Comput Chem 16: 1326
18. Eckert F, Werner H-J (1998) Theor Chem Acc 100: 21
19. Müller K, Brown L (1979) Theor Chim Acta 53: 75
20. Wolfram S (1988) Mathematica. Addison–Wesley, Redwood City, Calif, and associated computer programs
21. Stewart JJP (1983) QCPE Bull 3: 101
22. Olivella S (1984) QCPE Bull 4: 109
23. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107: 3902
24. Olivella S, Solé A, Bofill JM (1990) J Am Chem Soc 112: 2160
25. Hirsch M, Quapp W, Heidrich D (1999) Phys Chem Chem Phys 1: 5291